MOLECULAR AND CELLULAR BIOLOGY, Jan. 1991, p. 458–467
0270-7306/91/010458-10$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 11, No. 1

Sequence Organization and RNA Structural Motifs Directing the Mouse Primary rRNA-Processing Event

NESSLY CRAIG,[1]* SUSAN KASS,[2]† AND BARBARA SOLLNER-WEBB[2]

Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland 21228,[1] and Department of Biological Chemistry and the Human Genetics Program, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205[2]

Received 13 August 1990/Accepted 15 October 1990

The first processing step in the maturation of mouse precursor rRNA involves cleavage at nucleotide ca. +650, at the 5′ border of a 200-nucleotide region that is conserved across mammals and contains the sequences that direct the processing. To identify the relevant sequence elements, we used rRNAs with small internal mutations and short pre-rRNA substrates. Much of the region can be mutated without appreciable effect, but nucleotides +655 to +666 appear to be absolutely required and short segments surrounding +750 and +810 markedly stimulate processing. The minimal processing signal corresponds to rRNA nucleotides +645 to +672. Formation of a ribonucleoprotein complex of retarded electrophoretic mobility is evidently necessary but not sufficient for processing. Computer-assisted analysis suggested a phylogenetic- and mutant-supported secondary structure in which the minimal processing signal forms a stem with the +655 region in the loop, and there is a separate branched duplex containing the downstream stimulatory sequences. Use of antisense RNA, in *trans* and in *cis*, to sequester the +655 region in a duplex supported the hypothesis that this critical region was needed in a single-stranded conformation for processing and for specific complex formation.

rRNA in mammalian cells is initially transcribed as a large (ca. 14,000-nucleotide or ca. 47S) precursor, which is subsequently processed into the mature 18S, 5.8S, and 28S rRNAs. Nearly two decades ago it was established that there is a preferential, but not obligatory, pathway in the processing cleavages at the ends of the mature rRNA regions (reviewed in references 1, 4, 12, and 16). However, the first cleavage in mammalian pre-rRNA maturation has since been shown to occur near the 5′ end of the external transcribed spacer, 4 kb upstream from the 18S region (5, 13). The rRNA upstream of this processing site is degraded virtually concomitantly with cleavage, while the downstream rRNA gives rise to the major pre-rRNA form, 45S rRNA. This processing occurs at either of two adjacent sites, each with a 2- to 3-nucleotide heterogeneity, residues ca. +650 and ca. +656 in mice and residues ca. +414 and ca. +419 in humans (8).

The mechanism by which any of the rRNA-processing sites are recognized and cleaved in metazoan cells is unknown. In procaryotes, complementary sequences flanking the 16S and 23S rRNA regions have been shown to form stems which are cleavaged by RNase III to liberate the loop containing the mature rRNA regions (11, 22). In contrast, in the three metazoan examples studied to date, the sequences adjacent to the processing sites are sufficient to direct the cleavage. The processing at the 5′ and the 3′ ends of the 18S region does not require the 3′ portion of the 18S region in vivo (17, 21) and occurs in vitro on substrates containing only 0.3 and 0.5 kb surrounding the processing sites, respectively (6, 19). Furthermore, for the primary processing in the external transcribed spacer (13), the cleavage signal resides immediately 3′ of the processing site, within a 200-nucleotide sequence that is conserved between rodents and humans (3). These studies (3) defined the boundaries of the important region by using 5′ and 3′ deletion templates in a coupled transcription and processing system, but they did not address whether all of the large intervening region or only small portions of it were critical. Recent studies have shown that the primary rRNA-processing region associates with cellular components to form a large ribonucleoprotein complex (9) to which the U3 small nuclear ribonucleoprotein particle, which is essential for the processing (10), can also bind.

This paper reports analysis of the sequence signal used by the rRNA primary processing machinery. Using a series of "internal mutations" within an otherwise unaltered mouse rRNA context, we found that most small portions of the conserved 200 nucleotides can be altered without significantly inhibiting processing, but that the segment between +655 and +666 is critical and two other segments also contribute to the processing efficiency. In complementary studies with short T7 RNA polymerase-generated rRNAs, the sequence between +646 and +673 was shown to be sufficient for minimal levels of processing, and additional 3′ sequences markedly increased efficiency. The ability of these substrates to form the specific ribonucleoprotein processing complex was also found to correlate with their processing efficiency. Finally, computer-predicted secondary structure analysis suggested that a single-stranded loop containing the cleavage site is important for processing; when these residues were placed in duplex by using antisense rRNA oligonucleotides, processing was indeed abolished.

MATERIALS AND METHODS

Plasmid constructs. (i) Internal mutations. These plasmids join selected pairs of fragments containing 3′ and 5′ deletions of mouse rDNA (3). These fragments contain the rDNA segments from position −168 to the site of the 3′ deletion and from the site of the 5′ deletion to rDNA position +1290, respectively. They were excised from their M13mp9 (or M13mp19) vectors as *Sal*I-*Eco*RI and *Eco*RI-*Pst*I fragments,

* Corresponding author.
† Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115.

respectively, and were joined with a *SalI-PstI*-cleaved pUC18 vector (see Fig. 1C in reference 3). In essence, in these internal-mutation (IM) constructs, a short region between the noted 5′ and 3′ deletion endpoints was replaced with 10 bp of polylinker, a head-to-head duplication of the *EcoRI-SmaI* region of mp9 (or 21 bp of polylinker when one of the deletions came from a mp19 vector); otherwise, the rDNA region between positions −168 and +1290 is regenerated. Since these plasmids contain the entire rDNA promoter, they serve as efficient templates in the S-100 in vitro transcription system. The sequence of the mutated region in all plasmids was confirmed by dideoxy sequencing.

(ii) Mini-rRNA genes. The templates used for transcription of small rRNA regions by T7 RNA polymerase contained rDNA from the indicated 5′ position to residue +1290 and were subcloned as *EcoRI-PstI* fragments from previous 5′ deletions (3) into pGEM1. The T7-transcribed RNA has 28 nucleotides of polylinker sequence before the mouse rRNA sequence, which extends from the indicated 5′ position (e.g., pT7+645) to the site at which the template was truncated by restriction enzyme cleavage. The RNA designation T7+645/875 indicates such a transcript that includes mouse rRNA residues +645 to +875.

(iii) Antisense rRNA plasmids. The template for the *trans* antisense RNA was constructed by inserting an oligodeoxynucleotide complementary to the mouse rRNA sequence +649 to +664 (plus terminal *Eco*RI and *Sac*I sticky ends) into *Eco*RI- and *Sac*I-cleaved pGEM3z, using Taq DNA polymerase to synthesize the other strand (notch cloning [2]). The 4-bp *Eco*RI overhang was subsequently removed with *Eco*RI digestion, S1 treatment, and then religation. This effectively replaces 6 nucleotides of polylinker with 16 nucleotides of complementary rDNA sequence. The plasmid was linearized with *Hin*dIII in the distal polylinker region and transcribed by using T7 RNA polymerase. Control nonspecific RNA was from *Hin*dIII-cleaved pGEM3z. (Antisense RNA rather than antisense DNA was used because the cell extracts have high levels of RNase H activity, which would cleave the pre-rRNA.)

The *cis* antisense plasmids contained a 38-bp double-stranded oligodeoxynucleotide (5′-AATTCATCGATCTAA GAGTGAGCAGGGTACCGAGCTCG-3′) inserted into the unique *Eco*RI site adjacent to position +645 of pT7+645 (recloned in pGEM3z). With the oligodeoxynucleotide inserted in the sense orientation, the rRNA region +645 to +664 plus downstream polylinker is duplicated; when it is inserted in the antisense orientation, the T7 RNA polymerase-catalyzed transcript forms a 34-bp perfectly duplexed stem (containing residues +645 to +664 of mouse rRNA and the adjacent polylinker sequence) terminated by a 4-nucleotide loop (see Fig. 6A).

In vitro transcription and processing reactions. (i) Coupled systems. For the coupled transcription and processing system, we used mouse cell S-100 extract and templates truncated at the *Hin*dIII site just 3′ to nucleotide +1292, as described previously (3, 13). After a 45-min transcription period in the presence of radiolabeled CTP, one-half of the reaction was stopped (pulse); the other half was supplemented with 3 mM unlabeled CTP (a 60-fold excess, reducing the effective $MgCl_2$ concentration to 0.5 mM) and incubated for a further 60 min at 30°C to allow processing of the labeled precursor rRNA (chase). The RNA products were isolated, analyzed on 3% polyacrylamide–9 M urea gels, and detected by autoradiography at −70°C with preflashed X-ray film. Band intensities were quantitated by using a Bio-Rad densitometer and autoradiograms exposed within the linear range.

(ii) Uncoupled processing. Labeled pre-rRNA was synthesized from linearized pGEM-based plasmids (8) by using [α-^{32}P]CTP (10 to 20 Ci/mmol) and T7 RNA polymerase, as specified by the supplier (Promega or Bethesda Research Laboratories). Competitor RNAs were synthesized with [^{3}H]CTP (0.2 Ci/mmol) to facilitate their quantitation (9). The small *trans* antisense RNA was synthesized as described previously (14), using [α-^{32}P]CTP at 0.05 Ci/mmol. Labeled RNA was phenol extracted, ethanol precipitated, generally purified by polyacrylamide gel electrophoresis, and counted.

For processing, 0.002 to 0.010 pmol of the T7 polymerase-synthesized [^{32}P]rRNA was added to a 25-μl reaction mixture containing 5 to 10 μl of S-100 extract, 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.9), 100 to 120 mM KCl, 2 mM $MgCl_2$, 1.5 mM ATP, 1 to 2 mM dithiothreitol, 0.1 mM EDTA, and 10% glycerol. This gives a free $MgCl_2$ concentration of ca. 0.5 mM, the optimal concentration for processing (8). "Unlabeled" competitor RNA (12-fold molar excess, determined from its internal ^{3}H labeling) was preincubated with the extract for 10 min prior to assembly of the complete processing reaction (9). The *trans* antisense RNAs (500- to 1,000-fold molar excess) were preincubated with the labeled substrate RNA for 10 min at 65°C and then 10 min at 37°C prior to addition of the other components of the processing reaction. After a 45- to 60-min incubation at 30°C, the RNA was analyzed as above.

Assembly and electrophoretic resolution for the processing complex. T7 polymerase-synthesized rRNA was incubated for 25 min at 30°C in the ascites S-100 extract or a fraction thereof that elutes from a double-stranded DNA cellulose at 300 mM KCl (9). Heparin was then added to 50 or 16 μg/ml, respectively, and incubation was continued for another 10 min. The samples were then loaded onto a nondenaturing 4% polyacrylamide gel (acrylamide/bisacrylamide ratio, 65:1; preelectrophoresed for 60 min at 13 V/cm) and run at ca. 4 V/cm for 18 h at room temperature. The complexes were visualized by autoradiography at −70°C. When competitor RNAs were used, they were preincubated at a 12-fold molar excess in the extract for 10 min prior to addition of the substrate RNA.

Computer prediction of RNA secondary structure. The current RNA folding program of Zuker (23) that identifies suboptimal as well as optimal structures was run on a VAX 8600 with the energy values of Jaeger et al. (7). Structures within 5% of the optimal free energy were examined. The same major stem-loop structures were seen whenever the folded region contained the 200 nucleotides that made up the cleavage site and the conserved 3′ region.

RESULTS

Processing of pre-rRNAs with internal mutations. Our earlier analysis of 5′ and 3′ deletion mutants of mouse rRNA had defined the region required for efficient cleavage at the +650 primary processing site as residing within the adjoining 200-nucleotide conserved region (3). However, because of the nature of deletion analysis, it was not possible to determine whether the whole identified region was important or whether only a few small positions were critical and, if so, whether their spacing was flexible or fixed.

To examine these questions, we constructed small internal mutations (IM constructs) within this identified region by combining the 5′ and 3′ deletions (3) to produce constructs

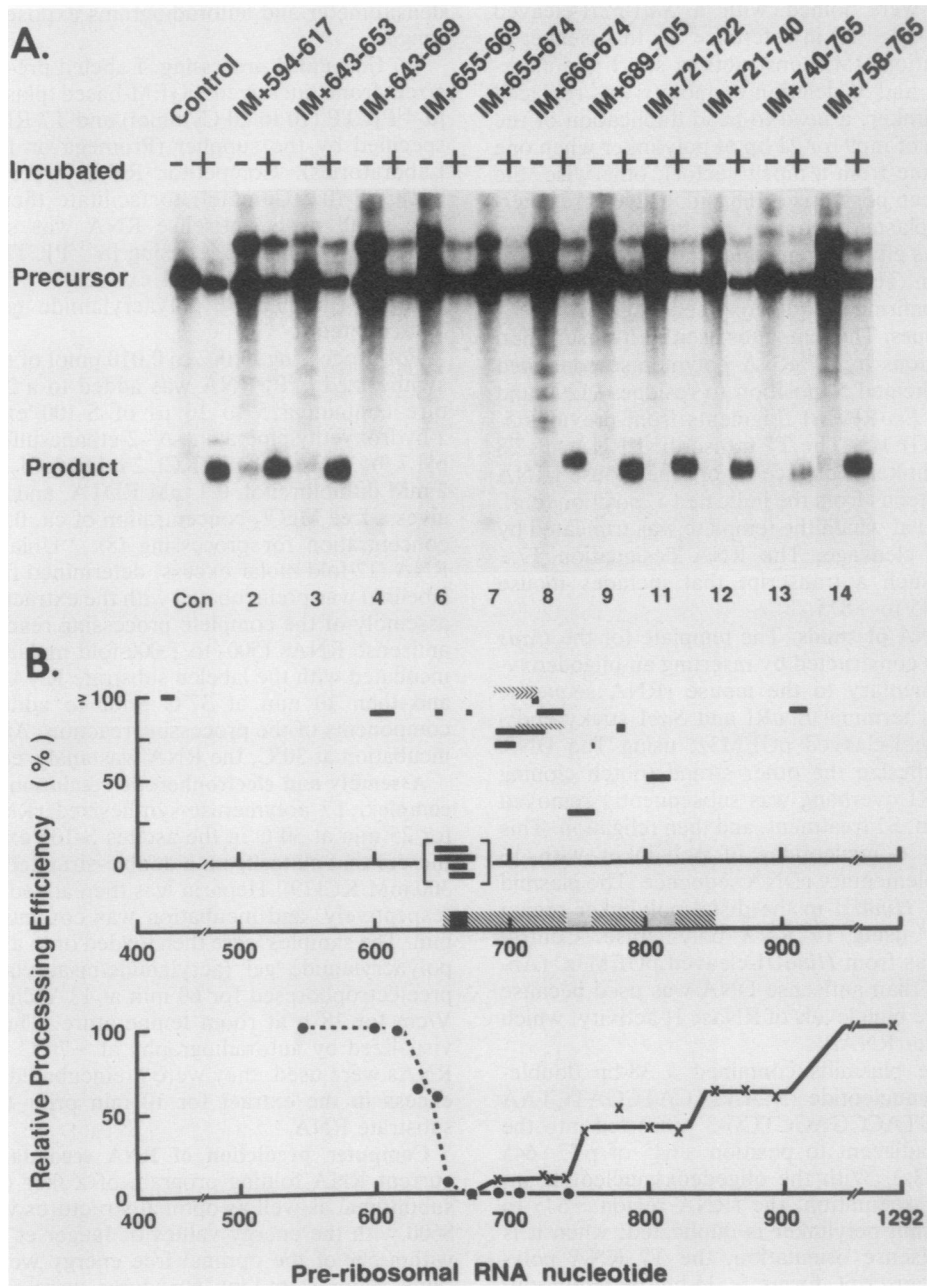

FIG. 1. Processing of mouse pre-rRNA with internal mutations. (A) Typical data. Each mutated plasmid, linearized just 3′ to nucleotide +1292, was used as a template in a mouse S-100 transcription reaction producing the ^{32}P-labeled pre-rRNA shown in lanes labeled −. An equal amount of the labeled pre-rRNA was then incubated to allow processing, as described in Materials and Methods (lanes labeled +). The numbers below each pair of lanes correspond to the designations in Table 1. Because the size alterations of the internal mutations are not identical (see Table 1), the mobilities of the rRNA precursor and processed product vary slightly. (B) Summary of the processing efficiency of each internally mutated pre-rRNA (upper graph) compared with the processing efficiency of the 5′ and 3′ deletion mutant pre-rRNAs (lower graph) (3). In the upper graph, the extent and location of each internal mutation are shown by the position of the bars (■) along the *x* axis, whereas the position of each bar on the *y* axis indicates the processing efficiency, relative to unmodified control pre-rRNA; ▨, ▧, duplications of rRNA sequence. In the middle line, the location of the 200-nucleotide mouse region that is conserved between humans and rodents is indicated, as are the location of the three regions in which mutations inhibit processing completely (■) or partially (▩) or not at all (▧).

with wild-type mouse rDNA from residues −168 to +1290, except that a short region between the deletion endpoints was replaced by polylinker (e.g., in IM+655/669 the nucleotides between +655 and +669 are replaced). Radiolabeled RNA was synthesized from these templates and then allowed to be processed during a cold chase period in the S-100 coupled transcription-processing system (3). As observed previously (3, 8, 13), the upstream rRNA cleavage

TABLE 1. Relative processing efficiency of precursor rRNA with internal mutations and duplications

rRNA no. (Fig. 1)	Template designation	rRNA nucleotides deleted or duplicated[a]	Processing efficiency relative to control[b]
1	p417/427[c]	−9 (+1)	1.3
2	p594/617	−22 (−12)	0.9
3	p643/653	−9 (+12)	0.9
4	p643/669	−25 (−4)	0
5	p643/674	−30 (−9)	0
6	p655/669	−13 (+8)	0
7	p655/674	−18 (+2)	0
8	p666/674	−7 (+14)	0.9
9	p689/705	−15 (−6)	0.7
10	p690/726	−35 (−28)	0.8
11	p721/722	−0 (+13)	1.1
12	p721/740	−18 (0)	0.9
13	p740/765	−25 (−14)	0.3
14	p758/765	−6 (+3)	0.7
15	p780/784	−3 (+7)	0.8
16	p801/819	−17 (−9)	0.5
17	p907/919	−11 (−2)	0.9
18	p721/689[d]	+33 (+61)	1.6
19	p721/705[d]	+17 (+37)	1.3
20	p740/705[d]	+36 (+45)	0.8

[a] The number in parentheses is the total spacing change in nucleotides.
[b] The processing efficiency of the control p5'Sal (+1 to +1292) was 44 ± 14% (n = 7).
[c] Template p417/427 has mouse rRNA nucleotides 1 to 417 and 427 to 1292, separated by the polylinker segment from the original M13 vectors. See Materials and Methods for details.
[d] Duplications.

product (residues +1 to ca. +650) is degraded virtually concomitantly with the processing reaction, so only the downstream product is observed. In all cases, the cleavage site was unchanged. Representative data showing the RNAs prior to and following the processing incubation are shown in Fig. 1A (lanes − and +, respectively), and the results of multiple experiments are summarized in Fig. 1B and Table 1.

Most of the assayed residues in the 200-nucleotide conserved region are not required for efficient cleavage, since small regions throughout can be altered without appreciable effects (Fig. 1; Table 1). In three small regions, however, mutations significantly affect the processing efficiency. The most critical of these is defined by IM+655/669, which includes the +656 processing site and directs no detectable processing. All larger IM sequences that contain this region are also inactive. In contrast, adjacent regions can be mutated without significantly inhibiting processing (e.g., IM+643/653 and IM+666/674). These data indicate that the critical residues reside between positions +654 and +666. A second important region is defined by IM+740/758, in which the processing efficiency is reduced to 30%, and a third region is defined by IM+801/819, which reduces the processing efficiency to 50%. IM sequences which map on either side of these regions have little to no effect on processing, suggesting that the ca. +750 and ca. +810 regions indeed are separate sites and are not part of one large contiguous recognition region.

The processing-active IM sequences that flank the critical ca. +660 region are insertions of 12 and 14 nucleotides, respectively, and other active IM sequences have length alterations of up to 28 nucleotides. Extensive duplications of the RNA sequence between residues +689 and +740 also have no adverse effect on the processing efficiency (Table 1). Therefore, the RNA spacing between the critical ca. +660 region and the downstream efficiency determinants at ca. +750 and ca. +810 cannot be rigidly constrained.

Minimal processing-competent rRNA region. Since internal mutations assess processing of large (1.3-kb) rRNAs altered only in small regions, if any required sequences were redundant, they would not have been identified. We therefore determined the minimal contiguous sequence capable of processing. To increase the sensitivity and flexibility of the assay, we used short, gel-purified mouse rRNA substrates that had been synthesized by using T7 RNA polymerase and were added to the standard in vitro processing system. As shown previously (8), exogenously added rRNA is processed identically to rRNA synthesized in the presence of the processing components in the coupled transcription-processing system.

The processing of T7-synthesized rRNAs with different 5' and 3' termini is shown in Fig. 2. Consistent with previous analysis in which large RNA molecules were used in the coupled transcription-processing system (3), the 5' boundary of the region required for processing maps between residues +645 and +669 (Fig. 2A). The template with the smallest 5' region that directs active processing (T7+645) was truncated at various positions to allow production of a nested series of 3' deletions, whose processing efficiencies are shown in Fig. 2B. To verify that RNA products derived from substrates with low processing efficiencies indeed arise from the processing under study, we demonstrated their specific competition (9) by supplementing the reaction with an excess of unlabeled processing-competent RNA (labeled Sp, for specific; T7+645/875) or processing-incompetent RNA (labeled Ns, for nonspecific; T7+669/875) (compare the second and third lanes of each set in Fig. 2B). Relative to T7+645/875, which directs active processing, T7+645/762 is 30% efficient, T7+645/741 and T7+645/693 are ca. 5% efficient, and T7+645/672 is still ca. 3% efficient at specific processing (Fig. 2B). T7+645/660 has never been observed to direct any specific processing. Thus, only 28 nucleotides of rRNA (which includes the essential ca. +660 domain) are needed to direct low levels of processing, and inclusion of the ca. +750 domain (identified as stimulatory above) causes a marked increase in processing efficiency.

Minimal region needed for complex formation. Processing-competent RNA that extends from residues +645 to +875 forms a specific complex with factors of the cell extract that migrates slowly in nondenaturing gels, whereas processing-incompetent rRNA that extends from +669 to +875 does not (9). To better define the borders of the region involved in complex formation and to correlate them with processing efficiency, we assayed the small T7 rRNAs for complex formation. Figure 3A confirms that T7+645/875 RNA is active for complex formation whereas T7+669/875 RNA is not. The presence of additional rRNA regions 5' of residue +645 does not appreciably increase the ability of RNA to form the specific complex (Fig. 3A). The specificity of the complex is confirmed since its formation is inhibited by processing-competent T7+645/875 RNA (lanes Sp) but not by the processing-incompetent T7+669/875 RNA (lanes Ns). To determine the 3' border for complex formation, the 3'-deleted RNAs of Fig. 2B were used. Relative to RNA extending to residue +875, which efficiently forms the specific complex, RNAs extending to residues +762, +741, and +693 or +672 are increasingly less efficient and RNA extending to residue +660 shows no detectable complex formation (Fig. 3B). Therefore, not only do processing and

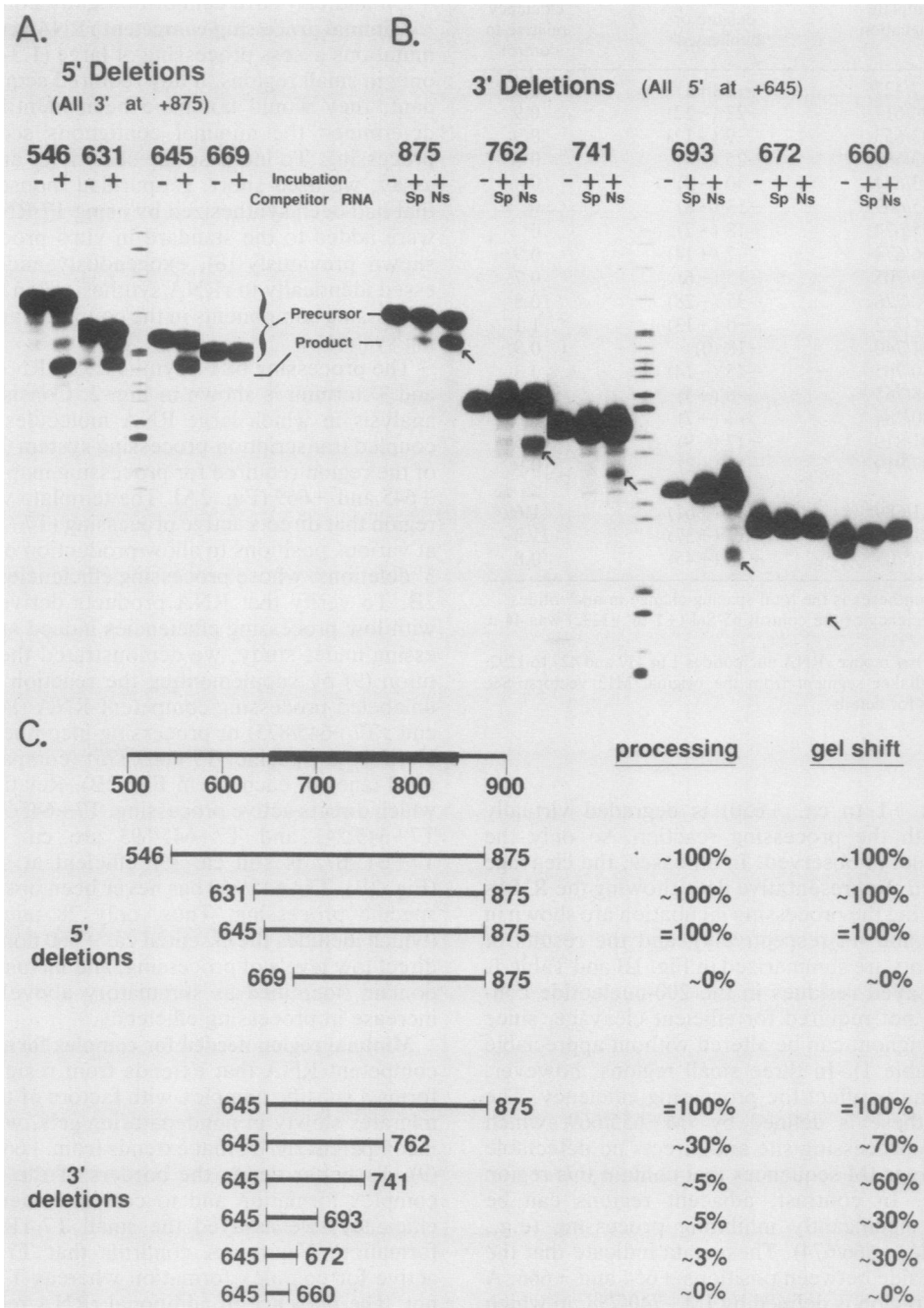

FIG. 2. Analysis of the minimum size of precursor rRNA required for processing. (A) Effect of varying the 5′ extent of pre-rRNAs on processing ability. Each 5′ deletion template (pT7+546, etc.; see Materials and Methods) was linearized at nucleotide +875 with *Ava*II before transcription by T7 RNA polymerase. Equal amounts of the labeled pre-rRNAs (shown in the lanes labeled −) were incubated in the standard S-100 processing system (shown in the lanes labeled +) before gel analysis. If pT7+669 RNA had processed at the position corresponding to 650 to 656, the resultant RNA would be 9 to 15 nucleotides shorter than the precursor, i.e., the same length as the processed RNA in the other lanes. The marker lane contains *Hpa*II-cleaved pBR322 DNA. (B) Effect of varying the 3′ endpoint of pre-rRNAs. Pre-rRNAs with different 3′ endpoints were transcribed from pT7+645 linearized at +875 (*Ava*II), +762 (*Apa*I), +741 (*Xma*III), +693 (*Eco*RII), +672 (*Mbo*II), or +660 (*Sau*3AI). These processing reactions were conducted in the presence of a 12-fold molar excess of either a specific (Sp) or nonspecific (Ns) competitor RNA, as described in the text. (C) Summary of multiple experiments measuring the relative efficiency of processing and specific complex formation (as in Fig. 3) of various 5′- and 3′-deleted pre-rRNAs.

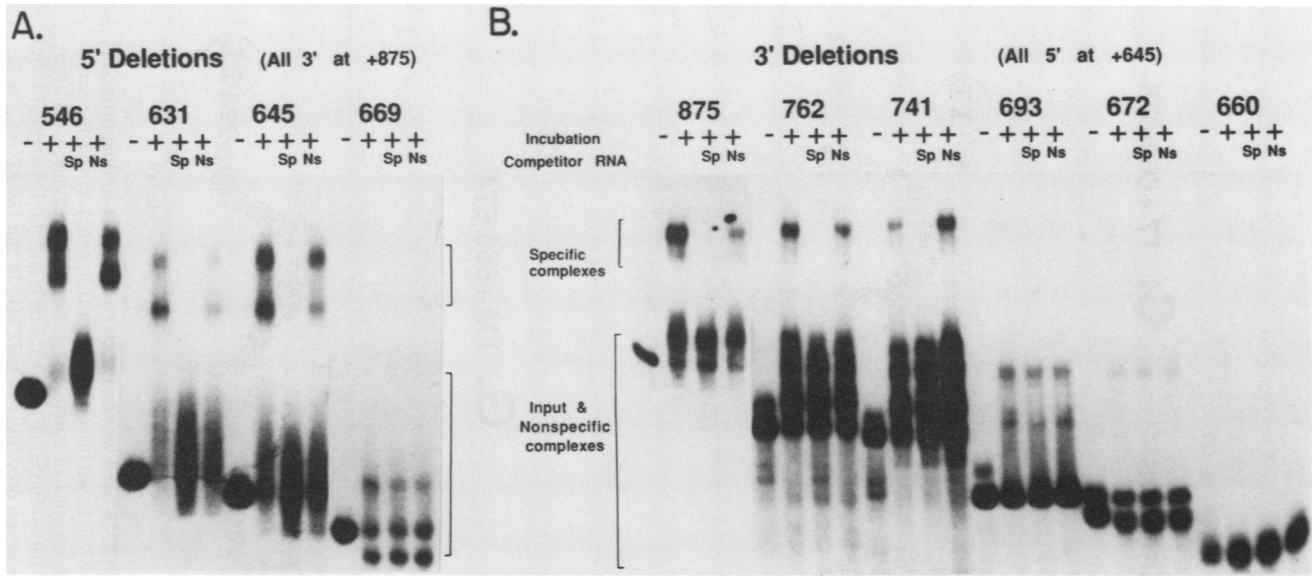

FIG. 3. The 5′ and 3′ boundaries of the pre-rRNA sequence necessary for specific complex formation. The same pre-rRNAs and the same specific (Sp) and nonspecific (Ns) competitor RNAs as in Fig. 2 were used for analysis of gel-shifted complexes.

complex formation require the same minimal sequences and both utilize 3′ stimulatory regions, but also the deleted rRNA substrates show a considerably greater impairment of processing than of complex formation. This suggests that complex formation may be necessary but not sufficient for the complete processing reaction.

Predicted secondary structure of the nucleotide sequence important for rRNA processing. To begin to examine the possible secondary structure of the nucleotide sequence required for efficient processing and complex formation, we used the program of Zuker (23) to predict the potential secondary structure of our rRNA substrates, both processing competent and processing incompetent. This program provides optimal as well as suboptimal energetic solutions, and, in conjunction with the energy values of Jaeger et al. (7), identifies >90% of the known tRNA helices within 5% of the lowest free energy (7). The predicted secondary structure of the natural rRNA extending from residues +600 to +875 (which includes the processing site and all the sequence required for efficient processing) is shown in Fig. 4A. Inclusion of extensive additional sequences further 5′ and 3′ does not change the predicted structure of this region (data not shown). Consistent with this predicted secondary structure, analysis of the homologous human and rat pre-rRNA-processing regions shows that their sequences superimpose on the predicted mouse secondary structure with the ca. 15% nucleotide changes in the 200-nucleotide conserved regions mainly in predicted single-stranded portions (loops and bulges) or, when in predicted duplex regions, mainly in pairs of compensatory base alterations or in an additional small stem-loop (data not shown). The nucleotides flanking the conserved 200-nucleotide rRNA region are predicted to be less highly base paired.

The structure of mouse pre-rRNA in this required region appears to have two domains. The first (nucleotides ca. +625 to ca. +680, designated as I in Fig. 4A) consists of a large articulated duplex stem that terminates in a single-stranded loop which contains the processing sites. Since it contains the minimal region required for processing and for complex assembly, it thus may represent a primary recognition domain. The second domain (residues ca. +680 to ca. +875, designated as II in Fig. 4A) has an extensive series of branched stems and loops and may be an efficiency domain, since it contains the +750 and +810 regions, which affect the processing efficiency, and since the predicted structure of this domain is most altered in the pre-rRNAs with less efficient processing (data not shown). Figure 4B to G shows the predicted structures of some of the mutant templates, focusing on the possible recognition domain. (In these structures, except for T7+645/672 [Fig. 4D], the predicted structure for domain II is the same as that shown in Fig. 4A, and so is represented by a II.) All of the processing-competent rRNAs, including the mutants shown in Fig. 4B, D, E, and G, yield a domain I structure in which the ca. +655 region is in a loop at the end of an articulated duplex stem, like the natural rRNA. In contrast, for the processing-incompetent rRNAs (T7+669/875, T7+645/666, and IM+655/669), the sequence whose position corresponds to ca. +655 is predicted to be missing or within a stem region (Fig. 4C and F; data not shown). This suggests that the occurrence of the +655 region in a single-stranded loop could be an important determinant for the rRNA-processing machinery.

Testing the recognition domain loop hypothesis with *trans* and *cis* antisense RNA. To examine whether sequestration of the +655 putative loop sequence in a base-paired conformation would inhibit processing, we used two types of antisense RNAs to alter the secondary structure of this region. The first was a short *trans* antisense RNA complementary to residues +649 to +664, which was synthesized by using T7 RNA polymerase, purified, and annealed to the rRNA substrate before addition of the processing extract. It markedly inhibited processing (Fig. 5, lane 4 versus lane 2). In contrast, a nonspecific RNA oligonucleotide caused no diminution of processing (Fig. 5, lanes 3 and 4). Furthermore, if the antisense RNA and the pre-rRNA were not annealed prior to addition of extract or if they were preincubated under conditions minimizing annealing (e.g., at 0°C), processing would remain efficient (data not shown), demonstrating

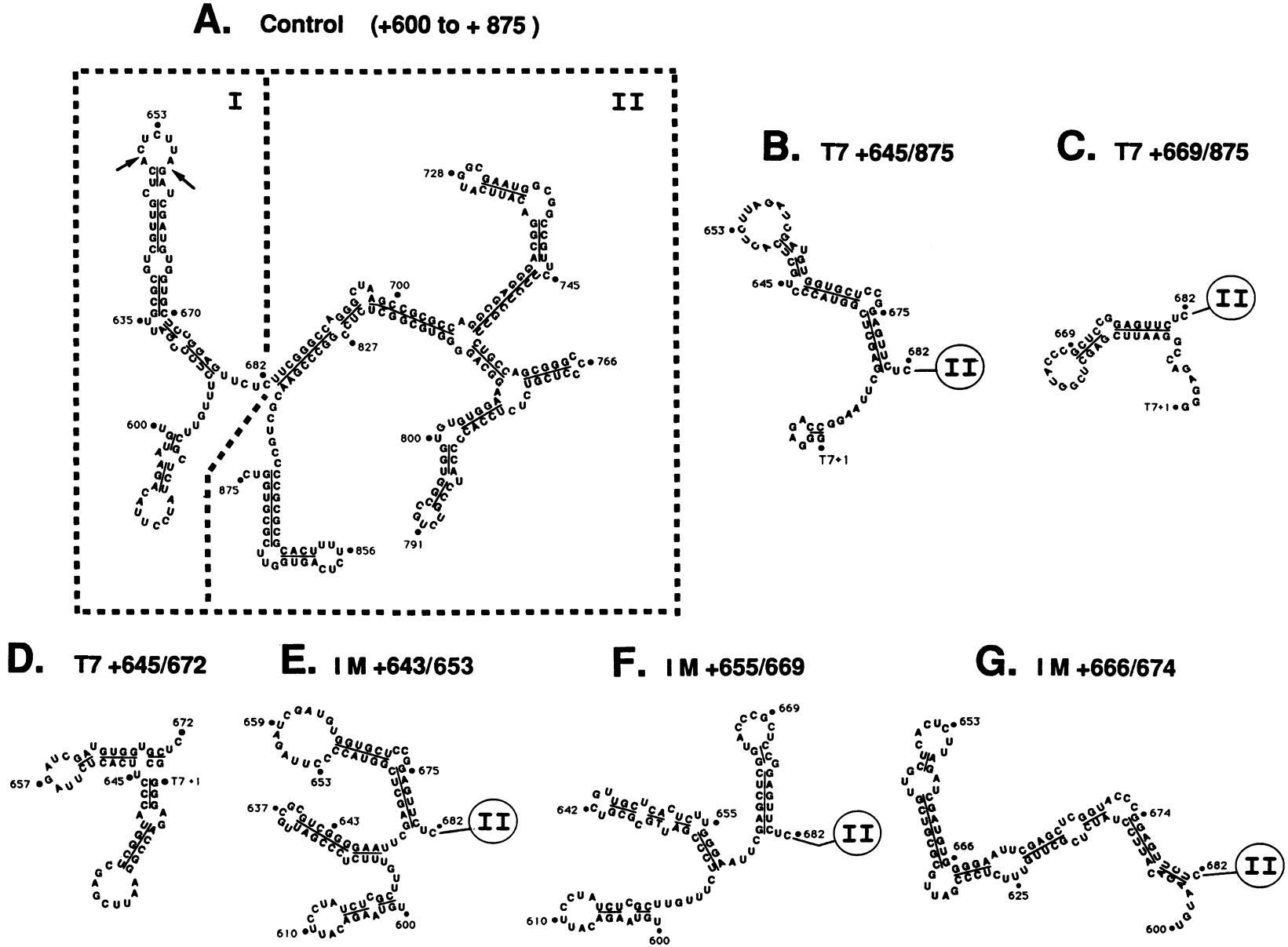

FIG. 4. Computer predictions of the secondary structures of pre-rRNAs. The region from +600 to +875 of each pre-rRNA used was analyzed by the RNA-folding program of Zuker (23), using the energy values of Jaeger et al. (7). The predicted structure extending from nucleotides +683 to +875 (shown boxed as region II in panel A) was the same for the RNAs of panels A to C and E to G, and it is represented by the symbol II in the latter panels. The arrows in region I of panel A represent the cleavage sites after nucleotide +650 and +656. (A) Unmodified, control pre-rRNA; (B) T7+645/875; (C) T7+669/875; (D) T7+645/672; (E) IM+643/653; (F) IM+655/669; and (G) IM+666/674. In the T7 RNAs, the nucleotides before the indicated 5′ rRNA position (+645 or +669) are derived from the T7 promoter and polylinker of pGEM1. In the IM RNAs, the 21 nucleotides between the indicated deletion endpoints are polylinker derived (see Materials and Methods).

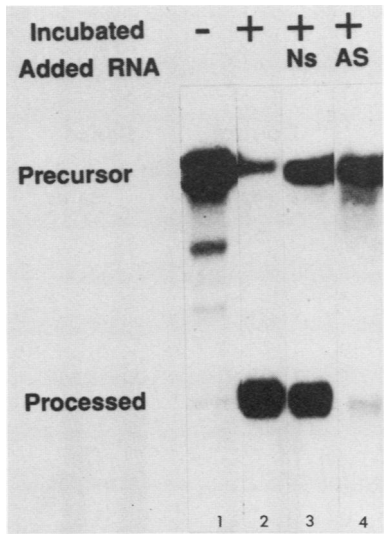

FIG. 5. Inhibition of rRNA processing by *trans* antisense RNA complementary to nucleotides +649 to +664. The T7+546/875 rRNA remained unincubated (lane −) or was incubated (lanes +) for 3 h at 30°C in the standard S-100 processing system after preincubation with either no added RNA or a 600-fold molar excess of either the 61-nucleotide nonspecific (lane Ns) RNA or the 71-nucleotide antisense (lane AS) RNA, which contain nucleotides complementary to rRNA positions +649 to +664, as described in Materials and Methods. The nonspecific RNA was synthesized from *Hin*dIII-cleaved pGEM3z.

specificity and suggesting that trapping the single-stranded loop in a duplex inhibited processing.

We have also examined the effects of a *cis* antisense RNA transcribed from a plasmid in which the antisense oligonucleotide was cloned 20 nucleotides upstream of the processing signal, causing the +645/664 region to be sequestered in a snapback duplex structure (Fig. 6A, part ii). This pre-rRNA is not processed (Fig. 6B, lanes 5 and 6). When the oligonucleotide was instead cloned into the rRNA plasmid at the same position but in the reverse (or sense) orientation, the resultant transcript carries a duplication of the +645/664 region (Fig. 6A, part i), and this sense RNA is efficiently processed at the natural +650 site (Fig. 6B, lanes 3 and 4). Thus, if the critical +645/664 region is base paired, the rRNA is not processed.

To determine whether the lack of processing of rRNA with +645/664 in a duplex structure correlates with a lack of formation of the specific nucleoprotein complex, we analyzed the *cis* antisense RNA (Fig. 6A, part ii) for complex-forming ability. This RNA is clearly deficient at complex formation (Fig. 6C, lanes 9 to 12), whereas the control RNA bearing the oligonucleotide sequence in the sense orientation is fully active (lanes 5 to 8). Most probably, a single-stranded conformation of the +645/664 region is a critical determinant for formation of the specific processing complex and this complex is essential for the processing reaction.

DISCUSSION

The first event in the maturation of mammalian rRNA occurs not at an end of one of the mature rRNA regions but within the 5′ external transcribed spacer, at ca. 20% of the distance between the initiation site and the 18S region (5, 13). Amphibian rRNA shows an analogous processing (15), and several lower eucaryotes also undergo processing in this region (reviewed in reference 10). In contrast to the rest of the external transcribed spacer, whose sequence diverges very rapidly in evolution, the 200 nucleotides 3′ of this primary processing site (which contain the sequences that direct the processing) are ca. 80% conserved between rodents and humans (3, 8), and the critical 5′ region is even conserved in *Xenopus* species. These data suggest that this processing serves an evolutionarily selective function. Our dissection of the sequence elements required for this processing provides the first determination of the minimal signal needed for any metazoan rRNA maturation event.

Two approaches have been used to determine the nucleotide sequences important for this primary processing of mouse pre-rRNA. Using a series of small internal mutations within a large, otherwise wild-type pre-rRNA (Fig. 1; Table 1), we found that much of the conserved region could be altered without affecting processing. However, mutations in the region +655 to +666 abolished processing, and alterations at nucleotides ca. +750 and ca. +810 partly inhibited processing. In contrast, mutations and spacing changes of more than 40 nucleotides between these three important regions did not substantially affect processing, indicating that their action is not highly dependent on their spacing in the linear rRNA.

The converse approach, determining the minimal contiguous nucleotide sequence capable of directing processing (Fig. 2), showed that a short rRNA region (residues ca. +645 to +672) containing the critical ca. +655 region was sufficient to direct processing, albeit at a low level. The low efficiency of this minimal processing region is not attributable to the small size of the T7 substrate rRNAs, because alternate substrates containing these segments embedded within larger regions of procaryotic sequence exhibited no greater processing (data not shown). A sixfold stimulation of processing was observed when the rRNA was extended in the 3′ direction from residues +741 to +762, consistent with the ca. +750 element identified with the internal mutations (Fig. 1C), and an additional threefold stimulation of processing by sequences still further 3′ is consistent with the ca. +810 element. These experiments extend our earlier analysis of large 5′ and 3′ deletions (3) and support the hypothesis that the +655 region is necessary for processing and sufficient to specify low levels of processing but that other regions within the conserved 200-nucleotide segment, notably at ca. +750 and ca. +810, substantially augment processing efficiency.

Mouse pre-rRNA has been shown to form a specific complex with factors of the cell extract that migrates slowly on electrophoretic gels (9). Although the formation of the complex by the various rRNA substrates is confirmed to qualitatively correlate with their processing efficiencies (Fig. 3), complex formation is considerably less sensitive to the mutations than is the complete processing reaction (Fig. 2C), suggesting that formation of the nucleoprotein complex is necessary but not sufficient for the processing cleavage.

Computer-assisted secondary-structure analysis of the pre-rRNA substrate (Fig. 4) suggested that the processing region is made up of two distinct and highly structured domains whose existence is consistent with phylogenetic data obtained from similarly analyzing the analogous human and rat processing regions. In all of the processing-competent RNAs, a presumptive specificity domain which contains the minimal processing signal shows the ca. +655 processing region in the loop at the end of an extensive duplex stem; in the processing-incompetent rRNAs, this analogous region is

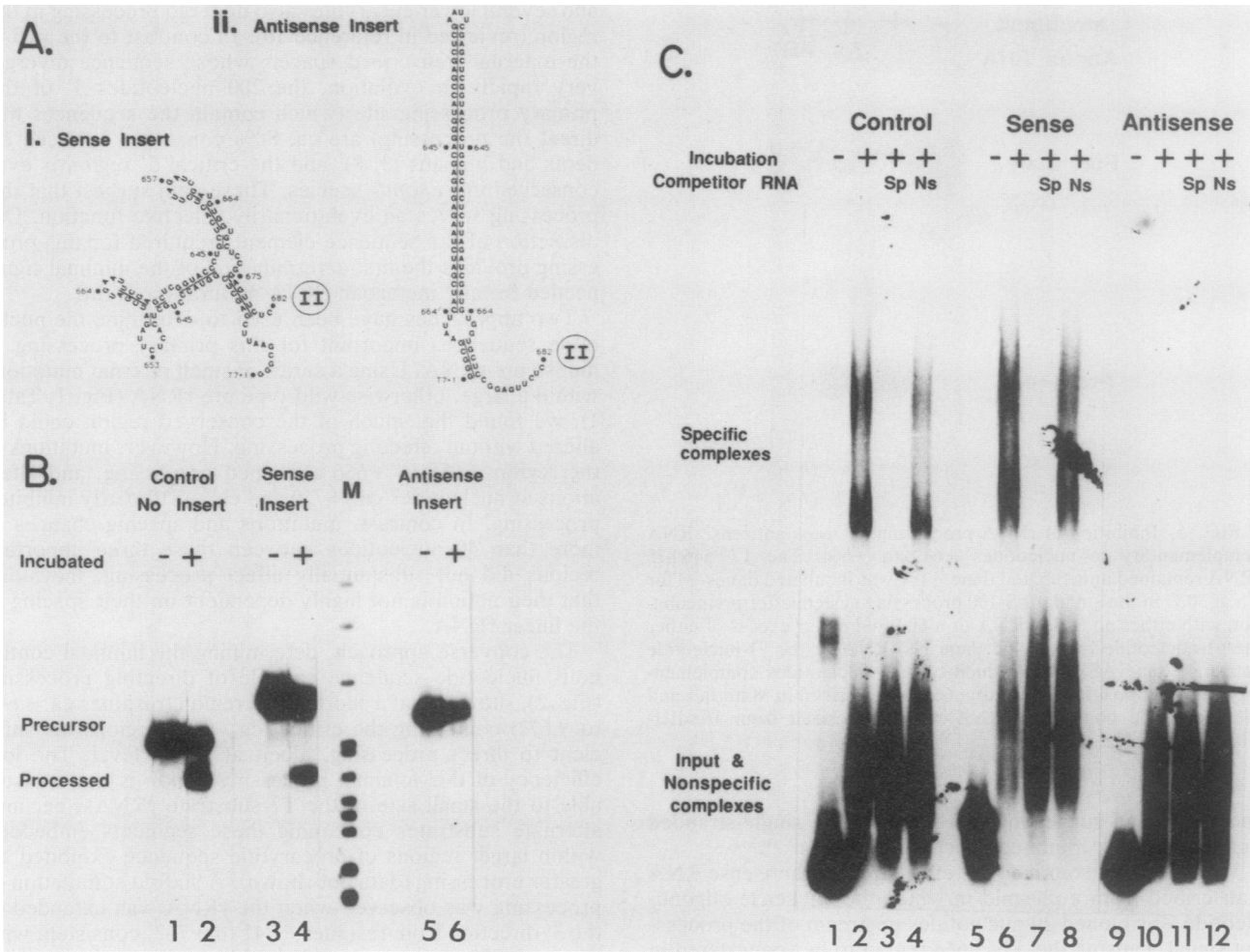

FIG. 6. Effect of *cis* antisense RNA on rRNA processing and on the formation of the specific processing complex. (A) Computer prediction of the secondary structure of the T7 *cis* RNAs (see Materials and Methods) containing the antisense oligonucleotide sequence in the sense orientation, forming a duplication of nucleotides +650 to +661 (panel i), or in the antisense orientation, forming a hairpin (panel ii). Compare this with the proposed structure of the rRNA with no insert (Fig. 4B). (B) Processing ability of pre-rRNA with no insert, the sense insert, and the antisense insert. The marker is the *Hpa*II-cleaved pBR322 DNA. (C) Formation of the specific rRNA-processing complex by using a control transcript with no insert, or transcripts bearing the sense insert or the antisense insert. The methods were as described for Fig. 2 and 3.

within a base-paired structure. The central and 3′ portion of the 200-nucleotide conserved region, including the two downstream stimulatory regions, is predicted to be part of a branched, highly base-paired efficiency domain, whose structure is significantly altered by the mutations in this region that diminish processing efficiency.

To assess the importance of ca. +655 region being single stranded, we placed residues +645 to +664 in duplex structure, both by using short antisense oligoribonucleotides added in *trans* and by using rRNA substrates with the antisense sequence present in *cis* (Fig. 5 and 6B). In both cases processing was selectively abolished, demonstrating the importance of the availability of the ca. +655 processing signal in a single-stranded conformation. The formation of the specific processing complex was also found to be crucially dependent on the ca. +655 region being single stranded (Fig. 6C). Thus, even within the context of a large surrounding rRNA region, the ca. +655 region is critical for complex formation.

The precise role of the +655 region in directing processing remains to be determined. Since the U3 small nuclear ribonucleoprotein particle is required for this rRNA-processing event (10), U3 RNA could a priori duplex with the single-stranded +655 region. However, no extensive or evolutionarily conserved duplex has been found between this region and the apparent single-stranded regions of U3 (10, 20; data not shown), suggesting that this small nuclear ribonucleoprotein particle may use other recognitions for binding, possibly involving proteins that bind directly to the +655 region. Thus, the +655 single-stranded region may have a different function. One possibility comes from the work of Shumard and Eichler (18), who reported that an extensively purified single-strand-specific nucleolar RNase could cleave the +650 region as well as a number of other upstream sites in purified mouse pre-rRNA. If that RNase functions in the primary processing in vivo, the single-stranded configuration of this rRNA region would be crucial.

It appears that the primary rRNA-processing event in-

volves a considerable number of cellular factors, in addition to the U3 small nuclear RNA (10). At least six polypeptides, most of which differ in size from the identified U3 proteins and from the nucleolar single-strand-specific nuclease, can be specifically cross-linked to the pre-rRNA substrate in the processing complex (9). Indeed, this processing complex may constitute the large 5′ terminal knobs characteristic of pre-rRNAs in electron-microscopic preparations (Miller spreads; discussed in reference 10). Determination of how the rRNA directs the primary processing should provide important information in the eventual elucidation of the mechanism of this rRNA maturation event.

ACKNOWLEDGMENTS

We thank Cathy Martin for help in the gel shift experiment of Fig. 6. A few of the lanes of Fig. 3B were previously shown in reference 9.

This research was supported by American Cancer Society grant NP-731A.

REFERENCES

1. **Bowman, L. H., W. E. Goldman, G. I. Goldberg, M. B. Hebert, and D. Schlessinger.** 1983. Location of the initial cleavage sites in mouse pre-rRNA. Mol. Cell. Biol. **3**:1501–1510.
2. **Carter-Muenchau, P., and R. E. Wolf.** 1987. A method for cloning mixtures of long, synthetic oligodeoxynucleotides. Gene Anal. Tech. **4**:105–110.
3. **Craig, N., S. Kass, and B. Sollner-Webb.** 1987. Nucleotide sequence determining the first cleavage site in the processing of mouse precursor rRNA. Proc. Natl. Acad. Sci. USA **84**:629–633.
4. **Crouch, R.** 1984. Ribosomal RNA processing in eukaryotes, p. 214–226. *In* D. Apririon (ed.), Processing of RNA. CRC Press, Inc., Boca Raton, Fla.
5. **Gurney, T.** 1985. Characterization of mouse 45S ribosomal RNA subspecies suggests that the first processing cleavage occurs 600±100 nucleotides from the 5′ end and second 500±100 nucleotides from the 3′ end of a 13.9 kb precursor. Nucleic Acids Res. **13**:4905–4919.
6. **Hannon, G. J., P. A. Maroney, A. Branch, B. J. Benenfield, H. D. Robertson, and T. W. Nilsen.** 1989. Accurate processing of human pre-rRNA in vitro. Mol. Cell. Biol. **9**:4422–4431.
7. **Jaeger, J. J., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86**:7706–7710.
8. **Kass, S., N. Craig, and B. Sollner-Webb.** 1987. Primary processing of mammalian rRNA involves two adjacent cleavages and is not species specific. Mol. Cell. Biol. **7**:2891–2898.
9. **Kass, S., and B. Sollner-Webb.** 1990. The first pre-rRNA processing event occurs in a large complex: analysis by gel retardation, sedimentation, and UV cross-linking. Mol. Cell. Biol. **10**:4920–4931.
10. **Kass, S., K. Tyc, J. A. Steitz, and B. Sollner-Webb.** 1990. The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. Cell **60**:897–908.
11. **King, T. C., R. Sirdeskmukh, and D. Schlessinger.** 1986. Nucleolytic processing of ribonucleic acid transcripts in procaryotes. Microbiol. Rev. **50**:428–451.
12. **Mandal, R. E.** 1984. The organization and transcription of eukaryotic ribosomal RNA genes. Prog. Nucleic Acid Res. Mol. Biol. **31**:115–156.
13. **Miller, K. G., and B. Sollner-Webb.** 1981. Transcription of mouse rRNA genes by RNA polymerase I: in vitro and in vivo initiation and processing sites. Cell **27**:165–174.
14. **Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. **15**:8783–8798.
15. **Mougey, E., L. Pape, and B. Sollner-Webb.** Unpublished data.
16. **Perry, R. P.** 1981. RNA processing comes of age. J. Cell Biol. **91**:28S–38S.
17. **Raziuddin, R. D. Little, T. Labella, and D. Schlessinger.** 1989. Transcription and processing of RNA from mouse ribosomal DNA transfected into hamster cells. Mol. Cell. Biol. **9**:1667–1671.
18. **Shumard, C. M., and D. C. Eichler.** 1988. Ribosomal RNA processing: limited cleavages of mouse preribosomal RNA by a nucleolar endoribonuclease include the early +650 processing site. J. Biol. Chem. **263**:19346–19352.
19. **Shumard, C. M., C. Torres, and D. C. Eichler.** 1990. In vitro processing at the 3′-terminal region of pre-18S rRNA by a nucleolar endoribonuclease. Mol. Cell. Biol. **10**:3868–3872.
20. **Stroke, I. L., and A. M. Weiner.** 1989. The 5′ end of U3 snRNA can be crosslinked in vivo to the external transcribed spacer of rat ribosomal RNA precursors. J. Mol. Biol. **210**:497–512.
21. **Vance, V. B., A. E. Thompson, and L. H. Bowman.** 1985. Transfection of mouse ribosomal DNA into rat cells: faithful transcription and processing. Nucleic Acids Res. **13**:7499–7513.
22. **Young, R., and J. A. Steitz.** 1978. Complementary sequences 1700 nucleotides apart form a RNase III cleavage site in E. coli ribosomal precursor RNA. Proc. Natl. Acad. Sci. USA **75**: 3593–3597.
23. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244**:48–52.