Predicted Structures of Apolipoprotein II mRNA Constrained by Nuclease and Dimethyl Sulfate Reactivity: Stable Secondary Structures Occur Predominantly in Local Domains via Intraexonic Base Pairing*

(Received for publication, October 31, 1988)

Sheng-Ping L. Hwang‡§, Moisés Eisenberg‡, Roberta Binder‡¶, Gregory S. Shelnutt¶, and David L. Williams‡**

From the §Department of Pharmacological Sciences and the ¶Molecular Biology Graduate Program, Health Sciences Center, State University of New York, Stony Brook, New York 11794

Analyses of apolipoprotein II mRNA with chemical and enzymatic probes showed that double- and single-stranded regions were distributed uniformly along the mRNA except for a large (72 nucleotides) single-stranded region containing the translation stop codon. Secondary structure models constrained by the experimental data were made by varying the distance (along the mRNA) over which base pairing was allowed. Four prominent secondary structures were seen with restrictions of 165, 330, or 659 nucleotides suggesting that such structures form via local interactions over distances of 50–120 nucleotides. Predicted long range interactions involve only 2–3 base pairs while local interactions involve helices of 4–10 base pairs. Predicted helices of ≥ 4 base pairs occur primarily within exons, raising the possibility that prominent secondary structures in mRNAs may be largely due to intraexonic base pairing. Tests of single- and double-stranded domains by oligonucleotide-directed RNase H cleavage and primer extension were in accord with the structure model and with nuclease and chemical modification data. The model predicting base pairing between the coding and the 3' noncoding regions was tested by RNase H cleavage followed by oligo(dt)-cellulose chromatography to separate 5' and 3' mRNA fragments. Most (82%) of the 5' fragment remained associated with the 3' noncoding region in a structure with a t1/2 = 50 °C in 0.2 M Na+. Predicted long range interactions involving two 24 nucleotides stems and a 101 nucleotide stem remained associated with the 3' noncoding region in a structure with a t1/2 = 50 °C in 0.2 M Na+. In the medium of serum albumin mRNA is not altered by hormone withdrawal and the decay rate of another estrogen-regulated mRNA, apob mRNA, is unaffected by the duration of estrogen treatment. These results suggest that apoII and VTG II mRNAs selectively targeted for decay by an estrogen-sensitive process.

Little is known about the mechanisms which account for selective mRNA degradation or how such mechanisms are regulated by hormones or other agents. Selective mRNA decay likely involves factors which recognize sequence or structural features of the mRNA. Various lines of evidence suggest that RNA secondary structure may play a decisive role in the regulation of mRNA turnover. In bacteria and bacteriophage, RNase III cleavage at a specific "RNase III site," which can potentially form a stem-loop structure, can regulate mRNA stability (Schmeissner et al., 1984; Panayotatos and Truong, 1985; Takata et al., 1987). In eukaryotic systems, a conserved stem-loop structure in histone mRNA is required for cleavage of the primary transcript for correct 3' end formation (Birchmeier et al., 1983). This same stem-loop structure in cytoplasmic histone mRNA may serve as the initial site of cleavage for degradation which proceeds in a 3' to 5' direction (Ross et al., 1986; Ross and Kobs, 1986; Levine et al., 1987; Pandey and Marzluff, 1987). In human immunodeficiency virus mRNA, it has been shown that a stable stem-loop structure formed in the 5' leader region may interact with a trans-activator protein to promote the accumulation of this mRNA (Muensing et al., 1987). These and other examples suggest that efforts to understand mechanisms of mRNA decay may require an understanding of the various regulatory mechanisms which control mRNA stability.

In addition to regulating gene expression at the level of transcription, estrogens are known to alter the turnover of cytoplasmic mRNAs (Darnell et al., 1986; Shapiro et al., 1987). One system in which altered mRNA turnover occurs is the avian liver in which apoI mRNA and VTG II mRNA are expressed at high levels in laying hens or in roosters or immature birds after the administration of exogenous estrogen (Bergink et al., 1974; Chan et al., 1976). Recent studies have shown that the degradation rates of apoII and VTG II mRNAs during hormone withdrawal are dramatically altered by the duration of prior estrogen treatment (Gordon et al., 1988). When hormone is withdrawn after 1 day of estrogen treatment, these mRNAs decay with a t1/2 = 13 h. After 14 days of estrogen treatment, apoII and VTG II mRNAs decay with a t1/2 = 1.5 h when hormone is withdrawn. This increase in the turnover rates of apoII and VTG II mRNAs is specific since the steady state level of serum albumin mRNA is not altered by hormone withdrawal and the decay rate of another estrogen-regulated mRNA, apob mRNA, is unaffected by the duration of prior estrogen treatment (Gordon et al., 1988). These results suggest that apoII and VTG II mRNAs are selectively targeted for decay by an estrogen-sensitive process.

The abbreviations used are: apo, apolipoprotein; VTG, vitellogenin; CV, cobra venom; oligo, oligonucleotide.

* This work was supported in part by Grant DK18171 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by National Institutes of Health Predoctoral Award (National Research Service Award) GM 08065.
‡ Supported by National Institutes of Health Predoctoral Award (National Research Service Award) GM 07518. Current address: Laboratory of Cell Biology, Rockefeller University.
§ To whom correspondence should be addressed. Tel.: 516-444-3078.

1 The abbreviations used are: apo, apolipoprotein; VTG, vitellogenin; CV, cobra venom; oligo, oligonucleotide.
standing of mRNA secondary structure.

We have described previously an approach for high resolution structure mapping of a specific mRNA in a complex population (Shelness and Williams, 1985). In combination with limited cleavage by base and structure specific nucleases, primer extension was used to map possible nucleotides in the 3' noncoding region of apoII mRNA. When constrained by RNase T1 cleavage data, RNA folding rules suggest the occurrence of a complex stem-loop structure. In the present study we have extended this analysis to the entire length of the mRNA using both chemical and enzymatic probes of RNA structure. Secondary structure models constrained by experimental data were generated by varying the distance (along the mRNA) over which base pairing could occur. The results suggest that prominent secondary structures consist of domains formed from relatively local interactions involving nucleotides within 50-120 residues along the mRNA chain. In addition, 12 of 13 predicted helices of ≥4 base pairs involved base pairing within exons, raising the possibility that stable secondary structure in mRNAs may be largely due to intranucleon base pairing. A predicted stem-loop structure involving the coding region and 3' noncoding region was experimentally tested by oligonucleotide-directed RNase H cleavage followed by oligo(dT)-cellulose chromatography. The results were consistent with the proposed structure. Within the 3' noncoding region, the complex stem-loop structure noted previously (Shelness and Williams, 1985) was also present in the current models. This stem-loop structure contains endonuclease target sites for apoII mRNA degradation suggesting that secondary structure may play a role in apoII mRNA turnover.2

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase, phenol, cesium chloride, and RNase H were obtained from Bethesda Research Laboratories or New England Biolabs. S1 nuclease, dioxyonucleotides, and deoxyonucleotides were obtained from Pharmacia LKB Biotechnology Inc. RNase T1 was obtained from Boehringer Mannheim. RNase T2 and F285 and cobram RNase were obtained from Behring Diagnostics. Avian myeloblastosis virus reverse transcriptase was obtained from Life Sciences, Inc. Dimethyl sulfate, 3-mercaptoethanol, yeast tRNA, morpholinopropanesulfonic acids, 17P-estradiol, and guanidinium isothiocyanate (GTC) and sodium N-lauryl sarcosine (30%) were obtained from Fluka Chemical Corp. [T-32P]ATP (3000 Ci/mmol) and [a-32P]dCTP (800 Ci/mmol) were purchased from Amersham Corp. Dimethyl sulfate methylation, RNA (0.65 mg/ml) was equilibrated for 5 min at 37°C in dimethyl sulfate buffer (0.15 M NaCl, 0.005 M MgCl2, 0.001 M EDTA, 0.05 M sodium cacodylate, pH 7.5). A dilution series of dimethyl sulfate (3.4-39.6 μM) in dimethyl sulfate buffer was prepared and kept on ice. Reactions were initiated by adding various concentrations of dimethyl sulfate to RNA and terminated after 10 min at 37°C by adding β-mercaptoethanol to a final concentration of 0.2 M. The modified RNAs were adjusted to 0.3 M sodium acetate, precipitated with 2.5 volumes of ethanol at −70°C for 1 h, washed with ethanol, dried, and redissolved in 50 μl of sterile deionized water.

RNA Isolation—RNA samples were isolated by extraction with phenol-chloroform extractions as described (Protter et al., 1982).

Enzyme and Chemical Treatment

RNA Isolation—Total liver RNA was isolated with the guanidinium isothiocyanate-ethanol precipitation procedure followed by phenol-chloroform extractions as described (Hilfiker et al., 1982).

RNA Digestion—Total liver RNA (2.8 mg/ml) was digested with RNase A (0.2 M NaCl, 0.005 M MgCl2, 0.001 M dithiobisreitol, 0.025 M Tris-HCl, pH 7.5) containing oligo(B) (in a 270-fold molar excess to the estimated amount of apoII mRNA) in a final volume of 40 μl. Digestion was initiated by adding RNase H (0.12 unit/μl) and terminated after 45 min at 37°C by adding sodium dodecyl sulfate to a final concentration of 0.1%. The sample was split into two tubes. To one tube, 1.5 ml of binding buffer (0.2 M NaCl, 0.005 M MgCl2, 0.001 M EDTA, 0.1% sodium dodecyl sulfate, 0.025 M Tris-HCl, pH 7.5) was added, and the sample was applied to the oligo(dT)-cellulose column. To the other tube, morpholinopropanesulfonic acids, pH 7.5, and EDTA were added to final concentrations of 0.1 and 0.02 μM, respectively. This sample was incubated at 100°C for 1 min. Binding buffer (1.5 ml) was added rapidly, and the sample was applied to the oligo(dT)-cellulose column. Mock digestion procedures were the same except that an appropriate amount of sterile deionized water was used instead of RNase H. For RNA melting experiments, total liver RNA (3.5 mg/ml) was equilibrated at 32°C for 5 min in RNase H buffer containing oligo(B) and digested with RNase H for 2 h at 32°C. The reaction was terminated with sodium dodecyl sulfate as above. Digested RNA samples were incubated at various temperatures (32, 37, 42, 47, 50, 55, 60, 65, and 100°C) for 15 min prior to loading onto the oligo(dT)-cellulose column.

Oligo(dT)-cellulose Chromatography—RNA samples in binding buffer were applied to the column, and the flow-through was reapplied seven more times to ensure maximal binding. Unbound poly(A)-RNA was collected by washing with 5 column volumes of binding buffer. Poly(A)-RNA fractions were collected with 5 column volumes of elution buffer (0.1% sodium dodecyl sulfate, 0.025 M Tris-HCl, pH 7.5). An appropriate amount of yeast tRNA was added to both poly(A)-RNA and poly(A)+RNA to obtain the same final mass of total RNA. Sodium acetate was added to 0.3 M, 2.5 volumes of ethanol were added, and RNA samples were precipitated −70°C for 1 h.

Energy Minimized Secondary Structures—apoII mRNA was folded into a secondary structure by means of a computer program provided

---

**TABLE I**

<table>
<thead>
<tr>
<th>DNA oligonucleotides</th>
<th>Nucleotide coordinates</th>
<th>RNase H cleavage</th>
<th>Primer extension</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>70-90</td>
<td>N.D.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>250-300</td>
<td>N.D.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>261-280</td>
<td>N.D.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>323-333</td>
<td>N.D.</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>334-344</td>
<td>N.D.</td>
<td>+</td>
<td>SS</td>
</tr>
<tr>
<td>B</td>
<td>380-399</td>
<td>N.D.</td>
<td>N.D.</td>
<td>SS</td>
</tr>
<tr>
<td>H</td>
<td>402-411</td>
<td>N.D.</td>
<td>N.D.</td>
<td>SS</td>
</tr>
<tr>
<td>K</td>
<td>414-430</td>
<td>-</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>435-444</td>
<td>-</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>446-468</td>
<td>N.D.</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>545-567</td>
<td>N.D.</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>623-653</td>
<td>N.D.</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>643-664</td>
<td>+</td>
<td>SS</td>
<td></td>
</tr>
</tbody>
</table>
by Dr. M. Zuker (Zuker and Stiegler, 1981). This program was compiled and run on an IBM 4361. The energy tables used were those originally developed by Tinoco and co-workers (Tinoco et al., 1971) and further refined by Salser (1977). These include values for the free energy gained by forming Watson-Crick base pairs. Closed bifurcations were allowed. Single-stranded regions had appropriate energy values assigned according to their length and kind (hairpin loops, interior loops, and bulges). The minimum acceptable length for hairpin loops was set to 3. G-U base pairs were allowed to occur and were assigned a free energy of 0 kcal/mol. The reference value of 0 kcal/mol corresponds to a single-stranded linear segment. The maximum distance (along the RNA) over which base pairing was allowed was set to the values indicated in the text. To display graphically these secondary structures, we used an interactive program developed by Dr. B. Shapiro (Shapiro et al., 1984) run on a DEC Vax 8700 with a Tektronix graphics display.

RESULTS

mRNA Structure Analyses

The specific enzymatic and chemical probes, RNases T1, T2 and S1, and dimethyl sulfate, were used to map single-stranded regions on apoI1 mRNA. Cobra venom (CV) RNase was used to map double-stranded regions. The molecular weights of these probes are given in Table II. Primer extension using apoII mRNA-specific DNA primers and reverse transcriptase was used to locate cleavage and chemical modification sites (Shelness and Williams, 1985). Five end-labeled probe-primers (I, II, III, E, and G, Table I) spaced at approximately 150-nucleotide intervals permitted cleavage and modification sites to be mapped with single nucleotide accuracy approximately 150-nucleotide intervals permitted cleavage and modification sites to be mapped with single nucleotide accuracy throughout the 659-nucleotide apoII mRNA.

In order to identify primary cleavage or modification sites, titrations were conducted with increasing concentrations of enzyme and dimethyl sulfate. Primary sites were identified as those detected while the majority of apoII mRNA molecules were intact. Representative titrations with T2 RNase and dimethyl sulfate as assayed with oligo(I1) are shown in Fig. 1. Cleavage and modification sites were mapped to single nucleotide resolution by comparison with a homologous sequencing ladder as shown in Fig. 2. Primer extended cDNAs resulting from RNase T2 cleavage at nucleotides 361, 362, 363, 364, 368, 385, 399, 407, and 428 are indicated (nucleotide 1 corresponds to the transcription start site), most occurring at adenosine residues which are preferred by RNase T2. cDNAs resulting from dimethyl sulfate methylation at nucleotides 358, 359, 361, 362, 366, 367, 368, 370, 371, 381, 383, 385, 406, 409, 410, 411, and 427 are also indicated. Dimethyl sulfate methylation at adenosine and cytidine residues yields cDNAs terminating one nucleotide 3' to the modified base (Inoue and Cech, 1985).

In addition to T2 RNase and dimethyl sulfate, RNases T1, S1, and CV were used as probes. With the exception of RNase S1, all digestion and modification reactions were done at 0.2 M Na+ and 0.005 M Mg2+, pH 7.5. S1 RNase reactions contained the same Na+ and Mg2+ concentrations but were at pH 5 and also contained 0.001 M Zn2+.

A total of 65 cleavage sites were detected by CV RNase, and 134 sites were detected by the various single-strand-specific probes (Fig. 3). This represents 30% of the nucleotides in apoII mRNA. As previously noted in the 3' noncoding region (Shelness and Williams, 1985), base-paired nucleotides detected by CV RNase do not overlap with nucleotides detected by single strand-specific probes throughout the length of the mRNA. This result suggests that apoII mRNA possesses a relatively stable secondary structure under these experimental conditions. S1 RNase was the least informative probe producing relatively few cleavages. Cleavages produced by S1 RNase were largely confirmed with one or more of the other probes suggesting that single-stranded domains are not dramatically different at pH 5 as compared to pH 7.5. The double- and single-stranded regions are distributed moderately uniformly throughout apoII mRNA. There is, however, a rather long single-stranded region bound by nucleotides 355–413, which includes the translation stop codon at nucleotides 399–401. Several other regions showed no reactivity with any of the probes suggesting that such regions may be inaccessible in the three-dimensional structure of the mRNA. One such region shown in Fig. 3 contains the AUG translation start codon.

Modeling of ApoII mRNA Secondary Structure

Using the nuclease cleavage and chemical modification data as constraints, secondary structure models of apoII mRNA corresponding to the lowest free energy configurations were determined. Calculations were performed on an IBM 4361 computer using the program of Zuker and Stiegler (1981), as described under "Experimental Procedures."

At different times in its lifetime, a mRNA may be subject to very different restrictions as to the distance (along the mRNA) over which base pairing can occur. These restrictions may reflect the presence of intron sequences in the primary transcript, the association of nuclear and cytoplasmic mRNA binding proteins with the mature mRNA, or the association with ribosomes and other components of the translation apparatus. A polyribosomal mRNA may have base pairing restricted to transient short range interactions in mRNA segments between ribosomes in the coding region or, possibly, between preinitiation complexes in the 5' noncoding region. It was of interest, therefore, to ask how the folding pattern of apoII mRNA would depend on the maximum distance along the nucleotide chain allowed for base pairing. The results of folding the apoII mRNA sequence with maximum allowed base pairing distances of (a) 659 (the full length), (b) 330, (c) 165, and (d) 41 are shown in Fig. 4. The calculated free energies are -76, -76, -76, and -34 kcal/mol, respectively. It can be seen that the folding pattern as well as the final free energy obtained with the shortest distance restriction are significantly different from the others. This is the expected result, since only small stem-loop structures and other short range interactions can occur with a distance restriction of 41 bases as might occur in the interribosomal region of polyribosomal apoII mRNA. With the maximum base pairing distance set at 165, a number of more complex secondary structures appear that are indicated by structures A–D in Fig. 4c. Structures A–D represent regions of the mRNA for which considerable experimental data exist. Interestingly, these structures are largely preserved when the maximum base pairing distance is set at 330 bases (Fig. 4b) or 659 bases (Fig. 4a).
FIG. 1. Concentration curves for RNase T2 and dimethyl sulfate (DMS) reaction with apoII mRNA. Total liver RNA was incubated with RNase T2 (final concentration ranged in 1.5-fold steps from 1.0 x 10^{-2} to 1.14 x 10^{-2} units/ml) (panel A) or with dimethyl sulfate (final concentration ranged in 1.5-fold steps from 3.4 to 39.6 mM) (panel B) as described under "Experimental Procedures." Samples were subjected to primer extension using oligo(II) (Table I), and cDNA products were analyzed by electrophoresis on a 6% polyacrylamide gel followed by autoradiography. Bracketed lanes show effects of increasing concentrations of T2 RNase or dimethyl sulfate. Lane M contains labeled DNA markers (MspI digest of pH322). Primer extension of mock-treated RNA samples (−) are shown. Primer extension of untreated total liver RNA is shown next to the DNA marker lane. Arrows or brackets labeled with a nucleotide coordinate, or a range of coordinates, are considered to represent cDNAs resulting from primary cleavage or modification of apoII mRNA. The nucleotide coordinates were determined by comparison to a homologous dideoxy sequencing ladder (Fig. 2) generated with cDNA clone mp8/apoIM3 (Shelness and Williams, 1985) and oligo(II) as primer.

4a). Overall, models 4b and 4c are similar even though the degrees of freedom in the folding of model 4b are far greater than in model 4c. Much of this similarity, including structures A−D, is also evident in model 4a which was folded with no distance restrictions. Only in the case of model 4a is there significant interaction between the coding region and the 5' or 3' noncoding region of apoII mRNA. Even in this case, however, most of the predicted secondary structures consist of domains formed from relatively short range interactions ranging between 50 and 120 bases. In addition, a number of the long range interactions in model 4a are due to only 2 or 3 base pairs, whereas most of the stems of 4 or more base pairs are present in relatively local structures. This is illustrated in Fig. 5 in which the model of Fig. 4a is shown complete (panel a) or with stems of 3 or less contiguous base pairs eliminated (panel b). With one exception, all base pairing interactions occur within 120 nucleotides along the mRNA.

Experimental Tests of the Predicted Structures

Three experiments involving hybridization of synthetic DNA oligonucleotides were designed to test the validity of the structures obtained by the folding method. Experiments were conducted with conditions of ionic strength, buffer components, and pH similar to those used to determine enzymatic cleavage and chemical modification sites on apoII mRNA. The results, therefore, can be interpreted in the context of the models in Fig. 4.

Does Base Pairing Take Place between the 3' Noncoding Region and the Coding or 5' Noncoding Regions?—Model 4a (Fig. 4) is the only model which predicts significant base pairing between the coding region and the 3' noncoding region. In this structure a large loop domain (bases 355−413) containing the translation stop codon (bases 399−401) is closed by a 10-base pair stem. This predicted loop domain was very susceptible to single strand-specific probes (Fig. 3).

The experimental strategy, therefore, was to cleave this loop with RNase H directed by oligo(B) which is complementary to nucleotides 380−399. Subsequent chromatography of the cleaved RNA on oligo(dT)-cellulose should show the mRNA fragment 5' to the cleavage site in the bound (poly(A)+) fraction according to model 4a (Fig. 4) or in the unbound (poly(A)+) fraction according to models 4b, 4c, or 4d. Northern blot analysis showed that apoII mRNA was completely cleaved by oligo(B)-directed RNase H digestion (data not shown).

ApoII mRNA fragments in the poly(A)+ and poly(A)+ fractions were assayed by primer extension with oligo(M) and -(E) (Table I) which sit 5' and 3', respectively, to the RNase H cleavage site. Fig. 6 shows the results of primer extension with the 5' side oligo(M). Analysis of the RNase H-cleaved apoII mRNA (lanes 1 and 2) showed that most of the 5'-apoII mRNA fragment was retained in the poly(A)+ fraction when this sample was heated to 100 °C prior to oligo(dT)-cellulose chromatography, the 5'-apoII mRNA fragment was recovered entirely in the poly(A)+ fraction (lanes 3 and 4). This result confirms that RNase H cleavage was complete and shows that after denaturation, the 5'-apoII mRNA fragment did not reassociate with the 3'-apoII mRNA fragment or with other mRNAs that might cause it to be retained in the poly(A)+ fraction. Mock-treated RNA samples typically showed 8−10%
of the mRNA in the poly(A)\(^+\) fraction (lanes 5 and 6). When this background value was subtracted from the experimental sample (lane 1), 82% of the 5\(^\prime\)-apoII mRNA fragment was retained in the poly(A)\(^+\) fraction due to stable interactions with the 3\(^\prime\) noncoding region. This result is consistent with model 4a (Fig. 4) but not models 4b, 4c, or 4d.

The stability of the association between the 3\(^\prime\) noncoding and coding regions was tested by determining the melting temperature. In this case, oligonucleotide-directed RNase H cleavage was carried out at 32 °C, and the RNA was incubated at various temperatures before chromatography on oligo(dT)-cellulose. RNase H cleavage was carried out at 32 °C instead of 37 °C to distinguish between the possibilities that (a) 18% of the 5\(^\prime\)-apoII mRNA fragments were not retained on oligo(dT)-cellulose because 37 °C is in the melting range of the structure or (b) 18% of the apoII mRNA molecules lack the association between the coding and 3\(^\prime\) noncoding regions. As in Fig. 6, the 5\(^\prime\)-apoII mRNA fragment was assayed by primer extension with oligo(M). The background value of apoII mRNA that was not retained on oligo(dT)-cellulose due to short poly(A)\(^+\) tails (typically 8–10%) was determined for each sample by primer extension with oligo(E). As shown in Fig. 7, the retained 5\(^\prime\)-apoII mRNA fragment was melted in a sharp transition with a tm of 50 °C. The plateau below 42 °C is consistent with 18% of the apoII mRNA molecules lacking this base-pairing interaction.

Test of Single- and Double-stranded Segments in the 3\(^\prime\) Noncoding Region by RNase H Cleavage—Oligonucleotides complementary to predicted single- and double-stranded domains of the 3\(^\prime\) noncoding region were used to test for susceptibility to RNase H cleavage. RNase H digestions were carried out for 45 min at 37 °C. As shown in Table I, oligo(H) and oligo(D) were targeted to nucleotides 402-411 and 548-567, respectively, which were predicted to be single-stranded. Oligo(J) and oligo(K) were targeted to nucleotides 435-444 and 414-430, respectively, which were predicted to be predominantly double-stranded. After RNase H digestion, reaction products were run on a denaturing agarose gel, blotted to nitrocellulose, and probed with full-length apoII cDNA. Fig. 8 shows a representative experiment in which oligo(H) did direct RNase H cleavage while oligo(J) and oligo(K) did not. Oligo(D) also directed RNase H cleavage (data not shown). These data support the assignments for regions D and H as single-stranded. The failure of oligo(J) and oligo(K) to direct RNase H cleavage is consistent with the assignment of the regions as predominantly double-stranded. The regions targeted by oligo(J) (435-444) and oligo(K) (414-430) are susceptible to cleavage by CV RNase, suggesting that the lack of RNase H cleavage is not due to the inaccessibility of these regions. The absence of cleavage in region K during the 45-min digestion with RNase H at 37 °C also indicates that this stem is very stable as was supported by the melting profile (Fig. 7). This also appears to be the case for region J.

Test of Single- and Double-stranded Segments by Primer Extension—The ability of oligonucleotides to prime reverse
Secondary Structure of ApoII mRNA

FIG. 4. Secondary structure models of mRNA. Energy-minimized secondary structure models of apoII mRNA were obtained using Zuker and Stiegler’s RNA folding program (1981). Base paired and unpaired nucleotides as determined by enzyme and chemical modification experiments were specified. Structure models were generated with four different constraints of maximum distance (659, 330, 165, and 41 nucleotides) between base-paired nucleotides and are shown in panels a, b, c, and d, respectively. The secondary structure drawings were done with the interactive computer graphics program of Shapiro et al. (1984). Four stem-loop domains common to models a, b, and c are stippled and indicated as A, B, C, and D. The start codon (AUG), stop codon (UAG), and polyadenylation signal (AAUAAA) are shown. The overlapping structures near the 3′ end in model 4d could not be untangled by the graphics program.

transcription was also used as an indirect probe of single- or double-stranded regions (data not shown). Table I shows the locations of oligo(J), oligo(K), and oligo(N) which did not prime reverse transcription and oligo(D) and oligo(L) which did prime reverse transcription. These results are consistent with the assignments of these regions in the structure model and with the RNase H data.

DISCUSSION

Chemical and enzymatic probes have been used to define the single- or double-stranded character of nucleotides throughout the 659-nucleotide apoII mRNA. By using CV RNase to detect base-paired nucleotides and a combination of probes to detect unpaired nucleotides, 199 nucleotides (30% of the apoII nucleotides) have been assigned as either single- or double-stranded. No overlap between single- and double-stranded nucleotides was seen throughout the entire length of apoII mRNA suggesting that this mRNA has a predominant and stable structure under these experimental conditions. This result was seen previously when the 3′ noncoding region of apoII mRNA was examined with CV RNase and RNases T1 and S1 (Shelness and Williams, 1985). These data do not establish that apoII mRNA has a unique secondary structure, since data are available on only 30% of the nucleotides and since alternate conformations were seen experimentally in one region of the mRNA (Fig. 7). Nevertheless, the absence of overlap between single- and double-stranded sites indicates that apoII mRNA does not exist in a large number of stable forms or in many alternate forms in rapid equilibrium.

Single- and double-stranded regions were distributed moderately uniformly along the mRNA except near the translation stop codon which resides in a large loop domain (nucleotides 355-413) that is highly susceptible to single strand-specific probes and shows no cleavage by cobra venom nuclease. Secondary structure models show the translation stop codons in rabbit α-globin mRNA and mouse β-globin mRNA in loop domains while this codon is base-paired in rabbit β-globin mRNA (Vary and Vournakis, 1984; Lockard et al., 1986). In none of these cases, however, was nuclease cleavage data available in this region to confirm the assignment. The polyadenylation recognition sequence in apoII mRNA is also assigned to a single-stranded domain (Fig. 4) as is the case with rabbit α- and β-globin mRNAs and mouse β-globin
mRNA (Pavlakis et al., 1980; Lockard et al., 1986). The translation start codon in apoII mRNA was not accessible to any of the probes and has two of the three nucleotides assigned as double-stranded in the structure model (Fig. 4). Whether the start codon is actually base-paired or is inaccessible because of the three-dimensional structure remains to be tested. The start codon in rabbit and mouse α-globin mRNAs was also inaccessible to nucleases while the start codon in the respective β-globin mRNAs was highly accessible (Pavlakis et al., 1980; Lockard et al., 1986).

The secondary structure analyses of globin mRNAs required purified mRNA that was end-labeled after removing the 5'-cap structure (Pavlakis et al., 1980; Lockard et al., 1986) or the 3' poly(A) tail (Vary and Vournakis, 1984). As a result, nuclease data at the two ends of globin mRNA were obtained on two RNA samples which were subject to different procedures prior to end labeling and nuclease analysis. In the present study, we have obtained approximately as much nuclease cleavage and chemical modification data as is available for any of the globin mRNAs. In this case, however, the use of primer extension to map sites of nuclease or chemical modification (Shelness and Williams, 1985) avoids the need to purify the mRNA or to modify the ends of the mRNA. In addition, with appropriate oligonucleotides for primer extension, data are obtained throughout the mRNA and are not limited to the 5' and 3' regions as is the case with end-labeled mRNAs. The present approach will permit the analysis of mRNA secondary structure in many mRNAs which cannot be purified and in internal regions of larger mRNAs which cannot be studied by end-labeling techniques.

The secondary structure of the exterior portion of stem-loop D (nucleotides 440–511, Fig. 4a) is nearly identical to that described earlier (region A in Shelness and Williams, 1985) in which the RNA folding was constrained only by RNase T1 data. This prominent structure also was seen when this region of 121 nucleotides was folded by itself (not shown) or when folded within the complete apoII mRNA sequence with no distance restriction (Fig. 4a) or restrictions of 330 bases (Fig. 4b) or 165 bases (Fig. 4c). Other studies show that loop regions within the stem-loop D structure contain cleavage sites for the degradation of apoII mRNA in vivo. The localization of cleavage sites to loop regions but not merely to
Secondary Structure of ApoII mRNA

mRNA in polyribosomes via dimethyl sulfate modification and nuclease cleavage indicate that the 3' noncoding region of apoII mRNA has a very similar, but probably not identical, structure in polyribosomal mRNP as in naked mRNA.\(^3\) Experiments to evaluate the functional role of secondary structure in stem-loop D are in progress.

The secondary structure model with no distance restriction (Fig. 4a) contains the translation stop codon within a large loop that is closed by a 10-base pair stem formed between the coding and 3' noncoding regions. The occurrence of this structure was tested by determining whether the 5'-apoII mRNA fragment remained associated with the 3' noncoding region when the loop region was cleaved by RNase H. The result showed that 82% of the 5'-apoII mRNA fragment remained bound to the 3' fragment. The \(t_m\) for the dissociation of this complex was 50 °C and occurred over a 10 °C span. Although it is not possible to calculate the \(t_m\) of a short double-stranded RNA region flanked by long single-stranded segments, this result is compatible with the melting of a 10-base pair stem.\(^3\) These results are consistent with the presence of this or a similar stem structure formed between the 3' noncoding region and nucleotides 5' to the site of RNase H cleavage in the coding region of apoII mRNA as in model 4a. While this stem may exist in stable form in the isolated mRNA, it would likely occur only transiently in polyribosomal apoII mRNA due to ribosome transit through the 5' side of the stem. One possibility is that alternate structures in this region may occur during ribosome transit and play a role in translation termination or in mRNA degradation. Interestingly, when the 3' noncoding region is folded by itself, those nucleotides which form the 3' side of the stem in the complete model (414–423, Fig. 4a) are largely base-paired to nucleotides further downstream to form an extension of the stem in stem-loop D (Shelness and Williams, 1985).

Additional tests of apoII mRNA structure in the region near the translation stop codon were made by determining whether specific oligonucleotides could direct RNase H cleavage or serve as primers for reverse transcription. These results (Table I) are consistent with the nucleotide assignments in the structure model (Fig. 4a) and in agreement with the experimental data obtained with nuclease and chemical probes. RNase H cleavage and primer extension are crude probes of RNA structure since they detect only those regions accessible to these large enzymes (Table II). In addition, since only 4–5 base pairs are required for oligonucleotide-directed RNase H cleavage (Donis-Keller, 1979; Vary and Vournakis, 1984), a positive result does not indicate which nucleotides within the target site were accessible. Primer extension is somewhat more specific since the 3' nucleotides of the oligomer must be base-paired for extension. Under the conditions employed in this study, oligo(dT)\(_4\) did not prime reverse

\(^3\) K. Breslauer, personal communication.
transcription of poly(A)-containing RNA while oligo(dT), did (data not shown). Thus, primer extension probes those nucleotides in the RNA corresponding to the five nucleotides at the 3' end of the oligonucleotide. Irrespective of the limitations of primer extension and oligonucleotide-directed RNase H cleavage, these techniques provide an alternate means to test for single-stranded structure in mRNA. Of particular interest for apoII mRNA, primer extension and RNase H cleavage are directed by oligo(E) (Table I) which sits at the junction with the poly(A) tail and terminates in the polyadenylation signal. This result provides direct evidence that the polyadenylation signal resides in a single-stranded region.

When apoII mRNA is folded with no restrictions on the distance over which base pairing can occur, the resulting structure (Fig. 4a) is very complex and appears to show extensive long range interactions which bring the ends of the molecule together. Extensive long range interactions also have been proposed in globin mRNAs with interactions between the 5' and 3' noncoding regions as well as between the coding and noncoding regions (Pavlakis et al., 1980; Vary and Vournakis, 1984; Lockard et al., 1986). In the case of apoII mRNA, prominent secondary structure features A-D (Fig. 4) are seen when the RNA is folded with no distance restriction or with restrictions of 330 or 165 nucleotides. This result may indicate that major secondary structures are formed through relatively local interactions over distances of 50–120 nucleotides along the mRNA.

The relative significance of local versus long range interactions is shown in a different way in Fig. 5 in which all interactions of 3 or less contiguous base pairs have been retained (panel a) or eliminated (panel b). As is evident, most of the long range interactions are due to only 2 or 3 base pairs while the prominent local interactions contain helices of 4–10 base pairs. The complex structure with many long range interactions (Fig. 5a) is largely unfolded and has primarily local structures when only helices of 4 or more base pairs are displayed (Fig. 5b). Also seen in Fig. 5b is that 12 of 13 helices are formed by base pairing within an exon while only one helix is formed by base pairing between two exons. This result raises the interesting question of whether prominent secondary structures in mRNAs may be largely due to intraexonic base pairing. It will be possible to approach this question as well as the overall question of local versus long range interactions through analyses of RNAs made in vitro that correspond to exon fragments or to intron-containing primary transcripts.

Acknowledgments—We would like to thank Clinton MacDonald and Mary Reyland for critical reading of the manuscript.

REFERENCES


