Isolation of recombinant plasmids bearing cDNA to hen ovomucoid and lysozyme mRNAs.

G N Buell, M P Wickens, J Carbon and R T Schimke
A large library of hen oviduct cDNA-pCR1 recombinant plasmids has been established in Escherichia coli χ1776. From this library, ovomucoid cDNA and lysozyme cDNA-bearing plasmids have been identified. One of these plasmids, pMu7, yielded the sequence of the 3' untranslated region of ovomucoid mRNA.

Estrogen administration to immature chicks results in cytodifferentiation of the oviduct (1) and synthesis of egg-white proteins (2). The synthesis of at least three of these proteins, conalbumin, ovalbumin, and ovomucoid, is regulated by the levels of the mRNAs encoding them as measured by in vitro translation and hybridization to cDNA (3-5). The mechanism of steroid-induced mRNA accumulation is not understood. The possibility that this induction requires exact nucleotide sequences in the induced mRNAs is an area of active investigation. This report describes the establishment of a large library of recombinant plasmids bearing cDNA to hen oviduct poly(A)-containing RNA and the isolation of plasmids bearing sequences complementary to ovomucoid mRNA and lysozyme mRNA.

EXPERIMENTAL PROCEDURES

Materials—HindIII and alkaline phosphatase were generously donated by Dr. C. Yanofsky, Department of Biological Sciences, Stanford University. Terminal transferase was from Dr. Winston Salszer, Department of Biology, University of California at Los Angeles. Other restriction endonucleases and polynucleotide kinase were purchased from New England Biolabs. DNA sequencing reagents were obtained from other sources (9). DNA bands were eluted electrophoretically, as described by Maniatis et al. (17) using the TBE (90 mM Tris, 90 mM NaCl, 1 mM EDTA, pH 8.0, 1 mM MgCl₂, 10 mM NaCl, 10 mM β-mercaptoethanol, and 100 μg of DNA per ml with the exception of EcoRI which required 50 mM NaCl. All restriction digests were precipitated with ethanol and washed with 70% ethanol before electrophoresis.

Terminal Transferase Reaction, Annealing, and Transformation—Terminal transferase homopolymer addition (without prior exonuclease digestion) was carried out as described (9). Poly(dA) was added to cDNA and poly(T) to EcoRI-cleaved pCR1 (pCR1 was the only plasmid vector certified EK-2 by the National Institutes of Health at the time these studies were initiated.) For poly(dA) addition, incorporation was linear for at least 15 min. The yield of double-stranded cDNA was 1.5%.

For poly(A)-containing RNA and Synthesis of Double-stranded cDNA—Hen oviduct poly(A)-containing RNA was prepared from polyacrylamide gel electrophoresis and ethidium bromide-purified from chromosomal DNA, we used the SDS lysis procedure of Guerry et al. (13) followed by ethidium bromide-cesium chloride equilibrium density gradient centrifugation (14). When less purified DNA was required as in screening by restriction (see Fig. 3), the SDS lysis protocol was used without density banding.

cDNA Synthesis and Solution Hybridization—Single-stranded cDNA was synthesized as described (7) except for the inclusion of actinomycin (100 μg/ml). cDNA was labeled with both [³²P]dCTP (22 Ci/mmol) and [α-³²P]dATP and polynucleotide kinase (11). Specific activity of the product was about 10⁹ cpm/μg. End labeled RNA fragments were used to detect colonies of interest by colony filter hybridization. The procedure of Grunstein and Hogness (12) was used with some alterations. Details have been published previously (9). Each filter (7.5 × 11 cm) had 96 colonies and was incubated with 10⁶ cpm labeled RNA for 24 h. Including 0.02% SDS during the hybridization helped to lower background. Nonspecifically bound probe was hydrolyzed with pancreatic ribonuclease (2 mg/filter) at 37°C for 1 h in 2× SSC.

Preparation of Plasmid DNA—For plasmid DNA substantially purified from chromosomal DNA, we used the S2D lysis procedure of Guerry et al. (13) followed by ethidium bromide-cesium chloride equilibrium density gradient centrifugation (14). When less purified DNA was required as in screening by restriction (see Fig. 3), the SDS lysis protocol was used without density banding.

DNA Molecular Weight Markers—The lengths of the HindIII-EcoRI double digestion products of A-DNA are 21.7, 5.15, 5.00, 4.27,
Recombinant Plasmids Bearing cDNA to Ovomucoid and Lysozyme mRNAs

In previous communications (7, 8) we have discussed the lengths of the four prominent mRNAs in differentiated chick oviduct and the requirements for their use as templates for double-stranded cDNA. Based on those results we have constructed a library of recombinant DNA-harboring bacterial colonies in E. coli χ1776. These colonies contain DNA which is complementary to hen oviduct poly(A)-containing RNA inserted into pCR1. As described under “Experimental Procedures,” pCR1 DNA was cleaved with EcoRI to generate linear DNA and “tailed” with poly(T) using terminal transferase. SI nuclease-treated, double-stranded cDNA was tailed with poly(dA). These preparations of DNA were annealed and used to transform CaC12-treated E. coli χ1776 recipient cells. Analysis of the 4300 kanamycin-resistant colonies has shown approximately 80% to have recombinant plasmid DNA. We wanted to screen these colonies for sequences complementary to conalbumin, ovomucoid, and lysozyme mRNAs, without having to purify the mRNAs for use as hybridization probes. (We have previously described the characterization of a plasmid which contains sequence complementary to ovalbumin mRNA (9).) At our disposal was a monospecific anti-conalbumin serum. We expected that even minor degradation of poly(A)-containing RNA in CH3HgOH/agarose. Poly(A)-containing RNA (10 μg), prepared as described previously (7), was electrophoresed in 20 mm CH3HgOH, 1.5% agarose, Lane 3. Other lanes display the molecular weight markers employed, total HeLa RNA in Lane 1, E. coli rRNA in Lane 2, and rabbit globin mRNA (plus contaminating rRNA) in Lane 3. labeled RNAs in probe-excess hybridizations. However, we also expected that, whereas large mRNA sequences might contaminate small mRNAs after electrophoresis in CH3HgOH, the converse would not be true. In Fig. 2 are shown the results of four replica-platings of 96 colonies from the library onto nitrocellulose filters. Colonies were grown and the DNA affixed as stated under “Experimental Procedures.” Each filter was incubated with one of the four end-labeled RNAs from Fig. 1. As expected, all of the colonies which registered positively with the ovalbumin mRNA probe are also positive with size-fractionated ovomucoid and lysozyme probes. Ovalbumin mRNA must therefore contaminate the smaller mRNAs.

The ovomucoid probe, however, also labels three additional colonies (D6, E3, and F12) not labeled by the ovalbumin probe. In the same manner the lysozyme probe labels two colonies (B7 and G6) which larger mRNA probes fail to detect. The conalbumin mRNA probe scores two colonies (D10 and H5) which are negative for all other probes. It appears that insignificant amounts of conalbumin mRNA sequences contaminate the other probes and that contamination by ovumucoid mRNA is the major concern in this experiment.

Using this methodology we screened 1600 colonies (four replicas each). This group yielded 21 conalbumin candidates, 67 ovomucoid candidates, and 28 lysozyme candidates.

Second Screening

 Cultures (100 ml) were grown for 12 members of each of the three sets. About 4 μg of plasmid DNA was obtained from each culture. The DNA was restricted with Hpa II endonuclease and divided into two equal aliquots. Hpa II sites, near the EcoRI site in pCR1, have previously been mapped (9). One DNA aliquot for each of the ovomucoid and lysozyme candidates was prepared for hybridization to cDNA by dena-
Fig. 2. Colony hybridization analysis of hen oviduct total cDNA bank in pCR1. The independent isolates were replicated onto four nitrocellulose filters. Treatment of the four identical filters is described under “Experimental Procedures.” Each filter was incubated with one of the four 5'-phosphorylated mRNA hybridization probes, noted at the right of each filter. Arrows point to those positively registering colonies which are unique to each panel.

We had initially hoped to carry out quantitative solution hybridizations with the same probes used in the colony filter hybridizations. However, control experiments in which end-labeled ovalbumin mRNA fragments were hybridized with a full-length ovalbumin cDNA-bearing plasmid (pOvE12) showed that a large percentage of the RNA would not hybridize (data not shown). Undoubtedly, some RNA products of partial alkali hydrolysis are too small to form stable hybrids. Therefore, we synthesized cDNA to ovomucoid and lysozyme mRNA preparations which had been isolated by sucrose gradient centrifugation (7). cDNA products were electrophoresed in alkaline agarose and the full length transcripts were eluted from the appropriate gel slice (see “Experimental Procedures”). The Hpa II-digested plasmid DNAs were hybridized overnight with the appropriate probe and hybrids were scored by S1 nuclease digestion (Table I). Most of the candidates in each set hybridized with 40% or more of the probe DNA with at least one member hybridizing to more than 50%. Hybridization of the probes to pOvE12 indicated the level of ovalbumin cDNA probe contamination.

The other DNA aliquots were electrophoresed in TBE-agarose (ovomucoid candidates) or TBE-acrylamide (lysozyme candidates). Some of these are shown in Fig. 3. The EcoRI site in pCR1 resides in the largest Hpa II fragment which migrates as a doublet at 830 bp (9). Hpa II-digested plasmid candidates which display bands larger than 830 bp, turation in H2O at 100°C for 10 min followed by the addition of the appropriate salts. Further results with conalbumin candidates will be published elsewhere.

Fig. 3. Restriction analysis of ovomucoid cDNA plasmid candidates (A) and lysozyme cDNA plasmid candidates (B). Approximately 1.5 µg of DNA from each candidate was digested with Hpa II. In Panel A, Lane 1, pMu2; Lane 2, pMu3; Lane 3, pMu7; Lane 4, pMu8; Lane 5, EcoRI/HindIII-digested λ-DNA; Lane 6, pCR1; Lane 7, pMu10; Lane 8, pMu11; Lane 9, pMu12. Electrophoresis was in 1.5% agarose-TBE. In Panel B, Lane 1, pLys1; Lane 2, pLys2; Lane 3, pLys3; Lane 4, HindIII-digested SV40 DNA; Lane 5, pCR1; Lane 6, pLys4; Lane 7, pLys5; Lane 8, pLys6. Electrophoresis was in 3.5% acrylamide-TBE.

<p>| Table I |
| Candidate hybridizations to cDNA probes |
| Approximate 1.5 µg of DNA from each candidate was digested with Hpa II. Samples were phenol-extracted, ethanol-precipitated, and prepared for hybridization as described under “Experimental Procedures.” Hybridizations were in 20 µl for 12 h and then terminated by the addition of 1 ml of cold S1 nuclease buffer (3). All data have had 18 cpm of counter background subtracted. |</p>
<table>
<thead>
<tr>
<th>pMu candidates</th>
<th>pLys candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>cpm</td>
<td>S1 resistance</td>
</tr>
<tr>
<td>No DNA</td>
<td>12</td>
</tr>
<tr>
<td>Input</td>
<td>632</td>
</tr>
<tr>
<td>pOvE12</td>
<td>12</td>
</tr>
<tr>
<td>pOvE12</td>
<td>76</td>
</tr>
</tbody>
</table>
therefore, must contain inserted sequences. A number of each candidate set did appear to contain insertions into pCR1. Of
the ovomucoid candidates, the most promising is in Lane 3 of
Fig. 3A, pMu7, with a new band at 1650 to 1700 bp (insert size = 820 to 870 bp). When pMu7 was retransformed into fresh
E. coli, the diffuseness of this band was no longer found and
the size was 830 bp. Candidates in Lanes 7 to 9 in Fig. 3A
have inserts of at least 740 bp. Insert sizes could only be
interpreted as minimums because of the possibility that Hpa
II also cut the inserted sequence. Several of the lysozyme
candidates in Fig. 3B also must have inserts. For pLys2 and
pLys6 in Lanes 2 and 8, the minimum insert sizes are 400 and
280 bp, respectively.

Restriction Mapping and Sequencing of pMu7 and pLys6

Our next objective was to identify absolutely one member
of each set. It was possible to do this by chemically sequencing
the DNAs and comparing those results with the expected
nucleotide sequences based on the amino acid sequences of
the proteins. The order of amino acids for ovomucoid and
lysozyme has been determined (22, 23). We first prepared
a partial restriction map for two of the plasmids, pMu7 and
pLys6.

pMu7—pMu7 DNA was isolated by ethidium bromide-ce-
esium chloride density gradient centrifugation. It was restricted
with Hpa II, Hha I, Hae III, or HindIII. Each of these digest
was split into two portions and one-half was digested addi-
tionally with EcoRI. The minus and plus lanes for EcoRI
restrictions of pMu7 are shown in Lanes 3 and 4, respectively,
for the various digests in Fig. 4. They can be compared with
the minus (Lane 1) and plus (Lane 2) EcoRI restrictions of
the various digests of pCR1. Compared to pCR1, the Hpa
II and Hha I restrictions of pMu7 show one new band, both
indicating the insertion of about 820 bp of foreign DNA. This
foreign DNA is cut once by EcoRI 190 bp from site of
insertion.

Hae III and HindIII digest of pMu7 are more complicated in
that each enzyme cuts the plasmid DNA within the insertion
several times. For Hae III digestions all of these fragments
comigrate with other bands. However, Hae III plus EcoRI
yields two new bands, 340 and 220 bp. HindIII digestions of

pMu7 show four new bands 1060, 560 to 590, 230, and 190 bp.
None of these new bands, however, are cut by EcoRI. There-
fore, at least one more small HindIII band must be present in
the pMu7 insert.

In order to clarify the Hae III and HindIII digestions further,
a Hha I digest of pMu7 was conducted and the large insert
bearing band was excised after electrophoresis. This band of
DNA was then restricted with Hae III (Fig. 5, Lane 3) or HindIII (Lane 4). Lane 1 shows the
smaller fragments of EcoRI/HindIII-digested λ-DNA and Lane 2,
HindIII-digested pBR322. pLys6 (5 μg) was digested with Hha I and
two insert-containing bands (950 and 900) were recovered after elec-

trophoresis and end-labeled with polynucleotide kinase. Panel B is
an autoradiogram of the HindIII digestion products of the 900 bp Hha
I fragment. The 530-bp Hha I-HindIII fragment in Panel B was re-
duced and digested with either Hpa II (Panel C, Lane 1) or Hae III
(Lane 2) or Alu I (Lane 3).

Fig. 4. Restriction analysis of pMu7 and pLys6. Each panel is
labeled with the enzyme used. The DNA in Lanes 2 and 4 of each
panel were also restricted with EcoRI endonuclease. For each panel,
Lanes 1 and 2, pCR1; Lanes 3 and 4, pMu7; and Lane 5, pLys6.
Unmarked lanes contain a combination of molecular weight markers,
EcoRI/HindIII λ-DNA plus HindIII pBR322 DNA. Electrophoresis
was in 3.5% acrylamide-TBE (Hpa II and Hha I) or 4% acrylamide-
TBE (Hae III and HindIII). Distortion of the high molecular weight
lanes in the 4% acrylamide gel were the result of overloading the gel
in order to visualize the shorter DNA fragments.

Fig. 5. Analysis of isolated restriction fragments of pMu7
and pLys6. Approximately 20 μg of pMu7 was digested with Hha I
and the 2140-bp fragment (containing the insert) recovered after elec-
trophoresis. The DNA fragment was digested with either Hae III
(Panels A, Lane 3) or HindIII (Lane 4). Panel A, Lane 1 shows the
smaller fragments of EcoRI/HindIII-digested λ-DNA and Lane 2,
HindIII-digested pBR322. pLys6 (5 μg) was digested with Hha I and
2 inserts containing bands (950 and 900) were recovered after elec-
trophoresis and end-labeled with polynucleotide kinase. Panel B is
an autoradiogram of the HindIII digestion products of the 900 bp Hha
I fragment. The 530-bp Hha I-HindIII fragment in Panel B was re-
duced and digested with either Hpa II (Panel C, Lane 1) or Hae III
(Lane 2) or Alu I (Lane 3).
Recombinant Plasmids Bearing cDNA to Ovomucoid and Lysozyme mRNAs

FIG. 6. Partial restriction maps of pMu7 and pLys6. Restriction maps were prepared from data shown in Figs. 4 and 6 and DNA sequencing analysis (not shown). For pMu7 this included fragments labeled at each of the three Hae III sites and for pLys6, fragments labeled at the Hpa II sites. Restriction sites in pCR1 (external to the Kan^R') junctions) have been mapped previously (9). Symbols are I, Hha I; II, Hpa II; III, Hae III; and I, HinfI. Maps are oriented with kanamycin resistance determinants (KAN^R') at the left terminus and the 5'-untranslated region and the HindIII site at nucleotide 24. Each sequence was verified by sequencing three independent isolates from the colony bank.

FIG. 7. Partial sequences of chick ovomucoid and lysozyme mRNAs. In each panel, the nucleotide sequence is aligned with the COOH-terminal amino acid sequence of the respective proteins and is given in a one-letter code. For pMu7 the sequence was obtained from fragments end-labeled at the EcoRI site at nucleotide 20 in the 3'-untranslated region and the HinfI site at nucleotide 25. Each sequence was verified by sequencing three independent isolates from the colony bank.

had expected this EcoRI site to occur at the sole possible site within the translated region, amino acids 50-51 (glutamic acid-phenylalanine). However, no EcoRI site occurs at that position and the EcoRI site in pMu7 is encoded by nucleotides which are 20 to 26 bp 3'-ward (relative to mRNA polarity) from the last translated nucleotide (GAAUUC at position 20 in Fig. 7). By sequencing we also determined that HinfI sites lie 54 bp 5'-ward and 25 bp 3'-ward from the EcoRI site.

pLys6—Restriction analysis of pLys6 DNA are shown in Lane 5 of Fig. 4. The Hae III digest displays two new bands of 810 bp and 315 bp (insert size equals 560 bp) and the HinfI digest also displays two new bands, 1060 bp and 810 bp (insert size equals 570 bp). Each of these enzymes has only one site in the pLys6 insert and is mapped in Fig. 6b relative to sites in pCR1. The Hpa II and Hha I digests of pLys6 in Fig. 4 also display two new bands each. They are 1120 bp and 240 bp (apparent insert size equals 530 bp) and 950 bp and 900 bp (apparent insert size equals 450 bp), respectively.

We proceeded to map the relative orientation of the two Hha I fragments from the pLys6 insert by isolating them from the gel seen in Fig. 4 and end-labeling them with [γ-32P]ATP and polynucleotide kinase. Each fragment was redigested with HinfI. The Hha I 950-bp fragment yielded a 330-bp and a 110-bp fragment. It, therefore, overlaps the internal HinfI site and contains the rightward terminus (not shown). The Hha I 900-bp fragment yielded a 520-bp and 390-bp fragment (left
with AZu I to yield 150-bp and 90-bp fragments. AZu I restricts protein. The Hpa II 240 bp end-labeled fragment was digested fragment labeled at an internal Hpa II site was sequenced. shown in Fig. 7b, aligns with the COOH-terminal portion of which are third and second from the last amino acid in the lysozyme. The Hpa II site (CCGC) occurs at cysteine-arginine which are third and second from the last amino acid in the protein. The Hpa II 240 bp end-labeled fragment was digested with Alu I to yield 150-bp and 90-bp fragments Alu I restricts at a site very near the EcoRI site of pCR1 (9). The 150-bp fragment labeled at an internal Hpa II site was sequenced. Had that site been identical with the Hpa II site mentioned above, the first nucleotides would have encoded the final amino acid and the termination codon; however, this was not the case. pLys6, therefore, contains Hpa II sites separated by about 30 bp in its insertion into pCR1. Because pLys6 does not contain a full length copy of lysozyme mRNA, we cannot be sure it includes sequences complementary to the entire 3'-untranslated region. It does not in fact contain the sequence AAUAAA found in the 3'-untranslated region of other eukaryotic mRNAs. The restriction maps in Fig. 6 show the accumulated data for these plasmids and the direction of transcription for the respective mRNAs.

Homologies in the 3'-untranslated Regions of Ovalbumin mRNA and Ovomucoid mRNA—The 3'-untranslated region of ovomucoid mRNA, deduced from pMu7 (Fig. 7), was compared with the 3'-untranslated region of ovalbumin mRNA. The latter sequence has been published previously (24). The comparison was made with the aide of a computer program designed by Dr. L. Kom

1) 1267 CAGAGCAGUG 1276 → Ovalbumin mRNA
   8 CAGAC GUG 16 → Ovomucoid mRNA
2) 1484 UGUAA CAUG 1494
   1 UGA ACAUG 10
3) 1533 AGC CAGAU AUUC 1544
   92 AGCUCAGUAUU 104
4) 1584 AUUA AUAUAUAUAUA 1601
   124 AUAA AUAAUAUAUAUA 136
5) 1837 UCAUAA UAAA 1846
   115 UCUAU AAUUA 124

Fig. 8. Homologies in the 3'-untranslated regions of ovalbumin mRNA and ovomucoid mRNA. Comparisons were made by a published computer program (Ref. 21, see “Experimental Procedures”). For each of the five homologies the top row of nucleotides is from ovalbumin mRNA and the bottom row, ovomucoid mRNA. Numbers in the top rows refer to nucleotide number assignments made by L. McReynolds et al. (24) in which number 1 is at the 5' terminus of ovalbumin mRNA. Numbers in the bottom rows refer to the position of the nucleotide in the 3'-untranslated region of ovomucoid mRNA as shown in Fig. 7.

DISCUSSION

We have been interested in steroid regulation on egg-white protein synthesis in the chick oviduct (25). The experiments described were designed to yield pure probes for the minor mRNA species. The methods which we chose did not require the purification of individual mRNAs, but did rely on a knowledge of the mRNA lengths and the amino acid sequences of the proteins. We worked with a large library of recombinant DNA-harboring bacterial colonies in E. coli χ1776. These colonies contain cDNA to hen oviduct poly(A)-containing RNA. The size of the colony library was necessitated by the low reverse transcriptional efficiency of conalbumin mRNA (7). Our initial screening of the library was by the method best adapted to mass screening, colony filter hybridization. To obtain suitable probes we exploited the ability to visualize the four major mRNAs in hen oviduct poly(A)-containing RNA after denaturing electrophoresis (Fig. 1). Although the RNA excised from the CH3HgOH/agarose gel would not reverse transcribe, we were able to label the RNA in vitro by 5'-phosphorylation. This labeled RNA then served as a probe for the colony filter assay.

We further limited the number of colonies to be studied by conducting hybridization (Table I) and restriction analyses (Fig. 3). From those experiments we chose one ovomucoid cDNA plasmid candidate, pMu7, and one lysozyme cDNA candidate, pLys6, for further study. We decided to proceed directly to chemical sequencing of the DNA for absolute identification. This, however, first necessitated preparing a partial restriction endonuclease map of these plasmids. In preparing the maps we relied on previous work (9) which mapped a number of restriction endonuclease sites near the EcoRI site in pCR1. From that information plus information shown in Figs. 4 and 5, maps were prepared (Fig. 6). These maps are in agreement with the results of other workers (26, 27). By chemical sequencing of the plasmids we identified the nucleotides encoding the COOH termini of the respective proteins.

By sequencing we also identified a second Hpa II site in the cDNA portion of pLys6, estimated by poly(dA:dT) junction lengths to be 30 to 40 residues and determined the sequence of the 3'-untranslated region of ovomucoid mRNA. Poly(dA:dT) tracts were shorter than anticipated (see “Experimental Procedures”), indicating the addition of “tails” to nicks in the DNA.

The 3'-untranslated region of ovomucoid mRNA contains the characteristic AAUAAA sequence 14 nucleotides from poly(A). This same sequence is found in the 3'-untranslated regions of all eukaryotic mRNAs sequenced to date (24, 27-33). As in other mRNAs this region of ovomucoid mRNA is AU-rich. The sequence UUU which precedes AAUAAA by six nucleotides in ovomucoid mRNA is found in nearly the same position in rabbit α-globin mRNA (28), human α-globin mRNA (29), rat growth hormone mRNA (30), ovalbumin mRNA (24), and mouse dihydrofolate reductase mRNA. The

4 J. Nunberg, Department of Biological Sciences, Stanford University, personal communication.
3'-untranslated region of ovomucoid mRNA also contains other homologies with the same region of ovalbumin mRNA (Fig. 8); however, elucidating the biological significance of these sites must await future experiments.

Acknowledgments—We are grateful for the help and suggestions offered by Charles Yanofsky, Gerard Zurawski, and Gary Tschumper. Special thanks go to Eric Selker for preparing the computer analysis and to Bill Dower and Gary Crouse for critical reading of this manuscript. We would also like to thank M. Chi Nguyen-Huu for his sharing of data before publication.

REFERENCES