Drosophila 5 S RNA Processing Requires the 1–118 Base Pair and Additional Sequence Proximal to the Processing Site*

(Received for publication, August 9, 1991)

Peter R. Preiser‡ and Louis F. Levinger§
From the School of Life and Health Sciences, University of Delaware, Newark, Delaware 19716

Using an in vitro processing system, we have identified a required sequence surrounding the Drosophila melanogaster 5 S RNA processing site at nucleotide 126. Mutations in this region vary the processing rate from complete inhibition to a level equal to or greater than wild type. Analysis of mutants at +1 and in the region 118–122 separates the inhibitory effect into two parts. 1) Nucleotide 118 C, the base-paired nucleotide in helix I proximal to the processing site, plays an essential role. Changing it to a purine inhibits processing. The +1–118 base pair must be intact, but this alone is not sufficient for processing, since compensatory changes at +1 do not restore down-processing mutants at 118 to the wild type level. 2) The processing site has to be pyrimidine rich; multiple contiguous purines inhibit processing. On the other hand, multiple pyrimidines can largely negate the inhibitory effect of a purine at position 118. Thus a base-paired C at 118 followed by a stretch of pyrimidines is the processing signal, which may be recognized by the processing enzyme and/or a required accessory factor.

*This research was supported by grants from the University of Delaware Research Foundation, §National Institutes of Health Biomedical Research Support Grant and general program funds, and by Minority Biomedical Research Support Grant GM08153. 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.

†Present address: Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545.
‡Present address: Dept. of Natural Sciences/Biology, York College, City University of New York, Jamaica, NY 11431. To whom correspondence should be addressed.
§The abbreviations used are: nt, nucleotide; sub, substitution; wt, wild type.

5 S RNA is a component of the large ribosomal subunit (Hardicry and Kramer, 1986). In Drosophila melanogaster as well as other organisms, 5 S RNA is transcribed as a precursor and subsequently processed into mature 5 S RNA (Rubin and Hoggard, 1957; Jacq et al., 1977; Hamada et al., 1979; Tekamp et al., 1980; Rinke and Steitz, 1982; Brow, 1987). Thus, 5 S RNA processing may serve as a step in ribosome biogenesis.

In D. melanogaster the 5 S RNA primary transcript is 135 nt long with a stretch of 5 Us at its 3' end (Rubin and Hoggard, 1975). Processing removes 15 single-stranded nucleotides from the 3' end, resulting in a 120-nt mature 5 S RNA. Mature 5 S RNA consists of five phylogenetically conserved stem loops (Garrett et al., 1981). Using a T7 promoter/5 S RNA gene construct to obtain 5 S RNA primary transcript, we have studied the nucleotide sequences that influence processing in vitro. We have previously shown that processing is inhibited when all or part of domain II or III of mature 5 S RNA is deleted, while domain IV and half of domain V are dispensable, and that the 3' terminal U, tract inhibits processing (Preiser and Levinger, 1991).

Here we show that the nucleotides surrounding nucleotide 120 influence processing. Nucleotide substitutions in this region show the paramount importance of nucleotide 118C, the first base-paired nucleotide of helix I. Changing this nucleotide to a purine greatly reduces processing. Multiple downstream contiguous purines also strongly inhibit processing. On the other hand, keeping the region pyrimidine rich can partially negate the inhibitory effect of a purine at 118. Disruption of the 1–118 base pair by substitutions at +1 in wild type RNA disrupts processing, while 118 mutants with impaired processing cannot be restored to wild type processing ability by compensatory base changes at +1.

EXPERIMENTAL PROCEDURES

Cell Culture and Extract Preparation—D. melanogaster Kc16 cells adapted to growth in serum-free medium were carried in continuous culture using Eschhaar's D22 medium, and S100 extracts were prepared as previously described (Eschalier and Ohanesian, 1970; Price et al., 1987; Dingermann et al., 1981).

Construction of a Chimeric 5 S RNA Gene Linked to a T7 Promoter—Overlapping synthetic oligonucleotides were used to construct the T7/5 S gene as previously described (Preiser and Levinger, 1991) and cloned into the high copy number plasmid pH624 (Boros et al., 1984). A DraI site at +135 relative to the transcription start site at +1 allows production of a runoff 5 S RNA primary transcript identical in sequence to that found in vivo.

Construction of 3' End Deletions—Two approaches were used to obtain 3' end deletions. The deletion to nucleotide 118 made use of a Huell site at +118. T7/5 S DNA was digested with HaelIII and EcoRI, the insert was gel purified and subcloned into the EcoRI-Smal site of pH624. For the deletion to nucleotide 122, 10 µg of pHC T7/5 S DNA were digested with DraI followed by limited digestion with Ba31 to obtain a set of 3' end deletions. The DNA was blunt-ended with Klenow fragment and redigested with EcoRI using standard procedures (Maniatis et al., 1982). Fragments in the proper size range were gel purified and subcloned into the EcoRI-Smal site of the pH624 polylinker.

Plasmid DNA Preparation—Plasmid DNA was prepared from 40 ml of an overnight culture as described (Kraft et al., 1988). Dideoxy sequencing of supercoiled plasmid DNA (Sanger et al., 1977; Lim and Pene, 1988) was performed using Sequenase (US. Biochemical Corp., Cleveland, OH).

Mutagenesis Strategy—Mutant 5 S RNA genes were made by primer extension with mismatched oligonucleotides (Preiser and Levinger, 1991; Kunkel, 1985; Simmons et al., 1990). Mutation sub 118–1122, replacing the sequence CCCUG—GGGAU (see Fig. 5A), was made first, followed by single point mutations of all five nucleotides to their complements. Since the results indicated that pyrimidine content might be important, both in the subsequent mutations of the wt sequence (CCUCG) and the lack mutations of sub 118–122 (GGGAU), purines were replaced by pyrimidines, and vice versa.

Oligonucleotides were about 30 nt long with up to five mismatches. For template we used U-substituted Stratagene Bluescript M-13.
derivative Sk- which contained either the wild type T7/5 S or the sub 118–122 insert. The +1 substitutions were made from four different T7/5 S genes: wt, 118 C→G, 118,119 CC→AA, sub 118–122; the latter three were chosen for their reduced processing ability (Fig. 3 and Table I). Templates subcloned into pH624 were mutagenized by inverse polymerase chain reaction (Hemsley et al., 1989). The upstream mutagenic primers were from +15 → −12 including most of the T7 promoter with a single nucleotide mismatch at +1, and the universal downstream primer was from +16 → 38. Transformants were sequenced on both strands to confirm the +1 substitutions and to screen out occasional undesired substitutions introduced elsewhere by Taq DNA polymerase (obtained from U. S. Biochemical or from D. Engelke, University of Michigan (Engelke et al., 1990)).

**T7 RNA Polymerase Reactions**—The 3′ end deletion constructs (Fig. 1) were digested with HindIII (+160), while for all the other mutants the DraI run-off site at +135 was used. Labeled 5 S RNA was transcribed with T7 RNA polymerase and gel purified as previously described (Preiser and Levinger, 1991). For +1 substitutions, we observed the yield order G>A>C>>U, as reported earlier (Milligan et al., 1987).

**5′ End Analysis**—The +1 nucleotide in wild type and +1 G→C RNA was analyzed by 5′ end labeling, digestion with S1 nuclease, and one-dimensional thin layer chromatography as previously described (Preiser and Levinger, 1991); we determined that >95% of the labeled nucleotide at +1 was as expected for each template (data not shown).

**Processing Reaction**—5000 cpm of radioactively labeled RNA were incubated with 20 μl of S100 extract in a final volume of 40 μl as previously described (Preiser and Levinger, 1991); we determined that >95% of the labeled nucleotide at +1 was as expected for each template (data not shown).

**Processing Rate**—The +1 substitution was made from four different T7/5 S genes: wt, 118 C→G, 118,119 CC→AA, sub 118–122; the latter three were chosen for their reduced processing ability (Fig. 3 and Table I). The upstream mutagenic primers were from +15 → −12 including most of the T7 promoter with a single nucleotide mismatch at +1, and the universal downstream primer was from +16 → 38. Transformants were sequenced on both strands to confirm the +1 substitutions and to screen out occasional undesired substitutions introduced elsewhere by Taq DNA polymerase (obtained from U. S. Biochemical or from D. Engelke, University of Michigan (Engelke et al., 1990)). The +1 substitutions were made from four different T7/5 S genes: wt, 118 C→G, 118,119 CC→AA, sub 118–122; the latter three were chosen for their reduced processing ability (Fig. 3 and Table I). The upstream mutagenic primers were from +15 → −12 including most of the T7 promoter with a single nucleotide mismatch at +1, and the universal downstream primer was from +16 → 38. Transformants were sequenced on both strands to confirm the +1 substitutions and to screen out occasional undesired substitutions introduced elsewhere by Taq DNA polymerase (obtained from U. S. Biochemical or from D. Engelke, University of Michigan (Engelke et al., 1990)). The +1 substitutions were made from four different T7/5 S genes: wt, 118 C→G, 118,119 CC→AA, sub 118–122; the latter three were chosen for their reduced processing ability (Fig. 3 and Table I). The upstream mutagenic primers were from +15 → −12 including most of the T7 promoter with a single nucleotide mismatch at +1, and the universal downstream primer was from +16 → 38. Transformants were sequenced on both strands to confirm the +1 substitutions and to screen out occasional undesired substitutions introduced elsewhere by Taq DNA polymerase (obtained from U. S. Biochemical or from D. Engelke, University of Michigan (Engelke et al., 1990)).

**RESULTS**

We have previously used a T7 promoter/5 S gene construct to study in vitro 5 S RNA processing in D. melanogaster cells (Preiser and Levinger, 1991). Sequences past the normal transcription termination site at nt 135 have little effect on the processing rate.

**A Processing Signal Surrounds the Processing Site**—Using Bal31 we progressively deleted part of the 3′ end sequences found in the wt 5 S RNA primary transcript and replaced them with apparently innocuous sequence by using a HindIII runoff site in the polylinker of pH624 (see “Experimental Procedures”). The wt T7/5 S-HindIII runoff RNA was processed efficiently (Fig. 1A), as were 3′ end deletions up to nt 122 (Fig. 1B). Deletion to the HaeIII site at nt 122, which has been replaced by the polylinker sequence GGGAU in Fig. 1C, is important for processing. This suggests that the sequence CCUCG (nt 118–122), which has been replaced by the polylinker sequence GGGAU in Fig. 1C, is important for processing.

We used mismatch oligonucleotide mutagenesis to produce a 5 S RNA primary transcript in which the only change is the substitution of the sequence GGGAU at nt 118–122 (sub 118–122) for CCUCG. Sub 118–122 5 S RNA was not processed (Fig. 2).

**Mutations in the Region 118–122 Show That 118 C Is Critical for Processing**—Important nucleotides in region 118–122 were identified by mutagenesis. Single, double, and triple mutations which disrupt the wt sequence have a wide range of effects (Table I). Mutation of nucleotide 118 C→G (Fig. 3A) reduced processing to about 27% of wt (Fig. 2B), while deletion of nucleotides 118 and 118 C→U had practically no effect (Fig. 3B and Table I). Changing both nucleotides 118,119 CC→AA (Fig. 3D) almost completely inhibited processing (8% of wt), while the effect of mutation of nucleotides 119,120 CU→AA (Fig. 3C) was considerably less (77% of wt).

**Table I**

| Change | Sequence (118–122) | Processed | Adjusted to wt = 1.00 |
|--------|-------------------|-----------|----------------------|
| None   | Yes               | 1.00      |
| sub 118–122 | GGGAU               | No        | 0.90                  |
| 118 C→G | GGGCU               | Yes       | 0.27                  |
| Δ118    | CUGU                | Yes       | 1.04                  |
| 118 C→U | UCUCG               | Yes       | 1.12                  |
| Δ118→122| UCCAC               | Yes       | 0.41                  |
| 119 C→G | CGUCG               | Yes       | 0.69                  |
| 120 U→A | CCACG               | Yes       | 0.71                  |
| 121 C→G | CCUGG               | Yes       | ND                    |
| 122 G→C | CCUCG               | Yes       | 1.60                  |
| 118,119 CC→AA | AUUCG             | Very weak | 0.08                  |
| 119,120 CU→AA | CAACG             | Yes       | 0.77                  |
| 120,121 UC→AA | CCAAG             | Yes       | ND                    |
| 121,122 CG→AC | CCUCG             | Yes       | 1.69                  |
| 121,122 CG→AU | CCUAU             | Yes       | 0.57                  |

**Back Mutations from Sub 118–122**—Back mutations from sub 118–122 (Fig. 2A) also resulted in a wide range of effects (Table II). Point mutant 118 G→C (Fig. 2B) restored about
The sequence of nt 118–122 is given for all the mutants, along with the ratio of mutant to wt RNA processing rate (see “Experimental Procedures”).

| Change  | Sequence (118–122) | Processed | Adjusted to wt = 1.00 |
|---------|--------------------|-----------|----------------------|
| Sub 118–122 | GGGAU | No | 0.00 |
| 118 G → C | CGGAU | Yes | 0.39 |
| 119 G → C | CGGAU | No | 0.00 |
| 119 G → U | GUGAU | Weak | 0.13 |
| 120 G → U | GUGAU | Weak | ND |
| 121 A → C | GGCCU | Weak | 0.19 |
| 122 U → A | GGGAA | Weak | 0.10 |
| 118,119 GG → UU | UUGAU | Weak | 0.05 |
| 118,119 GG → UC | UUGAU | Weak | 0.13 |
| 119,120 GG → UC | UUGAU | Weak | 0.75 |
| 120,121 GA → CU | GGCUU | Weak | 0.16 |
| 120,121 GA → UC | GGCUU | Weak | 0.27 |
| 121,122 AU → CA | GGGCA | Weak | 0.27 |
| 118–120 GGG → CUU | CUUAU | Yes | 1.17 |
| 118–120 GGG → UCC | UUCAU | Yes | 1.14 |
| 119–121 GGA → UCC | GUUCCU | Yes | 0.60 |
| 119–121 GGA → CUU | GUUCCU | Yes | 0.82 |
| 120–122 GAU → UCA | GGUCA | Weak | 0.36 |

Fig. 4. Processing of 5 S RNA containing back mutations from nt 118–122 GGGAU. Mutants were made by introducing back mutations in the sub 118–122 sequence GGGAU. Panel A, 119 G→U. Lanes 1–5, gel-purified Dra I runoff RNA was processed for 0, 15, 30, 60, and 120 min. Panel B, same as panel A, except mutation 119,120 GC→AC. Panel C, same as panel A, except mutation 119,120 GC→AC. Panel D, same as panel A, except mutation 118,119 CC→AA.

39% of wt processing activity; on the other hand, change of nucleotide 119 G→U (Fig. 4A) only restored about 19% processing activity while 119 G→C did not restore processing (data not shown). Mutation of nucleotides 119,120 GG→UC (Fig. 4C) restored about 75% of processing, similar to the triple mutation 119–121 GGA→UCC (Fig. 4E). Change in nucleotides 120–122 GAU→UCA (Fig. 4D) resulted in only a moderate improvement in processing (37% of wt).

Compensatory Substitutions at +1 Do Not Restore 118 Mutants to Wild Type Processing—We made the substitutions +1G→A, C, and U in wild type, 118C→G, 118,119 CC→AA, and sub 118–122 5 S RNA genes (Fig. 5A). These templates (the latter three chosen for their reduced processing ability) were transcribed with T7 RNA polymerase and processed. All three substitutions which disrupt the 1–118 base pair in wt 5 S RNA strongly reduce processing (Fig. 5B), while the compensatory base substitutions in 118 C→G and 118,119 CC→AA mutant RNAs with reduced processing did not restore activity to the wild type level (Fig. 5C, lanes 11–15; Fig. 5D, lanes 16–20; cf. lanes 1–5 and Fig. 3A). The +1 mutants with sub 118–122 gave similar results (data not shown).

We note a slight but reproducible improvement in processing rate with +1G→C–118C→G RNA compared with 118C→G (Fig. 5C, lanes 11–15; cf. lanes 1–5). All 118C→G-processed RNAs display a litter of bands between precursor and mature product which is atypical of wild type 5 S RNA processing (Fig. 5C; cf. Fig. 5B, lanes 1–5; see “Discussion”).

DISCUSSION

Importance of 118C, the Base-paired Nucleotide Proximal to the Processing Site—Deletion analysis shows that nucleotides 118–122 influence the in vitro processing of D. melanogaster
5 S RNA (Fig. 1). Mutagenesis within this region indicates that this signal is more complex than just one or two specific nucleotides (Figs. 2–5). Changes in most nucleotides affect the overall processing rate, with sub 118–122 (CCUCG-GGGAU) completely inhibiting processing (Fig. 2A). Nucleotide 118, the first base-paired nucleotide of helix I in 5 S RNA, is important for processing. A single change from 118 C→G strongly reduces processing; no other point mutation has so great an effect (Table I). Change of nt 118 C→U is without effect (see below), while the double substitution 118,119 CC→AA almost completely inhibits processing.

Similarly, using sub 118–122 as the starting point, 40% of processing can be restored by changing 118 G→C (Fig. 4B), while other single back mutations have a minimal effect (Table II). Under these conditions, a C seems to be required at position 118, since change to a U does not restore processing to a comparable level. Back mutations beyond nucleotide 118 C→G further increase the processing rate.

The 1–118 GC Base Pair Is Necessary but Not Sufficient for Processing—Considering that nt 118, the base-paired nucleotide of stem I proximal to the processing site at 120 (Fig. 5A), is the most important nucleotide in the region 118–122, we investigated the effect of base pairing between +1 and 118 in wild type and several of the 5 S RNAs with severely reduced processing (Fig. 5, B–D and data not shown). Disruption of the 1–118 base pair in wt 5 S RNA inhibits processing (Fig. 5B). Compensatory substitutions at +1 in RNAs with processing defects at 118, on the other hand, do not restore processing to the wild type level (Fig. 5C and data not shown).

5 S RNA Processing in Drosophila and Other Organisms—In this important respect, D. melanogaster 5 S RNA processing differs from that in Bacillus subtilis (Stahl et al., 1989), in which nucleotides are removed from both the 3′ and 5′ ends by the endonuclease Rnase M5. B. subtilis 5 S RNA processing is greatly affected by mutations in the base-paired nucleotide stranded by the cleavage site. Compensatory mutations which restore base pairing largely reverse this effect in B. subtilis (Stahl et al., 1980), but less so in D. melanogaster (Fig. 5).

In Xenopus laevis, a 3′→5′ exonuclease processes some forms of 5 S RNA. This enzyme stops at the double-stranded boundary of helix I, and its disruption allows further degradation of the RNA (Xing and Worcel, 1989). Since in D. melanogaster several mutations around the processing site, including disruption of the 1–118 base pair, render the RNA refractory to processing without a great increase in nonspecific degradation, a 3′→5′ exonuclease, if active, would appear to be tightly regulated.

A litter of bands from 127 to 120 and at 111 nt in the patterns obtained with 118C→G RNA (Fig. 5C) show that a single nucleotide change can qualitatively affect processing. This could result from suppression of a more prominent endonuclease so that a weaker and less precise exonuclease is observed, analogous to the in vivo results with Saccharomyces cerevisiae (Piper et al., 1983, 1984).

Pyrimidine Richness of the Processing Site—While nucleotide 118C plays a dominant role in processing, the effect of a purine can be reduced if nucleotides 119 and 120 are pyrimidines (Table II). On the other hand, introduction of a purine in position 119 or 120 of the wt sequence reduces processing by about 30%.

The cis-acting processing signal surrounding the processing site could be due to local stacking of the run of three pyrimidines (118–120) starting with the +1–118 GC base pair. Stacking could also explain why the weaker GU base pair is tolerated at 1–118 (Table I). Alternatively, a pyrimidine-rich sequence could fit better than purines into a nucleic acid active site. A pyrimidine-rich region is common to the mature 3′ end of eukaryotic 5 S RNAs (8).

Trans-activation of 5 S RNA Processing—Trans effects arising from a finely tuned nuclease or 5 S RNA-specific accessory factor could account for these observations. Indeed, binding of large ribosomal protein BL16 is required for B. subtilis 5 S RNA processing (Stahl et al., 1984). The requirement for stems II and III for D. melanogaster 5 S RNA processing (Preiser and Levinger, 1991) is also consistent with this possibility. Signals which could identify the binding site for a nuclease and/or a functionally significant binding protein appear to be widely dispersed throughout 5 S RNA.

Acknowledgments—We are grateful to V. Greene for technical assistance, J. Pene for oligodeoxynucleotide synthesis (University of Delaware core facility), T. Pederson for a critical reading of the manuscript, and P. Romanik for helpful discussions.

REFERENCES
Boros, I., Posfai, F., and Venetianer, F. (1984) Gene (Amst.) 30, 257–260
Brow, D. A. (1987) J. Biol. Chem. 262, 13959–13965
Dingermann, T., Sharp, S., Appel, B., DeFranco, D., Mount, S., Heiermann, R., Pongs, O., and Soll, D. (1981) Nucleic Acids Res. 9, 3901–3918
Engelke, D. R., Krikos, A., Bruct, M. E., and Ginsburg, D. (1990) Anal. Biochem. 190, 396–400
Eschallier, G., and Ohannessian, A. (1970) In Vitro 6, 162–172
Garrett, R. A., Douthwaite, S., and Noller, H. F. (1981) Trends Biochem. Sci. 6, 137–139
Hamada, H., Muramatsu, M., Urano, Y., Onishi, T., and Kominami, R. (1979) Cell 17, 163–173
Hardesty, B., and Kramer, G. (1986) Structure, Function and Genetics of Ribosomes, Springer-Verlag, New York
Hemsley, A., Arnhem, N., Toney, M. D., Cortopassi, G., and Galas, D. J. (1989) Nucleic Acids Res. 17, 6545–6551
Jacq, B., Jourdan, R., and Jordan, B. R. (1977) J. Mol. Biol. 117, 785–795
Kraft, R., Tardiff, J., Krauter, K. S., and Leinwand, L. A. (1988) Biotechniques 6, 544–549
Kunkel, T. A. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 488–492
Lim, H. M., and Pene, J. J. (1988) Gene Anal. Techn. 5, 32–39
Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Mulligan, J. F., Groesbeck, M., Withrell, G. W., and Uhlenbeck, O. C. (1987) Nucleic Acids Res. 15, 8783–8798
Piper, P. W., Bellatin, J. A., and Lockeart, A. (1983) EMBO J. 2, 353–359
Piper, P. W., Patel, N., and Lockeart, A. (1984) Eur. J. Biochem. 141, 115–118
Preiser, P. R., and Levinger, L. F. (1991) J. Biol. Chem. 266, 7509–7516
Price, D. H., Sluder, A. E., and Greenleaf, A. L. (1987) J. Biol. Chem. 262, 3244–3255
Rinke, J., and Steitz, J. A. (1982) Cell 29, 149–159
Rubin, G. M., and Hoggins, D. S. (1975) Cell 6, 207–213
Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
Simmons, D. T., Loeber, G., and Teptmeyer, P. (1990) J. Virol. 64, 1973–1983
Stahl, D. A., Mayback, B., and Pace, N. R. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 5644–5648
Takahashi, P. A., Galea, R. L., and Rutter, W. J. (1980) J. Biol. Chem. 255, 9501–9506
Xing, Y. Y., and Worcel, A. (1989) Genes & Dev. 3, 1008–1018