Nucleotide sequence determining the first cleavage site in the processing of mouse precursor rRNA

(1-cell-free processing/deletion mutants/RNA polymerase I transcription)

NESSLY CRAIG*, SUSAN KASS‡, AND BARBARA SOLLNER-WEBB‡

*Department of Biological Sciences, University of Maryland Baltimore County, Catonsville, MD 21228; and ‡Department of Biological Chemistry, The Johns Hopkins School of Medicine, Baltimore, MD 21205

Communicated by Keith R. Porter, September 12, 1986

ABSTRACT The first step in the processing of 47S precursor rRNA in mouse cells is reproduced in vitro in an S-100 transcription reaction and consists of an endonucleolytic cleavage at residue +650 of the primary transcript followed by rapid degradation of the fragment upstream from residue +650. An analogous processing occurs in human rRNA. The mouse and human rRNA sequences are ~80% conserved for 200 nucleotides on the 3' side of these processing sites, suggesting that this conserved region may be important in specifying the processing. To test this hypothesis, we constructed a systematic series of deletion mutants approximating the mouse rRNA processing region from both the 5' and 3' directions and analyzed the processing of their transcripts in vitro. The 5' boundary of the region required for processing is quite sharp and corresponds to the rRNA cleavage site at the 5' end of the conserved sequence region. The 3' boundary is more complex: (i) The 3' deletions extending to between 250 and 130 nucleotides beyond the processing site cause about a 50% decrease in the amount of the processed RNA. (ii) A 3' deletion that extends to 109 nucleotides beyond the processing site greatly reduces the processing efficiency. (iii) Deletions to or beyond 91 nucleotides on the 3' side of the processing site virtually eliminate processing. Under altered ionic conditions, transcripts of 3' deletions extending to only 41 nucleotides beyond the processing site can still direct a low level of accurate processing. These results demonstrate that the mouse-human conserved sequence just on the 3' side of the primary rRNA processing site consists of several domains that direct and/or augment both the initial endonucleolytic cleavage and the closely coupled selective degradation of the upstream fragment that together constitute the primary rRNA processing event.

Ribosomal DNA in mammalian cells is initially transcribed as a long (>14 kilobase) precursor that is subsequently processed to yield the mature 18S, 5.8S, and 28S rRNAs (1-4). Studies analyzing nuclear rRNA processing intermediates and their labeling kinetics have provided information about the preferred cleavage pathways and probable cleavage sites. Early work from the laboratories of Dawid, Penman, and Perry (reviewed in ref. 2), confirmed and extended by studies of Bowman et al. (5), concluded that the major cleavage sites in mouse precursor rRNA are at or near the ends of the mature rRNA molecules. Subsequently, Miller and Sollner-Webb (6) found that mouse rRNA undergoes an additional processing event at residue +650 of the primary transcript, near the beginning of the ~4-kilobase external transcribed spacer. This processing evidently involves an endonucleolytic cleavage followed by rapid destruction of the upstream fragment (3, 7). In vivo, this is the earliest known rRNA cleavage event, and ~90% of the precursor 45-47S rRNA consists of this processed species (6, 8). In vitro, the processing is efficiently reproduced when cloned mouse ribosomal templates are transcribed in an S-100 extract of cultured mouse cells (6). We have recently found (unpublished data) that: (i) there is an analogous processing in the external transcribed spacer of human rRNA, (ii) the mouse and human rRNA sequences are conserved for 200 base pairs on the 3' side of these processing sites (although they are markedly divergent elsewhere in the external transcribed spacer), and (iii) human precursor rRNA is correctly processed in a mouse cell extract. These results suggest that sequences within the conserved region may direct this rRNA processing event. At present, however, virtually nothing is definitively known about the mechanism of any higher eukaryotic rRNA processing event, neither the rRNA determines of the cleavage sites nor the identity or mode of action of the relevant nucleases.

To directly identify which nucleotide sequences encode the primary mouse rRNA processing event, we have created systematic 5'- and 3'-deletion mutants of mouse rDNA and used these as templates in a coupled in vitro transcription-processing system. We find that the processing signal is largely contained within the 200-nucleotide mouse/human conserved region that occurs on the 3' side of the processing site. Sequences within the first 40 nucleotides are essential for processing and are sufficient to direct a low level of cleavage, while sequences within the next ~90 nucleotides greatly augment the efficiency of the processing reaction. The data further suggest that the initial cleavage and the very rapid degradation of the upstream fragment are closely coupled events.

MATERIALS AND METHODS

Preparation of rDNA Deletions. The 5' - and 3'-rDNA deletions were constructed using BAL-31 exonuclease as described in the legend to Fig. 1 and sequenced by the dideoxynucleotide method of Sanger et al. (9). The resultant deletion plasmids are designated according to the last nucleotide of mouse rDNA sequence they contain. To allow transcription of the 5' deletions, a fragment that contains the mouse rDNA promoter was inserted upstream from each deletion (Fig. 1A). The internal deletion/mutations were constructed by joining selected pairs of 3' and 5' deletions (Fig. 1C). Templates for run-off transcriptions were truncated with Pst I (residue +1300) for the 5' deletions, with Bgl I (653 nucleotides beyond the deletion endpoint), or with EcoRI (3 nucleotides beyond the deletion endpoint) for the 3' deletions, and with HindIII (residue +1310) for the internal deletion/mutations.

Transcription—Processing Reactions and Their Analysis. The S-100 extracts were prepared from mouse L1210 cells growing exponentially in suspension culture as described (6). The transcription reactions were performed essentially as

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†To whom reprint requests should be addressed.
RESULTS

Construction and Assessment of Processing Site Deletion Mutants. To directly determine which sequences of the mouse precursor rRNA specify the primary processing event, we constructed a systematic series of deletion mutants extending into the region of the processing site from both the 5' and 3' directions (Fig. 1A and B). The 5' deletions were formed in two steps. First, deletions were created by BAL-31 exonuclease digestion starting at residue +300 and extending toward the processing site (residue +650) from the upstream direction (Fig. 1A Upper). Then, a promoter region containing mouse rDNA sequence −170 to +55 was cloned upstream of each of these promoterless 5'-deletion mutants (Fig. 1A Lower). The 3'-deletion mutants were formed from a plasmid that contained the rDNA promoter and processing region by BAL-31 exonuclease digestion starting at residue +1300 and extending upstream toward the processing site (Fig. 1B).

The precise endpoints of each of these deletions was determined by DNA sequence analysis (Fig. 2A). The resultant plasmids were linearized several hundred base pairs downstream of the processing site and used as templates in S-100 in vitro reaction mixtures. Following a radioactive pulse-labeling period ("b"), half of each reaction mixture was stopped, and the other half was subjected to a cold chase period ("c"). The transcripts were isolated, resolved on a 3% acrylamide/9 M urea gel, and visualized by autoradiography; the band intensities were determined by densitometry.

Processing of the natural rRNA gene transcript involves a cleavage at residues +650 and very rapid degradation of the upstream fragment. The resultant processed fragment extends from residue +650 to the 3' end of the primary transcript (6). Thus, transcripts of the 5' deletions will vary in size according to the extent of the deletion, while their resultant processed fragments would all be expected to be the same size (Fig. 2B). In contrast, both the primary transcripts and the processed fragments derived from templates progressively deleted from the 3' direction should decrease in size according to the extent of the deletion (Fig. 2C).

The 5' Boundary of the Processing Signal. Fig. 3 shows a typical experiment assessing the ability of transcripts of the 5'-deletion mutants to be processed. As expected, these templates are transcribed with equal efficiency and yield transcripts of the predicted sizes. Notably, however, the efficiency of processing greatly depends on the particular 5' deletion. Transcripts of templates deleted up to residue +617 are processed to the same level as the parental rDNA, transcripts of mutants 5'A+631 and 5'Δ+646 are processed at

unit/ml, but the results were the same with or without its addition. Each extract was optimized for template concentration (usually 5-7.5 μg/ml) and for KCl concentration (usually ~100 mM). After a labeling period of 20-45 min at 30°C, the reaction mixture was split into two equal parts, and one-half was stopped ("pulse"). The other half was supplemented with 2-3 mM unlabeled CTP and was incubated for 45-60 min ("chase"). The RNA products were isolated, analyzed on 3% acrylamide/9 M urea sequencing gels, and detected by autoradiography as described (6). Band intensities were quantitated by densitometry (11). Processing efficiency is defined as the ratio of the number of radiolabeled, processed product molecules present after the chase period to the number of primary transcripts plus processed products present after the pulse labeling. The processing efficiency of the parental rDNA is remarkably high; generally ~50% of the initial transcript molecules are recovered as the processed species after the chase period. This is, in fact, a minimal estimate of processing efficiency, since RNAs that lack the processing site are also only recovered with ~50% efficiency following an equivalent chase period.

Fig. 1. Construction of deleted rDNA templates. (A) The 5'-deletion series. Plasmid p5'Pvu-Sal containing mouse rDNA sequences from +300 to +3000 (6) was linearized with Pvu II at position +300 and treated with BAL-31 for different time intervals. After repair of the ends with the Klenow fragment of DNA polymerase I and digestion at position +1300 with Xho I, the 5'-deleted fragments were gel purified and inserted between the Sma I and Sal I sites of M13mp9 (or mp19) to form plasmid p5'Δ(M13). After sequencing (9), the rDNA of each desired 5' deletion was isolated as an EcoRI-HindIII fragment and cloned downstream of a Sal I-EcoRI fragment of plasmid p170/+55 (10) that contains the mouse rDNA promoter sequence (residue −170 to +55) and 233 base pairs of vector sequence; the cloning vector was the large Sal I-HindIII fragment of pBR322. (B) The 3'-deletion series. Plasmid p5'Sal that contains mouse rDNA sequence −170 to +3000 was digested with Xho I at residue +1300, then with BAL-31 for various times, and then end-repaired. After cleavage at residue −170 with Sal I, the desired 3'-deleted fragments were gel-purified and cloned into Sal I-Sma I-cut M13mp9. (C) Internal deletions. The upstream rDNA portion from the desired 3' deletion (Sal I-EcoRI fragment) and the downstream rDNA portion from the desired 5' deletion (EcoRI-Pst I fragment) were cloned together using the large Sal I-Pst I fragment from pUC18 as a vector. The mouse rDNA sequence between the 5' and 3'-deletion endpoints is thereby replaced with a 10-base-pair head to head duplications of the mp9 Sal I-EcoRI polynucleotide region. E, EcoRI; B, Bgl II; H, HindIII; P, Pst I; S, Sal I; Sm, Sma I; X, Xho I; +, processing site at residue +650.

described by Miller and Sollner-Webb (6) in ~100 mM KCl/5 mM MgCl₂/1 mM dithiothreitol/0.1 mM EDTA/10% (vol/vol) glycerol/0.5 mM each of ATP, GTP, and UTP/0.05 mM CTP. In some experiments RNAsin was added at 0.4
The 3' Boundary of the Processing Signal. To determine the 3' boundary of the mouse rRNA region that directs processing at residue +650, the 3' deletions were transcribed in the mouse S-100 cell extract, and the reactions were subjected to a cold chase period. The results are somewhat more complex than those obtained with the 5' deletions since there is not one discrete 3' boundary for the relevant sequence but rather there appear to be several domains on the 3' side of the cleavage site that serve to successively augment the processing efficiency. Furthermore, the effect of these domains varies according to the ionic conditions. The primary data are shown in Fig. 4 and quantitated in Fig. 6. Under normal reaction conditions (100 mM KCl; Fig. 4A) processing was observed for all 3' deletions extending in to residue +691, while 3'Δ+666 directs no detectable processing. Thus, sequences extending only 40 nucleotides on the 3' side of residue +650 are able to direct processing at this site. Although accurate, the processing directed by 3'Δ+691 is only 10% as efficient as that of large rRNA transcripts, e.g., 3'Δ−2350 (data not shown), 3'Δ−1300, or 3'Δ−1100 (Fig. 4 and see Fig. 6). The less extensive deletion mutants 3'Δ+723, 3'Δ+728, and 3'Δ+741 also direct processing at this ~10% level. However, transcripts of 3'Δ+759, which contain 18 additional rRNA nucleotides, process several fold more efficiently, and a similar level of processing is obtained with the less extensive deletions through 3'Δ+823. Slightly more processed product is obtained from still less extensive deletions between 3'Δ+850 and 3'Δ+894 that yield over half as much processed product as does undeleted rRNA. This variation in processing efficiency is highly reproducible in duplicate experiments that used three different S-100 extract preparations.

When reacted under more stringent conditions (150 mM KCl; Fig. 4B), the overall level of transcription and processing is reduced by ~30%, and the rRNA regions required to direct processing are somewhat different. 3'Δ+741 and more extensive deletions yield no detectable processed product; transcripts of 3'Δ+759 are processed only ~10% as efficiently as control transcripts, and transcripts of 3'Δ+781 through 3'Δ+894 are processed 30–40% as efficiently. Thus, under these higher salt conditions, more extensive rRNA regions are needed to direct appreciable levels of processing.

These results strongly suggest that sequences within the first 40 nucleotides on the 3' side of the processing site are able to specify processing but that adjacent domains can markedly augment the processing efficiency. However, since processing efficiency is determined from the abundance of the processed rRNA species, one must consider whether the stability of the processed products from the different templates could vary and thus bias the results. For the 5' deletions, this is not a problem since the primary sequence of the processed fragment is identical for each deletion template, but, for the 3' deletions, the sequence of the processed fragment from each mutant differs, and this could conceivably result in differential stability. However, two lines of evidence indicate that this is not a problem, and that the observed level of precursor RNA does indeed reflect processing efficiency. First, in kinetic experiments, a set of reactions analogous to those of Fig. 4A was stopped after chase periods of between 15 and 90 min, and the relative processing efficiency of the various deleted RNAs was found to be very comparable, making unlikely the hypothesis that the processed RNAs resulting from the more extensive 3' deletions are less stable. Second, selected 3' deletions were tested in an S-100 extract in which the upstream-generated fragment was more stable than is usually observed. Since this upstream fragment (residue +1 to +649) has exactly the same primary sequence for all of the 3' deletions, it should be equally stable in all of the reactions. As seen in Fig. 4C, the upstream (650 nucleotides) and downstream (varying lengths) processed fragments are present in comparable relative molar ratios for all of the 3'-deleted rRNAs, strongly indicating that

Fig. 3. The effect of 5' deletions of mouse rDNA on processing. The designated 5' deletions were linearized with Pst I and used as template. After a 25-min pulse-labeling period (lanes p), CTP was added to a final concentration of 3 mM for a 45-min chase period (lanes c). The RNA was isolated and electrophoretically resolved.
**FIG. 4.** The effect of 3' deletions of mouse rDNA on processing. (A and B) The designated 3' deletion mutants were truncated with Bgl II and incubated under the standard conditions for 45 min (lanes p) and then chased for 45-60 min (lanes c). The reactions were performed with the same S-100 extract at two different ionic strengths, (A) 100 mM KCl and (B) 150 mM KCl, and RNA was assayed as in Fig. 3. (C) Representative 3' deletions were transcribed in an extract in which the upstream fragment (+1 to +649) is especially stable; reactions were pulse-labeled for 20 min and chased for 60 min at 100 mM KCl. Pulse and chase lanes are shown in pairs, with the position of the processed fragment indicated with arrowheads. The upstream generated fragment is 650 nucleotides long.

differential processing efficiency rather than differential RNA stability is responsible for the observed results. These results also show that the processing involves an endonucleolytic cleavage, rather than just a 5'-exonuclease action.

**Effect of Internal Deletions on Processing.** To further assess the effect of various regions in directing processing, we constructed internal rRNA mutations by joining together selected pairs of 5' and 3' deletions (Fig. 1C). The results of representative in vitro transcription–processing reactions with these templates carried out at 100 mM KCl and 150 mM KCl are shown in Fig. 5 A and B, respectively, and are tabulated in Table 1. As expected, mutations distant from the processing region have no discernable effect, while a mutation that removes the 40-base-pair core processing signal (p643/706) virtually abolishes processing. The effect of mutation p722/742 is dependent upon the ionic strength of the reaction. At 100 mM KCl, the transcript of p722/742 is processed at control levels, while, at 150 mM KCl, its processing is relatively inefficient. This result is consistent with the observation that the deleterious effect of 3'-deletion mutants is more pronounced at 150 mM than 100 mM KCl (Figs. 4 and 6). The efficient processing of p722/742 at 100 mM KCl is especially interesting in light of the results obtained with the 3'-deletion mutants under these conditions. As was observed in Fig. 4A, 3' Δ+741 extends to the downstream border of this region and reduces processing efficiency by ≈90%, while further 3' deletions that also remove the region between +691 and +741 show no additional processing deficiencies. This suggests that one domain important for processing is on the 5' side of residue +691 and that the 3' boundary of another important domain is on the 3' side of 741. In the mutant p722/742, 19 nucleotides of this intermediary region are replaced by 10 nucleotides of polylinker. Efficient processing of the RNA from this template demonstrates that this deleted sequence does not contain essential residues and that the relative position of the upstream and downstream processing domains can be varied by at least 9 nucleotides without impairing their effect.

**DISCUSSION**

Work from our laboratory (6) has identified the transcriptional initiation site of mouse rRNA and demonstrated that the primary rRNA transcript is rapidly cleaved at nucleotide 650, both in vivo and in vitro. Consistent with an earlier suggestion (12), we have demonstrated that there is an analogous processing site in human rRNA at position +414 (unpublished data). Moreover, the 200-nucleotide regions just on the 3' side of the mouse and human RNA processing sites are ≈80% homologous while surrounding regions show little sequence homology, and this human rRNA region directs accurate processing in the mouse S-100 extract (unpublished data), suggesting that the sequence of this

**Table 1.** The effect of internal deletions on rRNA processing

<table>
<thead>
<tr>
<th>Internal mutation</th>
<th>Processing efficiency, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 mM KCl</td>
</tr>
<tr>
<td>p5' Sal (control)</td>
<td>61</td>
</tr>
<tr>
<td>+417/+427</td>
<td>82</td>
</tr>
<tr>
<td>+594/+617</td>
<td>85</td>
</tr>
<tr>
<td>+643/+706</td>
<td>4</td>
</tr>
<tr>
<td>+722/+742</td>
<td>81</td>
</tr>
<tr>
<td>+802/+831</td>
<td>62</td>
</tr>
<tr>
<td>+909/+920</td>
<td>79</td>
</tr>
</tbody>
</table>

The average efficiency of processing in experiments such as are shown in Fig. 5 was quantitated.
The average KCl mM ing. in homologous 200-nucleotide sequence coincides with the processing and +850 boundary processing and +850 sequence inhomologous mouse/human. The data further specify several domains to a and downstream of this sequence. Involves refs. fragment upstream unstable (Fig. 3 region (residue +650 I). The endpoint 3' site but has only 5' deletion templates that are present and between +823 and +850 efficient deletions (Fig. 15). This is also true for 5'Δ+653, whose upstream fragment contains no rRNA sequences from the region of the processing site but has only vector sequences in its 3' region (Fig. 2A). In contrast, primary transcripts containing many different kinds of sequences are quite stable in the S-100 extract (15). In particular, when 5'Δ+653 is linearized with Kpn I at residue +647 (just upstream of the deletion endpoint) and used as a template, the resultant run-off transcript is also stable (data not shown). This transcript would be expected to be identical to the upstream fragment generated upon processing of 5'Δ+653 RNA in Fig. 3, lane 12, except that it lacks three 3' nucleotides of polylinker sequence. The fact that this run-off RNA is stable while the comparable fragment resulting from processing is very unstable suggests that the instability is not inherent in the RNA sequence. Instead, the rapid degradation of the upstream processed fragment is evidently closely coupled to the initial endonucleolytic cleavage. Possibly, the endonucleolytic and exonucleolytic activities are in a common processing complex and act in a concerted manner.

The maturation event that is examined in this paper is the first processing that mouse 47S rRNA precursor undergoes en route to becoming mature 18S, 5.8S, and 28S RNA. Like the other generally recognized rRNA processing sites (2, 4, 5), this mouse processing site at residue +650 occurs at the boundary between conserved and nonconserved sequence domains. As yet, none of the processing events that create the mature 18S, 5.8S, and 28S rRNA termini have been demonstrated to occur accurately in vitro, and thus, there is no information available on the sequence requirements at these other rRNA processing sites. However, the similarity in the asymmetric conserved/nonconserved sequence pattern between these sites and the mammalian primary rRNA processing site examined in this paper suggests that these other processing events may be similarly directed by residues extending over a 100- to 200-nucleotide region in the abutting conserved sequence domain.

We thank Ms. Sue Millionie for help in preparing the manuscript. This work was supported by Basic Research Grant 1-192 from the March of Dimes Birth Defects Foundation.