The 5'-terminal structure of ovalbumin mRNA was determined after labeling by oxidation with NaIO4 and reduction with NaB[3H]I. The presence of 7-methylguanosine in pyrophosphate linkage was demonstrated by release of [3H]pm7G' (ribose-ring-opened) by tobacco acid pyrophosphatase. The principal compounds labeled were pm7G' and pA' in a ratio of 0.22:1. The discrepancy from 1:1 is believed to be due to strand scission of the mRNA followed by selective enrichment of 3'-terminal fragments by oligo(dT)-cellulose chromatography. When labeled mRNA was digested with RNase T1 and oligonucleotides separated by DEAE-Sephal chromatography, a major peak of radioactivity was observed which eluted in the position of m'Gppp(Np). This assignment was confirmed by similar digestion and chromatography of labeled rabbit α- and β-globin mRNAs, which contain the 5'-terminal structures m'Gppp(Np) and m7Gppp(Np), respectively. The major labeled T1 oligonucleotide of ovalbumin mRNA was digested with RNase T1 and yielded labeled peaks on DEAE-Sephal chromatography consistent with the structures m'Gppp(Np) (cap 0, 15%), m7GpppNmpNp (cap 1, 35%), and m7GpppNmpNmpNp (cap 2, 50%). Digestion of the material in each peak with nuclease P1, followed by high performance liquid anion exchange chromatography permitted identification of the cap 0 material as a mixture of m'Gpppm6A (17%) and m'GpppA (83%), the cap 1 material as a mixture of m'Gpppm6Am (49%) and m'GpppAm (51%), and the cap 2 materials as a mixture of m'Gpppm6Am (78%) and m7GpppAm (22%). Thus, structures consisting of 7-methylguanosine in an inverted 5'-to-5' triphosphate linkage (1). Most of these structures have been deduced using mRNA labeled with various precursors in vitro or in cell culture, or which has been labeled in vivo with [y-32P]ATP and polynucleotide kinase after removal of the cap. Previous attempts to label hen ovalbumin mRNA either with [32P]P04, or [methyl-3H]methionine in vivo or in organ culture have proven unsuccessful, as have attempts to specifically label the 5' terminus with polynucleotide kinase (2). We report here the use of a different methodology for labeling ovalbumin mRNA, involving oxidation with NaIO4 and reduction with NaB[3H]I, which has permitted us to determine the structure of the cap and 5'-terminal RNase T1 oligoribonucleotide of this mRNA.

**EXPERIMENTAL PROCEDURES**

**Chemicals and Enzymes—**Ribonucleases T1 and T2 were obtained from Calbiochem, and DEAE-Sephal was from Pharmacia. Buffer A contained 25 mM Tris-HCl, pH 7.6 at 25°C, and 7 M urea (Schwarz Mann) which had been deionized prior to addition with a mixed-bed ion exchanger. Sources of all other reagents and enzymes have been given previously (3).

**Purification of Ovalbumin mRNA—**White Leghorn laying hens (Dekalb strain) were purchased from Yeiser Egg Company, Winchester, Ky. Ovalbumin mRNA was prepared from oviduct polysomes by the following criteria: (i) a single band was observed on polyacrylamide gel electrophoresis (5); (ii) when the mRNA preparation was translated in a &lease-treated rabbit reticulocyte cell-free system (6), the radioactivity associated with the cap of this mRNA was labeled using mRNA labeled with various precursors in vivo with [3H]methylmethionine in vivo or in organ culture have proven unsuccessful (3), as have attempts to specifically label the 5' terminus with polynucleotide kinase (2). We report here the use of a different methodology for labeling ovalbumin mRNA, involving oxidation with NaIO4 and reduction with NaB[3H]I, which has permitted us to determine the structure of the cap and 5'-terminal RNase T1 oligoribonucleotide of this mRNA.

**Purification of Ovalbumin mRNA—**White Leghorn laying hens (Dekalb strain) were purchased from Yeiser Egg Company, Winchester, Ky. Ovalbumin mRNA was prepared from oviduct polysomes by the following criteria: (i) a single band was observed on polyacrylamide gel electrophoresis (5); (ii) when the mRNA preparation was translated in a &lease-treated rabbit reticulocyte cell-free system (6), the radioactivity associated with the cap of this mRNA was labeled using mRNA labeled with various precursors in vivo with [3H]methylmethionine in vivo or in organ culture have proven unsuccessful (3), as have attempts to specifically label the 5' terminus with polynucleotide kinase (2). We report here the use of a different methodology for labeling ovalbumin mRNA, involving oxidation with NaIO4 and reduction with NaB[3H]I, which has permitted us to determine the structure of the cap and 5'-terminal RNase T1 oligoribonucleotide of this mRNA.

A large number of eukaryotic cellular and viral messenger RNAs have been shown to contain a 5'-terminal "cap" structure consisting of 7-methylguanosine in an inverted 5'-to-5' triphosphate linkage (1). Most of these structures have been deduced using mRNA labeled with various precursors in vitro or in cell culture, or which has been labeled in vivo with [y-32P]ATP and polynucleotide kinase after removal of the cap. Previous attempts to label hen ovalbumin mRNA either with [32P]P04, or [methyl-3H]methionine in vivo or in organ culture have proven unsuccessful, as have attempts to specifically label the 5' terminus with polynucleotide kinase (2). We report here the use of a different methodology for labeling ovalbumin mRNA, involving oxidation with NaIO4 and reduction with NaB[3H]I, which has permitted us to determine the structure of the cap and 5'-terminal RNase T1 oligoribonucleotide of this mRNA.

**Purification of Ovalbumin mRNA—**White Leghorn laying hens (Dekalb strain) were purchased from Yeiser Egg Company, Winchester, Ky. Ovalbumin mRNA was prepared from oviduct polysomes by the following criteria: (i) a single band was observed on polyacrylamide gel electrophoresis (5); (ii) when the mRNA preparation was translated in a &lease-treated rabbit reticulocyte cell-free system (6), the radioactivity associated with the cap of this mRNA was labeled using mRNA labeled with various precursors in vivo with [3H]methylmethionine in vivo or in organ culture have proven unsuccessful (3), as have attempts to specifically label the 5' terminus with polynucleotide kinase (2). We report here the use of a different methodology for labeling ovalbumin mRNA, involving oxidation with NaIO4 and reduction with NaB[3H]I, which has permitted us to determine the structure of the cap and 5'-terminal RNase T1 oligoribonucleotide of this mRNA.
Enzymatic Digestions of Ovalbumin [3H]mRNA—RNase T1 was added at 100 μg/ml to reaction mixtures containing 2 to 10 μl of ovalbumin [3H]mRNA, 1 mg/ml of unlabeled oviduct ribosomal RNA, 10 mM Tris-HCl, pH 7.4, and 1 mM EDTA. Incubation was at 37°C for 3 h. RNase T1 was added at 50 units/ml to reaction mixtures (80 μl) containing 1 to 5 × 10^9 cpm of ovalbumin [3H]mRNA. The column was eluted with a 200-ml linear gradient from 30 to 400 mM NaCl in Buffer A. The column then was washed with 30 ml of Buffer A containing 2 mM NaCl to elute any strongly bound material. Before chromatography, RNase T1 digests of ovalbumin [3H]mRNA on DEAE-Sepharose, the sample was mixed with markers consisting of AMP, ADP, and ATP (200 nmol each) or alternatively, was extracted twice with phenol/chloroform (9) and once with ether and then added to 0.5 mg of rRNA which had been digested with RNase T1. Pooled fractions from DEAE-columns were desalted on a Sephadex G-10 column (0.9 × 65 cm) equilibrated in 10 mM ammonium acetate, pH 4.5. Incubation was at 37°C for 5 h. Digestion conditions for nuclease P1 and tobacco acid pyrophosphatase have been described previously (3). Digestion with RNase A was similar to that with RNase T1; except enzyme was present at 200 μg/ml.

Chromatography on DEAE-Sepharose—The RNase T1 and A digests of ovalbumin [3H]mRNA were diluted 2-fold with Buffer A containing 30 mM NaCl and applied to a DEAE-Sepharose column (1.0 × 25 cm). Following a 25-ml wash with Buffer A containing 30 mM NaCl, the column was eluted with a 200-ml linear gradient from 30 to 400 mM NaCl in Buffer A. The column then was washed with 30 ml of Buffer A containing 2 mM NaCl to elute any strongly bound material. Before chromatography of RNase T1 digest of ovalbumin [3H]mRNA on DEAE-Sepharose, the sample was mixed with markers consisting of AMP, ADP, and ATP (200 nmol each) or alternatively, was extracted twice with phenol/chloroform (9) and once with ether and then added to 0.5 mg of rRNA which had been digested with RNase T1. Pooled fractions from DEAE-columns were desalted on a Sephadex G-10 column (0.9 × 65 cm) equilibrated in 10 mM ammonium acetate, pH 4.5. Incubation was at 37°C for 5 h. Digestion conditions for nuclease P1 and tobacco acid pyrophosphatase have been described previously (3). Digestion with RNase A was similar to that with RNase T1; except enzyme was present at 200 μg/ml.

RESULTS

Labeling Ovalbumin mRNA with NaBr3H14—The 5' and 3' termini of capped eukaryotic mRNAs contain cis-diol and trans-diol groups and are, therefore, subject to oxidation with NaIO4 and reduction of the dioldehydes with NaBH4, resulting in the incorporation of 3H into four nonexchangeable positions/mRNA molecule. Radioactivity is also incorporated into positions which exchange with H2O, presumably because of initial exposure of the RNA to H2O of high specific activity which is formed during the labeling reaction (10). Initially this exchangeable radioactivity represents an amount approximately equal to the stable radioactivity, but after a period of days it is lost. Commercial preparations of NaBH4 are contaminated with a small amount (0.2%) of acid-stable radioactive compounds, some of which are nondialyzable, polyatomic (11), and precipitable with ethanol. Although this is a small portion of the total radioactivity in the borohydride preparation, it still represents 10-fold more radioactivity than the maximum which could be incorporated into mRNA under our standard reaction conditions. Therefore, it was necessary to purify the labeled mRNA from this contaminating radioactive material by affinity chromatography on oligo(dT)-cellulose. Typically, a labeling reaction containing 20 μg of mRNA yielded 1.3 × 10^9 cpm bound to oligo(dT)-cellulose. This represented 74% of the theoretical incorporation based on the A260 of the mRNA preparation (assuming 25 A260/μg) and the specific activity of the NaBH4 (51.4 Ci/mmol). Presence of 7-Methylguanosine in Ovalbumin mRNA—Labeled ovalbumin mRNA was treated with tobacco acid pyrophosphatase and analyzed on HPLC using isocratic elution with 1 mM potassium phosphate buffer, pH 3.5. A single peak of radioactivity eluted with the pm'G' standard (Fig. 1B). No radioactivity eluted from the column when tobacco acid pyrophosphatase was absent from the incubation mixture (Fig. 1A). To identify other labeled compounds, ovalbumin mRNA was hydrolyzed to nucleoside 5'-monophosphates with nuclease P1, and analyzed by HPLC using isocratic elution with 30 mM potassium phosphate buffer, pH 3.5 (Fig. 1C). The major peak of radioactivity, eluting with a standard for pA', represents labeling of the 3'-terminal poly(A) portion of the mRNA. No pm'G' was observed in this chromatogram since this material was presumably present in structures of the form m7GpppN(m), which are resistant to nuclease P1 and are retained on the column under these conditions. When labeled ovalbumin mRNA was incubated with nuclease P1 plus tobacco acid pyrophosphatase (Fig. 1D), an additional

1 The abbreviations used are: HPLC, high performance liquid chromatography; pm'G, 7-methylguanosine 5'-phosphate; pA', ribose-ring-opened derivative of 5'-AMP; pm'G', ribose-ring-opened derivative of 5'-AMP; pm'GpppN, 7-methylguanosine linked to any nucleoside by a 5'-triphosphate linkage; Y, a pyrimidine nucleoside; (m), partial methylation.

2 The abbreviations used are: HPLC, high performance liquid chromatography; pm'G, 7-methylguanosine 5'-phosphate; pA', ribose-ring-opened derivative of 5'-AMP; pm'G', ribose-ring-opened derivative of 7-methylguanosine 5'-phosphate; m7GpppN, 7-methylguanosine linked to any nucleoside by a 5'-triphosphate linkage; Y, a pyrimidine nucleoside; (m), partial methylation.
peak of radioactivity was observed (cf. Fig. 1C) eluting with the pmG' standard. The radioactivity in pmG' represented 22% of that found in pA'. The incorporation of radioactivity into other nucleotides, primarily pU, was low, representing nonspecific labeling of about one nucleotide in 8000. This is considerably lower than previously reported (12). The difference in $^3$H$_2$O between Fig. 1, A and B, compared with Fig. 1, C and D, is due to the former experiment being performed within 2 h of oligo(dT)-cellulose chromatography, as opposed to the overnight incubation used with nuclease P$_1$. The shorter time permits less equilibration of tritium from exchangeable positions in the RNA with the medium.

The cause of the low ratio of pmG' to pA' was investigated. One possibility was that conditions for NaIO$_4$ oxidation were not optimal. In preliminary experiments, it was shown spectrophotometrically that cap analogues of the form m$^7$GpppNp react with NaIO$_4$ more slowly than nucleoside 5'-monophosphates. Thus, it was conceivable that oxidation could be complete at the 3' terminus but incomplete at the 5' terminus. To test this, the time of oxidation was varied, and the amount of 5'-terminal labeling was determined by treatment of the RNA with tobacco acid pyrophosphatase and quantitation of the released $[^3$H]pmG' by HPLC. Identical labeling was found at 1.5 and 3 h, but at 12 h, the amount of pmG' was substantially reduced. Furthermore, model reactions with cap analogues were complete within 3 h. Thus, it seemed unlikely that incomplete reaction of 5'-termini with NaIO$_4$ was occurring. Another possibility was that overoxidation and $\beta$ elimination of m$^G$ was occurring, due to an excessive molar ratio of NaIO$_4$ to mRNA. However, when the concentration of NaIO$_4$ was lowered from 2.5 to 0.4 mm, the amount of labeled pmG' decreased rather than increased, making overoxidation unlikely. A third possibility which could account for low 5'- to 3'-terminal labeling was that the mRNA preparation was subjected to random strand scission, either within the cell, during purification, or during the labeling reaction itself, and then only the 3'-terminal fragments were co-purified with intact molecules by the oligo(dT)-cellulose chromatography. Supporting this hypothesis is our previous finding (3) that the 5'- to 3'-terminal labeling ratio is 0.66 in rabbit globin mRNA, an mRNA which is isolated from a nuclease-free tissue and which also has only one-third the target size of ovalbumin mRNA. Further evidence was obtained by electrophoresis of the ovalbumin $[^3$H]mRNA preparation on methylmercury/agarose (13), cutting 1-mm slices through the ovalbumin $[^3$H]mRNA region, and determining ratios of pmG'/pA' in the material eluted from each slice. From lowest molecular weight to highest, the ratios were 0.25, 0.42, 0.48, and 0.86. These results are consistent with the presence of uncapped, 3'-terminal fragments of ovalbumin mRNA in our preparation.

5' Terminal T$_1$ Oligonucleotide of Ovalbumin mRNA—When the cap dinucleotides resulting from a P$_1$ digest of labeled ovalbumin mRNA were analyzed by HPLC, a complex pattern was observed. The compounds eluted in the positions of known cap standards and were hydrolyzed by tobacco acid pyrophosphatase (Fig. 4A), but precise identification and quantitation was not possible due to incomplete resolution. To simplify the HPLC pattern, labeled mRNA was hydrolyzed with RNase T$_1$, and the major oligonucleotide was purified from the other oligonucleotides and poly(A) by DEAE-Sepharose column chromatography (Fig. 2A). The major peak, accounting for 50% of the radioactivity associated with pmG', eluted with a charge of -10.5, which is the predicted charge for an oligonucleotide of structure m$^7$Gppp(Np)$_m$. To confirm this assignment, the 5'-termini of rabbit $\alpha$- and $\beta$-globin mRNA, which have the structures m$^7$Gppp(Np)$_m$, respectively (14, 15), were used as markers.

When the RNAse T$_1$ digest of labeled globin mRNA was analyzed by DEAE-Sepharose chromatography, two major peaks were observed eluting with the predicted charges (Fig. 2B). The major peak of the ovalbumin RNAse T$_1$ digest eluted at a position intermediate between the peaks for the 5'-terminal T$_1$ oligonucleotides of $\alpha$- and $\beta$-globin mRNA. Thus,
the major labeled oligonucleotide in the ovalbumin RNase T₁ digest has a structure with the same charge as m'Gppp(Np)₅m. Not shown in the figures are peaks which eluted with 2 M NaCl, corresponding to the poly(A) portion of each mRNA, which accounted for 68 and 40% of the total radioactivity for ovalbumin and globin mRNA, respectively.

2'-O-Methylation of Ovalbumin mRNA—The distribution of 2'-O-methyl groups in the cap of ovalbumin mRNA was determined by digesting the major T₁ oligonucleotide (Fig. 2A, Fractions 115 to 119) with RNase T₂, which does not hydrolyze phosphodiester bonds adjacent to 2'-O-methyl groups. The digest was analyzed by chromatography on DEAE-Sephacel (Fig. 3). Three major peaks of radioactivity were observed, eluting at positions on the gradient corresponding to charges of -3.9, -4.6, and -5.4. These charges are consistent with the structures m'Gppp(Np)₅m (cap 0), m'GpppNmpNp (cap 1), and m'GpppNmpNmpNp (cap 2). In the latter two cases, these agree with the charge assignments made to the cap 1 and cap 2 structures of NaB[³H]₄-labeled rabbit globin mRNA (14). However, no cap 0 compounds were detected in globin mRNA. The ratios of cap 0:cap 1:cap 2 were approximately 1:2.5:3.5.

Identification of Cap Dinucleotides—Each cap species generated by hydrolysis of the major T₁ oligonucleotide with RNase T₂ and separated by DEAE-Sephacel chromatography (Fig. 3) was hydrolyzed to cap dinucleotides with nuclease P₁, and identified by HPLC. The cap 0 compound yielded a major peak which eluted with the dinucleotide m'GpppA (Fig. 4B). A small shoulder on the main peak eluted between m'GpppG and m'GpppG. a position on the gradient where m'GpppM₄m would be expected to elute. Since we lack the m'GpppM₄m standard at present, assignment remains tentative. The cap 1 material yielded two peaks of approximately equal radioactivity upon P₁ digestion, which eluted with the dinucleotide standards m'GpppAm and m'GpppM₄m (Fig. 4C). The cap 2 material yielded the same two peaks of radioactivity as the cap 1 material, but in this case m'GpppM₄Am predominated over m'GpppAm (Fig. 4D).

Digestion of the Major T₁: Oligonucleotide with Ribonuclease A—If the general structure of a capped oligonucleotide is represented as m'GpppN₅m(m)pₙ(m)p₂ₙ(m)pₙPₚₙ ..., the preceding results have shown that the N' base consists of a mix of A and m'A. If the N₂ base were a pyrimidine, oligonucleotides terminated with cap 0 or cap 1 structures would be hydrolyzed by RNase A to yield m'GpppN(m)pₙNp. When the major T₁ oligonucleotide was subjected to RNase A digestion and DEAE-Sephacel chromatography, approximately half of the radioactivity eluted with a charge consistent with this structure (Fig. 5). Oligonucleotides terminated with cap 2 structures, on the other hand, should be hydrolyzed at the first pyrimidine residue after N'. The other half of the radioactivity in the major T₁ oligonucleotide eluted in the position of m'GpppN₅m(m)pₙ(m)p₂ₙ(m)pₙPₚₙ, indicating that N₅ is a pyrimidine (Fig. 5). Since RNase T₁ did not cleave at N₅, that residue must be adenosine. Thus, the structure of the oligonucleotide is m'Gppp(m₅A)(m)pₙY(m)pₙApYpNₜ(m)pₙGₚ, where Y represents a pyrimidine nucleoside. (We are assuming that the cap 0, 1, and 2 structures found in this peak are all derived from the same oligonucleotide.)

Relative Amounts of Adenosine-containing Caps—The quantities of the six cap structures identified in the major T₁ oligonucleotide of ovalbumin mRNA are presented in Table I, expressed as a percentage of the total caps present in that oligonucleotide. The most highly methylated structure, m'GpppM₄m(m)pₙYmₜ(m)pₙApYₜ(m)pₙGₚ, is the most abundant form. The ratio of m₅A/A is low for caps with no 2'-O-methylations, becomes unity when one 2'-O-methyl group is present, and is

![Fig. 4. HPLC analysis of ovalbumin [³H]mRNA digested with nuclease P₁. Arrowheads represent the positions of unlabeled standard compounds injected simultaneously. Radioactivity was determined in whole fractions of 300 to 400 µl to which 4 ml of Instagel (Packard) were added. The compounds derived from DEAE-Sephacel chromatography were desalted as described under "Experimental Procedures" prior to nuclease P₁ digestion. A, unfractionated ovalbumin mRNA (6.3 x 10⁴ cpm of which 1.1 x 10⁴ cpm were in pm'G₈); B, cap 0 compounds (Fractions 78 to 82, Fig. 3), representing 1.4 x 10⁴ cpm; C, cap 1 compounds (Fractions 85 to 89, Fig. 3), representing 4 x 10⁴ cpm; D, cap 2 compounds (Fractions 92 to 97, Fig. 3), representing 4.8 x 10⁴ cpm. The cap dinucleotide standard m'GpppNm was not present in the HPLC analysis of cap 1 and 2, but other experiments have shown that no radioactivity in the major 5' terminal T₁ oligonucleotide elutes with this standard.](image1)

![Fig. 5. DEAE Sephacel column chromatography of the major T₁ oligonucleotide of ovalbumin [³H]mRNA digested with RNase A. Fractions 115 to 119 from a DEAE-Sephacel column run similar to the one shown in Fig. 2A were desalted and concentrated as described above. The recovered material (1.4 x 10⁴ cpm) was digested with RNase A and separated on DEAE-Sephacel chromatography. Aliquots (300 µl) of the 1.0-ml fractions were counted to determine radioactivity. Arrowheads give the positions of oligonucleotide markers with the charges indicated contained in the RNase A digest. Details on digestion with RNase A are given under "Experimental Procedures."](image2)
significant greater than one when two 2'-0-methyl groups are present. Thus, the degree of base methylation increases with increasing 2'-0-methylation. Cap 2 structures accounted for roughly half of the total, which is in agreement with the finding that equal quantities of mGppp(Np)2 and mGppp(Np)3 were produced by RNase A digestion (Fig. 5).

Occurrence of Other Cap Structures in Ovalbumin mRNA—The profile of radioactivity on DEAE-Sephacel chromatography of RNase T1-digested ovalbumin mRNA revealed a second prominent peak (Fig. 2A, Fraction 78) which eluted at approximately the position of a cap 0 compound (cf. Fig. 3). Of the total radioactivity which was present in ovalbumin mRNA as mG (Fig. 1), 20% was represented by this putative cap 0 compound and 50% by the major peak which we have identified as mGppm(m'5)A(m)pY(m)pApY(p)G(Np)2Gp. Since this putative cap 0 compound was released by RNase T1, it should have the structure mGpppGp. Fractions 75 to 83 from Fig. 2A were pooled, desalted, digested with nuclease P1, and analyzed on HPLC. A major portion of the radioactivity eluted with the mGpppG standard (Fig. 6). The fact that this material was not degraded to nucleotides by nuclease P1 strengthens the assignment of its structure as mGpppG. A large peak of radioactivity was observed in the region where nucleotides elute which was not observed in Fig. 4, B to D. This was presumably derived from nonspecifically labeled compounds of the form NpNpNp which eluted in the same position as mGpppGp from the DEAE-Sephacel column.

Table 1

<table>
<thead>
<tr>
<th>Structure</th>
<th>Percentage of total</th>
<th>Ratio: m'A/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>mGpppApYpApYp(Np)2Gp</td>
<td>14.4</td>
<td>0.20</td>
</tr>
<tr>
<td>mGpppmApYpApYp(Np)2Gp</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Cap 1</td>
<td>35.3</td>
<td>0.96</td>
</tr>
<tr>
<td>mGpppAmpYpApYp(Np)2Gp</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>mGpppmAmpYpApYp(Np)2Gp</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>Cap 2</td>
<td>50.3</td>
<td>3.5</td>
</tr>
<tr>
<td>mGpppAmpYmpApYp(Np)2Gp</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>mGpppmAmpYmpApYp(Np)2Gp</td>
<td>39.1</td>
<td></td>
</tr>
</tbody>
</table>

* Relative occurrence of a particular structure in the major 5'-terminal T1 oligonucleotide of ovalbumin mRNA.

In an earlier study, we demonstrated the feasibility of using NaB[3H]4 for labeling mRNA cap structures by confirming the structure of rabbit globin mRNA as m'Gpppm'Am (3). The present work represents the first instance in which an unknown cap structure has been elucidated by this method, although others have used NaB[3H]4 labeling to demonstrate the presence of mG or unspecified cap 1 and cap 2 structures in mRNA (14, 16, 17). The method has certain advantages over labeling with polynucleotide kinase and [γ-32P]ATP. First, for reasons that are still not clear to us, ovalbumin is labeled to only a small degree with polynucleotide kinase (2). Second, it is not necessary to remove pppm7G before labeling. Third, the efficiency of labeling with polynucleotide kinase varies with the nature of the 5'-terminal nucleoside (18), the presence of 2'-0-methylations (19), and the secondary structure of the nucleic acid (20), whereas the reactions with NaI04 and NaBH4 proceed to completion when model compounds are used, regardless of the nature of the penultimate base (3).

A heterogeneous mixture of 5'-terminal structures, such as those found in ovalbumin mRNA, would not be expected to be labeled uniformly by polynucleotide kinase. The NaB[3H]4 labeling method also has advantages over the use of [methyl-3H]methionine since not all messengers are amenable to labeling in culture. Also, reduced levels of methionine, necessary to obtain high specific activities of precursors, could lead to incomplete methylation of mRNA. Thus, the distribution of labeled structures obtained with methionine labeling would not reflect true steady state levels of the different cap structures.

Six distinct cap structures in which adenosine was the penultimate nucleoside were found (Table I). The most abundant cap was the most highly methylated structure. Somewhat unexpected was the finding of cap 0 structures (Figs. 3, 4B, and 6), as these have previously been reported only in the cytoplasmic mRNA of yeast (21) and slime mold (22) and in the virion RNA of plant viruses (23-26). The finding of cap 0 structures in polysomes supports our earlier conclusion from studies with methylation inhibitors that 2'-0-methylation is not necessary for transport of mRNA from the nucleus or initiation of translation (27). Two other structures potentially could be present in ovalbumin mRNA, mGpppApYmpAp and mGpppmApYmpAp, but these would not have been distinguished from cap 0 structures by our methodology. It should be noted that these six structures represent polysomal mRNA only, due to our method of mRNA preparation. Differential stability of mRNA in vivo or differential selection of cap structures by the protein synthetic initiation machinery could produce a spectrum of cap structures in polysomes which differs from that of newly transcribed mRNA. A striking correlation between ribose methylation and base methylation was observed (Table I). Our observation that cap 1 structures contain more m'A than cap 0 structures is in agreement with the finding of Keith et al. (28) that 2'-0-methylated nucleosides are better substrates for the 6-methylase. The interpretation of our finding that cap 2 structures contain, in turn, more m'A than cap 1 structures is unclear since it appears that the 2'-0-methylation which forms cap 2 structures occurs in the cytoplasm whereas 6-methylation occurs in the nucleus (29). Nonetheless, since all possible patterns of methylation were found, it can be concluded that there is no obligatory order of methylation.

The data presented here allow us to locate the nucleotide sequence of ovalbumin mRNA reported by McReynolds et al. (30) relative to the cap. The evidence strongly suggests that the partially methylated adenosine residue we find in the penultimate position of the mRNA is the same as the first
adenosine residue in their sequence. The identification of the major 5'-terminal T1 fragment as m'Gppp(Np)Gp (Fig. 2A) corresponds to the presence of guanosine in the eighth position of their sequence. Furthermore, the pattern obtained with RNase A allows assignment of a pyrimidine nucleoside in the second position, adenosine in the third, and a pyrimidine in the fourth (Fig. 5), which corresponds to their sequence of A-C-A-U-A-C-A-G.

The occurrence of m'GpppGp as an RNase T1 digestion product of ovalbumin mRNA (Figs. 2A and 6) is puzzling, as it represents 20% of the total radioactivity in m'G. This is a large fraction when it is recalled that the sum of all adenosine-containing caps is only 50% of the radioactivity in m'G. Some of the remaining 30% of the total radioactivity in m'G may be present as m'GpppGm, as this compound was detected in unfractionated ovalbumin mRNA (Fig. 4A) but not in the major T1 oligonucleotide. The purity of our ovalbumin mRNA preparation is estimated to be at least 90%, so it is unlikely that the guanosine-containing caps are contributed by contaminants. Furthermore, when [3H]ovalbumin was subjected to methylmercury/agarose electrophoresis (see "Results"), material eluted from each slice throughout the ovalbumin mRNA region exhibited the same pattern of cap structures as shown in Fig. 4A, which includes m'GpppG and m'GpppGm. Since the sequence of the ovalbumin gene in the region of the promoter has recently been published as

5'-CTGTACATACAG-3'
3'-GACATGTATCCTC-5' (31), where the asterisk denotes the putative start of transcription, it is possible to speculate that transcription of some ovalbumin mRNA molecules may begin 2 residues to the left of the asterisk (upstream). These molecules would contain two additional residues at the 5' terminus and begin with either m'GpppG or m'GpppGm. The first of these would give rise to the peak at Fraction 78 in Fig. 2A, which we have identified as m'GpppGp. The second would give rise to a structure of the form m'Gppp(Np)10 upon digestion with RNase T1. A small peak eluting in the position of such a structure can be seen in Fig. 2A, fraction 123. The existence of two populations of ovalbumin mRNA molecules differing by 2 nucleoside residues would also explain the finding of McReynolds et al. (30) that reverse transcriptase elongates a DNA primer hybridized to ovalbumin mRNA to a "strong stop" position, corresponding to the 5'-terminal adenosine residue in the published sequence, but then further elongates a small percentage of molecules by 1 to 3 residues. Another possibility is allelism of the ovalbumin gene (32-34). Unequivocal proof that two populations of ovalbumin mRNA molecules exist, however, will require further investigation.

Acknowledgments—We are indebted to Drs. Wyman Dorough and Robert Lester for the use of their HPLC equipment and to Ms. Linda Berardino for expert secretarial assistance.

REFERENCES

27. Kaehler, M., Coward, J. & Rottman, F. (1979) Nucleic Acids Res. 6, 1161-1176