Ovalbumin Gene

ACTION OF RESTRICTION ENDONUCLEASES UPON DNA CODING SEQUENCE*

(Received for publication, November 23, 1976, and in revised form, March 1, 1977)

JOHN J. MONAHAN, SAVIO L. C. WOO, CHARLES D. LIARAKOS, AND BERT W. O'MALLEY

From the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

A full length, double-stranded ovalbumin complementary DNA (cDNA) was synthesized in vitro and used to assay the action of 35 restriction endonucleases upon this DNA sequence. Some 22 enzymes, Bal I, BamHI, Bgl II, and its isoschizomer Xho I, Ppu I, Xho I, Taq I, Sal I, Bgl I, Bst I, Kpn I, Sac II, Hae II, Hpa I, Hpa II, Hha I, Hind I, and its isoschizomer HincII, Hind III and its isoschizomer Hsu I, EcoRI, and Sma I, failed to cut such a DNA sequence. However, 13 enzymes did cut the DNA. They were Hae III, Pst I and its isoschizomer Xma II, Sac I, Hpa I, Xho I, Hph I, EcoRII, Hinfl, Mnl I, Mbo I, Mbo II, and Alu I. A detailed map of the restriction enzyme recognition sites for 11 of these restriction enzymes was determined.

Since the amino acid sequence of a phosphorylated peptide from hen ovalbumin is known, it was therefore possible to construct a partial nucleotide sequence corresponding to this peptide fragment. Such a sequence could possibly be cut by the three enzymes, EcoRII, Hinf I, and Hph I, in close proximity to each other. Since these three enzymes did actually cut the ovalbumin cDNA in such close proximity, we have therefore, tentatively assigned the coding portion of the ovalbumin gene to a region which begins within 200 nucleotides from 5'-terminal end of the gene.

We have recently reported the enzymatic in vitro synthesis of complete length double-stranded ovalbumin cDNA (1). Highly purified ovalbumin mRNA was used to synthesize a full length complementary DNA using the enzyme AMV reverse transcriptase. After removal of the mRNA by alkali, the second strand of the DNA was synthesized using AMV reverse transcriptase and the first cDNA strand as a template. We demonstrated that we had indeed generated a double-stranded molecule that was a faithful representation of the ovalbumin structural gene.

We now wish to describe the action of some 35 restriction endonucleases upon this DNA sequence. It is hoped that this data will be of use for DNA sequencing studies of the structural gene, since one may not only obtain information about the presence or absence of specific nucleotide sequences within the DNA, but one may also obtain small discrete DNA fragments ideally suited for such studies. A map of the sites at which restriction endonucleases cut the structural gene sequence will be useful to study the extent and orientation of the gene inserted into bacterial plasmids (2, 3). Such a map will also be of use for the construction of a physical map of the ovalbumin structural gene within its adjacent sequence in the cellular genome. Coupled with our knowledge of the amino acid sequence for ovalbumin, these data will permit localization of the protein coding sequences within the mRNA molecule. Information about what restriction enzymes cut the structural gene should allow the appropriate enzymes to be used to cut the cellular DNA into reasonable small size fragments suitable for their incorporation into bacterial plasmids using the "shotgun technique" (4, 5) or by other methods. Finally, by using selected restriction endonucleases it will be possible to prepare cDNA probes for RNA transcripts from the 5'- and 3'-terminal ends of the ovalbumin structural gene. Such probes will be of use in studying the relative extents of transcription of the 5' and 3' regions of the gene in chromatin in vitro.

METHODS

Isolation of Ovalbumin mRNA—Purified ovalbumin mRNA was isolated as we have described previously (6, 7). The purity of the ovalbumin mRNA was determined by analytical procedures such as acid-urea agarose gel electrophoresis, polyacrylamide gel electrophoresis in 8% formamide, sucrose gradient centrifugation in 70% formamide, translation in vitro in a heterologous protein-synthesizing system, and enzymatic digestion followed by a homochromatographic analysis of the resultant oligonucleotide maps. Only ovalbumin mRNA of greater than 99% purity was used in these studies.

Preparation of Reverse Transcriptase—Highly purified avian myeloblastosis virus reverse transcriptase was generously supplied by Dr. J. W. Beard (Life Sciences, Inc., St. Petersburg, Fla.). The purification procedure was a modification of that of Kacian and Spiegelman (8) and is described in detail elsewhere (9).

Preparative Scale Synthesis of Ovalbumin cDNA—The procedure for synthesis of the complete ovalbumin cDNA labeled with 32P to a specific activity of 10⁶ cpm/μg for use in these studies was as follows: 50 μCi of [32P]dATP were hyylized to dryness and resuspended in 10 ml of a solution containing 50 mM Tris/HC1 (pH 8.3), 20 mM dithiothreitol, 10 mM MgCl₂, 800 μM dTTP, dGTP, dCTP, and dATP, 2.5 μg/ml of oligo(dT)₂₀, 36 μg/ml of actinomycin D, 1% ethanol, and greater than 100 μg/ml of ovalbumin mRNA. The actinomycin D was made up as a stock solution of 360 μg/ml in 10% ethanol. The mixture was left on ice for 5 min. Then reverse transcriptase enzyme was quickly added to a final concentration of 60 units/ml. The solution was quickly vortexed and incubated for 15 min at 46 °C. The reaction mixture was then made 0.5% with sodium Sarkosyl, 10 mM EDTA, and 100 μg/ml with poly(rA). The entire sample was then placed on a Sephadex G-50 column equilibrated with 0.2 M Tris/HCl, pH 8.3.
Hga I, HincII, Hsu I, Hinfl, Hph I, Kpn I, Mbo I, Mbo II, Mnl I, Pst I, BarnHI, EcoRI, EcoRII, Hue II, Hue III, Hpa I, HindII, HindIII, HhaII, borate, pH 8.3, 1 mM EDTA. Electrophoresis was in the same buffer by BlairCraft (Cold Spring Harbor, N.Y.) (10). The gels were stained after electrophoresis with 5 μg/ml of ethidium bromide in the electrophoresis buffer for 10 to 20 min. The SV40 or phage A carrier DNA fragments were determined by electrophoresis on agarose slab gels.

The reaction was terminated by the addition of EDTA. The double-stranded cDNA was then deproteinized with phenol and the unretracted 3' labeled dTTP was removed by Sephadex G-50 chromatography. Slab Gel Electrophoresis - The size of the ovalbumin DNA fragments were determined by electrophoresis on agarose slab gels. Varying concentrations of agarose were dissolved in 90 mM Tris/borate, pH 8.3, 1 mM EDTA. Electrophoresis was in the same buffer for 6 to 8 h at 100 mA/slab. The gel apparatus (16 × 16 cm) was made by BlairCraft (Cold Spring Harbor, N.Y.) (10). The gels were stained after electrophoresis with 0.5 μg/ml of ethidium bromide in the electrophoresis buffer for 10 to 20 min. The 3840 or phage carrier DNA present was then visualized using ultraviolet light. When required cDNA were eluted from the agarose gels using the "freeze-squeeze method" of Thurung et al. (11).

Restriction Enzymes - The restriction endonucleases Alu I, BamHI, EcoRI, EcoRII, Hae II, Hae III, Hpa I, HindIII, Hph I, Hha II were obtained from BRL Inc. (Bethesda, Md.). Sma I and Hinf I were kindly furnished by Dr. Richard J. Roberts (Cold Spring Harbor, N.Y.).

RESULTS

Purification of Gene Sequence

One essential requirement for accurate mapping of restriction endonuclease sites within a DNA fragment is that the DNA itself be homogenous with respect to its size. We have recently developed conditions whereby it is possible to obtain predominantly complete length cDNA copies from an mRNA.  

Addition of [3H]Poly(dT) to 3'-Terminal of Double-stranded Oualbumin cDNA - The optimum conditions for the enzymatic synthesis of double-stranded cDNA, using reverse transcriptase have been described by us in detail elsewhere (1). In summary, the conditions used in these studies were 1 μCi of [3H]dGTP was lyophilized to dryness and resuspended in 0.5 μl of a solution containing 50 mM Tris/HC1 (pH 8.3), 20 mM dithiothreitol, 800 μM MgCl2, 10 mM dATP, dGTP, and dCTP, and 40 μg/ml of cDNA... Reverse transcriptase was added to a final concentration of 200 units/ml and the solution was quickly vortexed and incubated for 4 h at 46°. The reaction mixture was then made 10 mM with EDTA and phenol-extracted. The aqueous phase was placed on a Sephadex G-50 column as described above and the purified ovalbumin structural gene was collected and precipitated with 2 volumes of ethanol.

Fraction 11 and 12 of Fig. 1 was determined by agarose gel electrophoresis. However, the poor yield of material obtained by this method prompted us to try an alternative approach. It has been shown previously that single-stranded cDNA species are apparently synthesized as discrete size classes (1, 14). We have observed that if the largest product made was a complete length double-stranded DNA structure, the next smallest double-stranded cDNA, species do not have the resolution to discriminate between molecules within 100 nucleotides of each other in length, it appears that since there is not a continuous spectrum of lengths among the single-stranded cDNA species within the gradient, such a resolution was not needed. Indeed, the next smallest double-stranded cDNA, structure that could be made was approximately 100 nucleotide base pairs shorter than a complete length cDNA transcript (1).

The length of the 32P-labeled double-stranded DNA, structure made from the single-stranded DNA, species present in Fractions 11 and 12 of Fig. 1 was determined by agarose gel electrophoresis as shown in Fig. 2. The location of the 32P-labeled double-stranded DNA, was determined by autoradiography. Also run on the same gel was 32P-labeled SV40 DNA which was cleaved with the restriction endonuclease Hae III. Using these DNA fragments which have been well characterized with respect to size (15), a size estimate of 1850 ± 25 nucleotide pairs was obtained for the double-stranded cDNA species. This length, which is in agreement with all of the
EcoRI, pairs larger than our earlier estimates (1). It is, we believe, restriction enzymes that do not cut the double-stranded DNA sequence was concluded when the phage DNA was double-stranded 3H-labeled end.

termine which restriction enzyme DNA fragment corresponds to the 3' end of the double-stranded cDNA, simply by determining which of the DNA fragments in an agarose gel contain the free 3'-(OH) terminus with 3H using [3H]dTTP and the enzyme terminal deoxyribonucleotidyltransferase, one can determine the position of the ovalbumin DNA sequence as shown in Fig. 5.

Orientation of Restriction Enzyme Sites Within Gene

The double-stranded cDNA\textsubscript{a}, sequence made as described under "Methods" has only one free 3'-(OH) terminus. The other end of the molecule corresponding to the 5' terminus of the mRNA contains a closed loop structure (1, 3). The presence of such a loop conveniently allows the orientation of restriction enzyme sites within DNA sequence to be made. By labeling the free 3'-(OH) terminus with 3H using [3H]dTTP and the enzyme terminal deoxyribonucleotidyltransferase, one can determine which restriction enzyme DNA fragment corresponds to the 3' end of the double-stranded cDNA\textsubscript{a}, simply by determining which of the DNA fragments in an agarose gel contain the 3H-labeled end.

Restriction Enzymes That Do Not Cut the Double-stranded cDNA\textsubscript{a}, Sequence

The following enzymes failed to cut the ovalbumin DNA sequence: Bal I, HinII and its isoschizomer HindII, Hha I, Hpa II, Hae II, HindIII and its isoschizomer Hsu I, BamHI, EcoRI, Hpa I, Tag I, Sma I, Bgl I, Ppu I, Bsu I and its isoschizomer Xho I, Sal I, Bgl II, Kpn I, Sat I and Xba I. In all cases ~2 µg of phage λ-DNA were present within the same reaction mix during the digestion with the above enzymes. The absence of a restriction enzyme site within the ovalbumin DNA sequence was concluded when the phage DNA was digested to completion while the double-stranded [32P]-DNA\textsubscript{a}, remained intact (data not shown). For some enzymes (BamHI, EcoRI, Sma I, Sal I, Kpn I, Xba I, and Xho I) which cut phage λ-DNA only a few times (13), it was necessary to use a 1% agarose gel to observe the change in mobility of the phage DNA fragments upon restriction enzyme treatment. In other cases, however, digestion of the phage λ-DNA could be clearly seen in 2.5% agarose gel (data not shown).

Restriction Enzymes That Cut Ovalbumin DNA Sequence

The following enzymes cut the ovalbumin DNA sequence: Hae III, Pst I, Xma II, Sat I, Hga I, Xho II, Hph I, EcoRII, HindIII, Mnl I, Mbo I, Mbo II, and Alu I. Fig. 3 shows the restriction enzyme patterns obtained for the phage DNA which was digested along with a 32P-labeled ovalbumin DNA. Only these gels in which the phage λ-DNA had been digested to completion as shown were used in these studies. Fig. 4 shows the restriction enzyme patterns of the 32P-labeled ovalbumin DNA when treated with the corresponding enzymes shown in Fig. 3. We shall discuss the mapping data for each enzyme briefly.

Hae III—Treatment of the ovalbumin DNA with the enzyme Hae III isolated from Haemophilus aegyptius (16) yields two fragments 1100 and 750 nucleotide pairs long (Fig. 4). Using ovalbumin [32P]-DNA which had a 3' labeled 3'-terminal label, 98% of the 3' counts were associated with the 1100-nucleotide fragment. One can, therefore, assign the position of the Hae III site to a region 750 nucleotides from the 5' terminus of the ovalbumin DNA sequence as shown in Fig. 5.

Pst I and Xma II—These two enzymes isolated from Providencia stuartii 164 (17) and Xanthomonas malvaccarum (12) are isoschizomers and would, therefore, be expected to yield the same DNA fragments. In both cases a large fragment 1350 nucleotide pairs long and a smaller fragment 500 nucleotide pairs long were obtained. Digestion of the ovalbumin DNA sequence Pst I or Xma II together with Hae III reduced the size of the 1350-nucleotide fragment to 1100 nucleotides. It did not affect the size of the 500-nucleotide-pair DNA fragment; a smaller size of 250 nucleotide pairs fragment was also seen. This was consistent with the enzyme cutting the ovalbumin DNA sequence 500 nucleotides from the 5'-terminal and of the DNA as shown in Fig. 5.

Sat I—This enzyme isolated from Streptomyces stagneri (12) cut the double-stranded cDNA\textsubscript{a}, into a large fragment 1450 nucleotide pairs and a smaller one 400 nucleotide pairs long. Again simultaneous treatment of the double-stranded cDNA\textsubscript{a}, with both Sat I and Hae III lead to a reduction in size of the larger fragment to 1100 nucleotides. The size of the smaller 400-nucleotide fragment was unaffected and a new 350-nucleotide fragment appeared. The data suggest that the enzyme cut the DNA sequence 400 nucleotides from the 5'-terminal as shown in Fig. 5.

Hga I—This enzyme isolated from Haemophilus gallinarum (18) could be shown to yield only two DNA fragments; a larger one 1100 nucleotide pairs and a smaller one 750 nucleotide pairs long. However, a simultaneous treatment with both Hga I and Hae III yielded only two fragments 750 nucleotides and 350 nucleotides long. In contrast, treatment with Hga I and Pst I yielded fragments 750, 600, and 500 nucleotide pairs long. These data are consistent with a position of the Hga I site being 1100 nucleotides from the 5'-terminal of the ovalbumin DNA sequence.

Xho II—This enzyme isolated from Xanthomonas hollcota (12) cut the double-stranded cDNA\textsubscript{a}, sequence quite similarly to that of Hae III. A large fragment 1175 nucleotide pairs long

\textsuperscript{a} R. Roberts, personal communication.
and a smaller 675-nucleotide pair fragment were observed. The larger fragment was reduced in size, however, to 1100 nucleotides upon simultaneous digestion with Hae III. The 675-nucleotide fragment was unaffected by these conditions. This suggested a site for the enzyme 650 nucleotides from the 5'-terminal of the ovalbumin DNA sequence as shown in Fig. 5.

Hph I - The enzyme Hph I isolated from Haemophilus parahaemolyticus (19) cut the double-stranded cDNAα into two fragments 1300 nucleotide pairs and 275 nucleotide pairs. Since there are 1850 nucleotide pairs within the total double-stranded cDNAα sequence, there should be two species of the 275-nucleotide fragments for each 1300-nucleotide fragment. The high intensity of the band on the film due to this fragment was clearly in agreement with this. Three possible models could be drawn; either a 275-nucleotide species at each end of the DNA or both together at either the 5'- or 3'-terminal ends of the ovalbumin DNA sequence. Simultaneous digestion with both Hph I and Hae III split the large fragment into two pieces 825 and 475 nucleotides long, while treatment with Hph I and Pst I split the large fragment into two pieces 1075 and 225 nucleotides long. The 275-nucleotide band was unaffected by either Hae III or Pst I. These results were only consistent with the two 275-nucleotide fragments being located at the 3' and 5' termini of the ovalbumin DNA sequence as shown in Fig. 5.

EcoRII - A somewhat similar situation was observed for the enzyme EcoRII (20, 21). Again, a large fragment 1300 nucleotide pairs long was obtained together with only one other...
small fragment 275 nucleotide pairs long (Fig. 4). As in the case of Hph I, two identically sized 275-nucleotide fragments were suspected. Treatment with EcoRII and Hae III yielded a 1100-nucleotide and a 200-nucleotide fragment, and the intense 750-nucleotide band. The location of the two small 275-nucleotide fragments toward the 5'-terminal end of the ovalbumin DNA sequence as shown in Fig. 5 was confirmed by isolating from an agarose gel the 750-nucleotide fragment of the gene DNA liberated by Hae III treatment alone and treatment of the DNA with EcoRII. The products obtained were a 200-nucleotide and a 275-nucleotide fragment. In contrast, treatment of the similarly isolated 1100-nucleotide Hae III fragment with EcoRII was without effect; the only product obtained was the 1100-nucleotide fragment. From the size of the smallest EcoRII and Hph I fragments, it was not possible to locate precisely enough of the relative map positions of these two enzymes at the 5' end of the gene.

HinfI - This enzyme isolated from Haemophilus influenzae R (12) cut the double-stranded cDNA, in three places yielding three fragments, an 850-nucleotide, 450-nucleotide, and intense 275-nucleotide band as shown in Fig. 5. Using DNA in which the free 3'-terminal was labeled with 3H it was clear that the 850-nucleotide fragment belonged to the 3'-terminal end of the ovalbumin DNA sequence. Co-digestion of the double-stranded cDNAH in HinfI and Hae III brought about the cutting of the 550-nucleotide fragment into two fragments 400 and 150 nucleotides long. Treatment of the double-stranded cDNAH in both Mnl I and Hae III with Mnl I and Pst I in contrast led to the cutting of the 225-nucleotide fragment into two fragments 175 and 50 nucleotides long; the intensity of the 175-nucleotide band being twice as great as in the case of Mho I alone. The restriction sites for the enzyme could then be drawn as shown in Fig. 5.

Mbo I - This enzyme isolated from Moraxella bovis (12) cuts the ovalbumin DNA sequence in three places giving fragments 1000, 450, 225, and 175 nucleotides long as shown in Fig. 4. Using DNA in which the free 3'-terminal was labeled with 3H it was clear that the 1000-nucleotide fragment was located at the 3'-terminal of the ovalbumin DNA sequence. Co-digestion of the cDNAH with both Mbo I and Hae III led to the cutting of the 175-nucleotide fragment into two species 100 and 70 nucleotides long. Treatment with Mbo I and Pst I in contrast led to the cutting of the 225-nucleotide fragment into two fragments 175 and 50 nucleotides long; the intensity of the 175-nucleotide band being twice as great as in the case of Mbo I alone. The restriction sites for the enzyme could then be drawn as shown in Fig. 5.

Mbo II and Alu I - These two enzymes isolated from Moraxella bovis (12) and Arthrobacter luteus (12), respectively, cut the ovalbumin in too many places for us to obtain an accurate location of the restriction enzyme sites with the data presently available. We estimate at least five fragments were generated
by Mbo II and seven by Alu I. Many of these fragments were less than 200 nucleotides in length, making calculation of their size difficult to determine in an agarose gel electrophoresis system.

**DISCUSSION**

In this paper we have studied the ability of some 35 restriction endonucleases to cut the ovalbumin gene. Of the 22 enzymes which did not cut the gene, the sequence-specific sites of 14 are now known (12) and are summarized in Table I. This information will not only be of use in nucleotide sequencing studies of the gene itself, it will be of value in DNA recombinant studies using plasmid or phage DNAs, where one may want to cut the total cellular DNA with a particular restriction enzyme that does not cut into the structural gene itself. The fact, for example, that EcoRI did not cut the ovalbumin structural gene sequence makes it feasible to use this enzyme to insert cellular DNA fragments into bacterial plasmids using the "shotgun approach" (4, 5) and still hope to obtain a complete ovalbumin structure gene sequence in one clone. Of the 13 restriction enzymes which did cut the ovalbumin structural gene (Table I), the nucleotide sequence of the recognition site of 12 are now known (12). We can thus, assign specific sequences at the locations shown in Fig. 6 to 10 of the above nucleotide sequences. It is interesting to note in this study that when a restriction enzyme cuts toward the 3'-terminal end of the gene, that fragment (if it is small) always migrates as a somewhat broad band upon electrophoresis in the agarose gel. A clear example of this is seen in the case of the Hinfl (Fig. 5). The 400-nucleotide fragment is broader than either the 850- or 275-nucleotide fragments. We suspect that this has arisen due either to some heterogeneity in the lengths of the poly(A) tails in the mRNA, or during the synthesis of the poly(T) tail of the cDNA itself.

We have attempted to use the above data to predict the location of the region of ovalbumin mRNA that is actually translated into protein in vivo. Since there are 387 amino acids in hen ovalbumin, a total of 1161 nucleotides would be required to code for the protein itself (7). This leaves approximately 729 nucleotides of the 1890-nucleotide mRNA molecule which do not code for the above amino acids. Of these, approximately 62 adenylate residues (7) and a minimum of another 75 nucleotides (22) should be located at the 3'-terminal end of the mRNA. Because ovalbumin is a secretory protein, a precursor molecule with at least 25 extra NH2-terminal amino acids (corresponding to 75 nucleotides) is probably first synthesized in vivo (22, 24). Therefore, one could predict that the position for the initiation of protein synthesis within the mRNA would lie between a region approximately 50 nucleotides (thereby allowing for an mRNA ribosome binding site (25, 26)) from the 5'-terminal end of the mRNA and a region no more than 592 nucleotides from the 5'-terminal end of the mRNA. Since the amino acid sequence of a phosphorylated peptide located toward the NH2-terminal end of hen ovalbumin is known (27), a corresponding partial nucleotide sequence for this region of the gene can be generated taking into account the redundancy of the genetic code due to wobble in the third base of each amino acid codon. Analysis of this nucleotide sequence (Table II), reveals the presence of six possible restriction enzyme sites for the enzymes EcoRII, Hph I, and Hinfl in close proximity to each other. A similar clustering of these same sites was observed following digestion of the actual ovalbumin gene sequence (Fig. 6). No such cluster of sites for these three enzymes was observed elsewhere within the other potential nucleotide sequences generated from known amino acid sequence data. Thus, we could speculate that the codons for these amino acids are located at a region which is approximately 225 nucleotides from the 5'-end of the mRNA. If we allow for a precursor protein to ovalbumin with an extra 25 NH2-terminal amino acids to be synthesized in vivo (23, 34) and for the first five NH2-terminal amino acids which are known (27), then the start of the protein coding region of the mRNA may be located within 135 nucleotides of the 5'-terminal end of the mRNA. This is clearly within the predictable boundary region of 50 to 592 nucleotides mentioned above. Experiments are presently underway to determine the nucleotide sequence at the 5'-end of the ovalbumin structural gene to test the validity of these observations.

Restriction enzyme studies of this type have also been carried out recently by Maniatis et al. (3) for the rabbit β-globin synthetic gene. However, no similarities in the cutting by restriction enzymes of the rabbit β-globin and ovalbumin gene are apparent to use at this level of sequence resolution. We are

---

**Table I**

**Occurrence of specific sequences within ovalbumin structural gene**

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Absent</th>
<th>Sequence</th>
<th>Enzyme</th>
<th>Present</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bal I</td>
<td>TGGCCA</td>
<td>Hae III</td>
<td>GCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HincII</td>
<td>GTPyPuAC</td>
<td>Pet I</td>
<td>CTGGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hha I</td>
<td>GCGC</td>
<td>Xho I</td>
<td>CTGGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hpa II</td>
<td>CCGG</td>
<td>Sst I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HinIII</td>
<td>AAGCTT</td>
<td>Xho II</td>
<td>PGATCPY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hpd I</td>
<td>PuGCCCpy</td>
<td>Hpa I</td>
<td>GACCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HindI</td>
<td>GTPyPuAC</td>
<td>Hpd I</td>
<td>GGTGA or TCACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BamH I</td>
<td>GAATCC</td>
<td>EcoRII</td>
<td>CCATTOG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EcoRI</td>
<td>GAATTTC</td>
<td>Hinfl</td>
<td>GANTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hpa II</td>
<td>GTTACG</td>
<td>Mnl I</td>
<td>CCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toq I</td>
<td>TCGA</td>
<td>Mbo I</td>
<td>GATC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smal I</td>
<td>CCCGGG</td>
<td>Mbo II</td>
<td>GAAGA or TCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bgl I</td>
<td>AAT I</td>
<td>Sst I</td>
<td>AGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsu I</td>
<td>AAGCTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ppu II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blu I</td>
<td>CTCGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sal I</td>
<td>GTGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rsal I</td>
<td>AGATCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xho I</td>
<td>CTCGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kpn I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sst II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Pu = purine, Py = pyrimidine.

---

**THE OVALBUMIN DNA CODING SEQUENCE**

**NUCLEOTIDES X 10^7**

[Diagram of the ovalbumin DNA coding sequence]

---

**Fig. 6.** Summary map of the sequence specific site for 11 restriction enzymes upon the ovalbumin DNA coding sequence.
presently using cDNA, "probes" made from Hae III-cut double-stranded cDNAs, to look at the fidelity of transcription of the 5' and 3' regions of the ovalbumin gene in chromatin in vitro. It is hoped that these studies will in the future define the fidelity of in vitro chromatin transcription and also provide us with more defined information on the mechanism of gene transcription in vivo.

Acknowledgments — We wish to thank Richard R. Roberts (Cold Spring Harbor Laboratory, New York) for very kindly providing us most of the highly purified enzymes required for this work and the facilities at his laboratory for carrying out much of these studies. We thank also Bernard Weisblum (University of Wisconsin, Madison, Wisconsin) for supplying the enzyme Smn I. We thank also Joseph W. Beard and Ed Houts (Life Sciences Inc., St. Petersburg, Florida) and Michael A. Chirigos (Office of Program Resources and Logistics, Viral Cancer Program, National Cancer Institute, Bethesda, Maryland) for purified AMV reverse transcriptase. We thank also Michael Farrell (University of North Carolina, Chapel Hill, North Carolina) for purified 3P-labeled SV40 DNA. We wish to thank our colleagues Larry A. McReynolds, Howard Towle, and Jeffrey Rosen for their critical comments and helpful discussion. The excellent technical assistance of Joyce Reed and Timothy Adams was much appreciated.

REFERENCES

Ovalbumin gene. Action of restriction endonucleases upon DNA coding sequence.

J J Monahan, S L Woo, C D Liarakos and B W O'Malley