Structure and Transcription of the *singed* Locus of *Drosophila melanogaster*

Jamie Paterson and Kevin O'Hare

Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ, England

Manuscript received February 18, 1991
Accepted for publication August 8, 1991

ABSTRACT

Developmental and genetic studies of the *singed* gene of *Drosophila melanogaster* indicate that the gene has a role in somatic cells during the formation of adult bristles and hairs, and in the female germline during oogenesis. During metamorphosis a single 3.6-kilobase (kb) RNA is made, and this RNA is also present in adults and early embryos. Early embryos and adult females have additional 3.3- and 3.0-kb RNAs. The RNAs differ only in the length of the 3′ untranslated region and a single gene product of 57 kilodaltons is predicted. Analysis of RNA from females lacking ovaries suggests that the 3.3- and 3.0-kb RNAs are made only in ovaries. The absence of the 3.3- and 3.0-kb RNAs in pupae and the time course of their appearance in adult females after eclosion suggests that transcription of *singed* in the ovary is from middle to late stages of oogenesis. Analysis of RNA in embryos from the reciprocal crosses between wild type and *singed-3* showed that all three RNAs are maternally inherited with very little zygotic transcription in embryos. The mutation *singed-3* appears to separate the two requirements for *singed* function as it has an extreme effect upon bristle development, but does not obviously affect oogenesis. In *singed-3*, there is a deletion at the 5′ end of the gene, but the coding region is intact. Transcription in *singed-3* is from a cryptic promoter in the upstream flanking sequences which is sufficiently active during oogenesis for fertility, but less active than the wild-type promoter during metamorphosis. The role of the single *singed* gene product may be in the asymmetric organization and/or movement of cytoplasmic components.

THE *singed* (*sn*) gene of *Drosophila melanogaster* has a role in both somatic and germ cells. The bristles and hairs found over much of the body of wild-type flies are shortened, or twisted and gnarled in *singed* mutants. This phenotype is most easily seen in the large bristles (macrochaetes) on the dorsal surface of the thorax of the fly, but the smaller bristles (microchaetes) and hairs are also affected. Alleles can be distinguished as having extreme or weak affects on bristles and hairs. The phenotype is autonomously expressed, and somatic *sn* mutant clones are easily recognized in a wild-type background.

In addition to this bristle phenotype many *sn* mutants are female sterile. There is no effect on male fertility. Ovary transplantation showed that *sn* is expressed autonomously in the ovary (CLANCY and BEADLE 1937). Mutant *sn* germline clones do not make eggs, indicating a requirement for *sn* expression in the germline during oogenesis (PERRIMON and GANS 1983). The ovaries of female sterile *sn* mutants have few late stage egg chambers, and few eggs are laid. The eggs are flaccid with shortened filaments, and they do not develop. For most *sn* mutants, there is a correlation between the effect a mutation has on female sterility and its effect upon bristle morphology. Extreme bristle mutants (such as sn^1 and sn^{X2}) are female sterile while weak bristle mutants (such as sn^2) are female fertile. A notable exception to this is sn^3, which is one of the most extreme bristle mutants, but is fertile.

Many mutations reduce female fertility, but most of these are weak alleles of vital genes rather than mutations of genes with a specific function in the ovary (PERRIMON *et al.* 1986). No lethal alleles have ever been described for *sn*, and mutations at *sn* selected as female sterile mutations do have bristle phenotypes (PERRIMON and GANS 1983). Developmental genetic studies have not clarified the relationship between the role of *sn* in the development of bristles and hairs, and its role during oogenesis. To gain more insight into this, we need to understand how the gene is expressed in somatic and germ cells, and the nature of the gene product (s).

We have reported the isolation of sequences from the *sn* locus by chromosomal jumping and detected alterations in the DNA maps of some *sn* mutants (ROIHA, RUBIN and O'HARE 1988). We report here a molecular analysis of the structure of the gene and its pattern of transcription throughout development. This leads to the conclusion that there is only one *sn* gene product, and that mutations such as sn^3, which separate the bristle and fertility phenotypes, affect regulation of gene expression rather than the structure of the gene product.

MATERIALS AND METHODS

Drosophila stocks and genetics: *Drosophila melanogaster* stocks were reared on standard cornmeal/treacle/yeast me-

dium at 25°. In analyzing *sn* mutants for DNA lesions, we have found differences between stocks of supposedly the same allele. A stock of sn^2 analyzed earlier (ROIHA, RUBIN and O'HARE 1988) showed no differences from wild type in its DNA map, while a second stock analyzed here did (see RESULTS). We assume that this weak allele of *sn* has been lost from some stocks. We suspect that there may be problems with stocks of other *sn* mutants. Some mutants originally described as sterile appear in stocklists as apparently female fertile stocks. The wild-type P cytotype strain, Harwich, was from M. G. KIDWELL (University of Arizona). All other strains were from the Bowling Green Drosophila Stock Center.

The reciprocal crosses between Canton S (wild-type M cytotype) and Harwich (wild-type P cytotype) were made at 29°. The progeny were allowed to mature in fresh yeasted bottles for 3 days at 25° before harvesting. This regime allows the ovaries of females to develop fully so that egg chambers of all stages are present. Examination of these bottles showed that the progeny from the cross of Canton S females with Harwich males were sterile, while the progeny of the cross of Harwich females with Canton S males were fertile. Dissection showed that the sterile females lacked ovaries, as expected for this dysgenic cross.

The reciprocal crosses between Canton S and $y\ sn^3\ v$ were made at 25°. Embryos were collected after overnight laying (0–16-hr embryos) and at the end of the day (0–8-hr embryos), and pooled. Some embryos were allowed to develop and the genotypes were as expected. Heterozygous sn^+/sn^3 and sn^3/sn^+ females were harvested and shown to have both wild-type and mutant RNAs (data not shown, see RESULTS).

Molecular biology: Genomic DNA was prepared from frozen adult flies as described by LEVIS, BINGHAM AND RUBIN (1982). The construction of genomic libraries for sn^3 and sn^{X2}, screening of genomic and cDNA libraries, purification of recombinant clones, subcloning and DNA blotting were all by standard techniques (SAMBROOK, FRITSCH and MANIATIS 1989). The wild-type (Oregon R) library was that of MARIANI, PIRROTTA and MANET (1985) and cDNA libraries prepared from staged poly(A^+) RNA were those of POOLE *et al.* (1985). Preparation of total RNA from embryos, larvae, pupae and adults, selection of poly(A^+) RNA, and electrophoresis on formaldehyde agarose gels was as described by LEVIS, O'HARE and RUBIN (1984).

Double-stranded DNA probes were prepared as described by FEINBERG and VOGELSTEIN (1984). Strand-specific DNA probes for hybridizing to RNA blots were prepared by the method of BURKE (1984). A complementary radiolabeled copy of the inserts in M13 clones was synthesized using Klenow fragment or T7 DNA polymerase in the presence of dTTP, dGTP, dATP and [α-^{32}P]dCTP (Amersham). The products of the extension reaction were digested with a restriction enzyme at the end of, or at a site within, the insert, denatured in 0.1 M NaOH and electrophoresed on a 1.5% low-gelling-temperature agarose gel (FMC BioProducts) containing 30 mM NaOH and 10 mM EDTA. The single-stranded labeled DNA was located by autoradiography, excised, melted and added directly to the hybridization. RNA blots were reprobed with either the *yolk protein 1* gene (HUNG and WENSINK 1981) to control for sex, or the 5C *actin* gene (FYRBERG *et al.* 1983) to control for loading.

Regions of DNA to be sequenced were either subcloned into M13 vectors as restriction enzyme fragments, or DNA fragments were first progressively shortened using *Bal*31 nuclease (PONCZ *et al.* 1982). Fragments treated with exonuclease were phenol extracted and ethanol precipitated before ligation. To determine the terminal sequences of cDNAs, the cDNA fragment was subcloned into pBluescribe

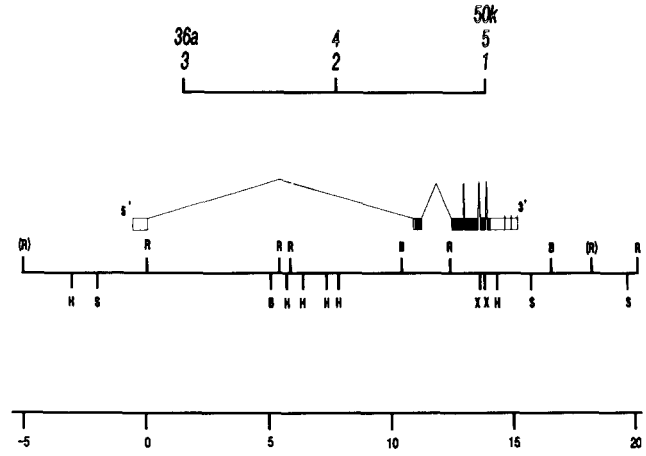

FIGURE 1.—Genetic and DNA map of the *singed* region. The genetic map of the *singed* locus is from HEXTER (1957). Distal to proximal with respect to the centromere is from left to right. Alleles with detectable DNA lesions are shown dark while those without detectable DNA lesions are in light letters. The DNA map was derived by restriction mapping overlapping phage containing inserts from either Canton S or Oregon R. The coordinate system, in kilobases, is that of ROIHA, RUBIN and O'HARE (1988), and refers only to the DNA map. The genetic map is shown for comparison with the DNA map, but cannot be precisely aligned. The orientation with respect to the chromosome was determined from analysis of chromosome inversions (for details see ROIHA, RUBIN and O'HARE 1988). Restriction enzyme sites are: R, *Eco*RI; H, *Hin*dIII; S, *Sal*I; B, *Bam*HI; X, *Xho*I. Restriction enzyme sites in brackets are polymorphic. The intron-exon map is derived from cDNA, sequencing and RNA blotting data. Noncoding sequences are open boxes and protein coding regions are filled boxes. The alternate 3′ ends are marked by vertical lines in the final exon.

(Stratagene) and sequencing template was prepared as described by GUO, YANG and WU (1983). DNA sequencing was by the dideoxy chain terminating method (BANKIER and BARRELL 1983). Reactions were performed using the Klenow fragment of DNA polymerase (Amersham) or T7 DNA polymerase (United States Biochemicals or Pharmacia).

Consensus sequences were built up from individual gel readings using the MicroGenie software package (Beckman). The DNA sequences have been entered into the EMBL Nucleotide Sequence Database as accession numbers X17548, X17549 and X17550. The sequences were analyzed using the SEQNET facility of the SERC Daresbury Laboratory. The nucleic acid sequence of cDNA clone P5 was compared with the GenBank and EMBL nucleic acid data bases using FASTA. The longest open reading frame in clone P5 was translated and the predicted amino acid sequence compared to the NBRF and SWISS protein databases using FASTA. This peptide sequence was also compared with the combined protein and nucleic acid OWL data base using PROSRCH at the Biocomputing Research Unit in Molecular Biology of the University of Edinburgh.

RESULTS

Molecular mapping of the *singed* locus: We have reported the use of chromosome inversions to clone sequences from the *singed* locus. A 35-kilobasepair (kb) region (−17 to +18 in Figure 1) was mapped and its orientation with respect to the centromere was determined (ROIHA, RUBIN and O'HARE 1988). DNA

blotting showed that two mutations which had been mapped by recombination to the distal end of the locus (HEXTER 1957), differed from wild type in the interval −1.2 to 0.0. In sn^3 there is a 0.3-kb deletion in the interval between −0.9 to 0.0, while in sn^{36a} there is a 5.5-kb insertion between −1.2 and 0.0 (ROIHA, RUBIN and O'HARE 1988). As no differences were detected for mutations which mapped to central or proximal sites within *sn*, it seemed possible that only the distal end of the gene had been cloned. We therefore extended the cloned interval in the proximal direction to +27 by chromosome walking using a wild-type Oregon R library.

We have analyzed *sn* mutants by DNA blotting using subclones covering the region from −5.0 to +20. A few restriction enzyme site polymorphisms were noted (sites in parentheses in Figure 1). Furthermore, the sizes of a number of fragments in the interval +10.0 to +15.0 were variable in mutants, but also between the wild-type strains Canton S and Oregon R. We have not precisely localized these differences, but as *sn* RNAs from the two wild-type strains comigrate (data not shown), we presume that these differences in the DNA maps are due to small differences in the sizes of introns of *sn*. In the course of this work, the 0.1-kb insertion mapped between −1.2 and −1.9 in sn^{X2} (ROIHA, RUBIN and O'HARE 1988) was found to be present in the parental *In(1)dl49* chromosome. This insertion is upstream from the position where transcription starts (see below) and is a polymorphism between wild-type chromosomes rather than the cause of the mutant phenotype in sn^{X2}. As no other differences were detected in sn^{X2}, we conclude that there is a point mutation elsewhere in the gene.

The 0.3-kb deletion reported for sn^3 between −1.2 and 0.0 (ROIHA, RUBIN and O'HARE 1988) was examined by cloning and DNA sequencing. The deletion is from −0.8 to −0.5 and corresponds to deletion of nucleotides 1184 to 1475 of the wild-type sequence (shown below in Figure 4A). A DNA breakpoint associated with the chromosome rearrangement $Dp(1;3)sn^{13a1}$ was mapped between 0.0 to +5.5 (data not shown). This duplication stock is derived from the extinct transposition stock $Tp(1;3)sn^{13a1}$ (VALENCIA 1966) which broke the *X* chromosome within *sn*. In contrast to our earlier results we did detect a DNA alteration in a stock of sn^2. This is between +8.0 and +10.5, and is a large insertion or a cytologically undetectable rearrangement (data not shown). This allows us to extend our comparison of the genetic and physical maps of *sn* (Figure 1). The distal allele sn^3 has a deletion from −0.8 to −0.5 while the central allele sn^2 has an alteration in the region +8.0 to +10.5. This indicates that mutations in the proximal alleles, sn^1, sn^5 and sn^{50k}, lie to the right of +8.0, and that at a minimum, the *sn* gene spans from −0.5 to +8.0.

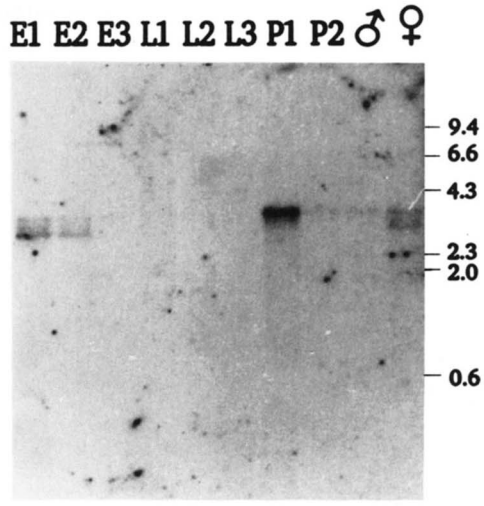

FIGURE 2.—Profile of *singed* RNA during development. Poly(A⁺) RNA samples from different stages of development were probed with sequences from the *sn* locus (−2.0 to 0.0 of Figure 1). E1, 0–4-hr embryos; E2, 4–8-hr embryos; E3, 0–24-hr embryos; L1, 1st instar larvae; L2, 2nd instar larvae; L3, 3rd instar larvae; P1, early pupae; P2, late pupae; ♂, adult males; ♀, adult females. Lambda DNA digested with *Hin*dIII was used as molecular weight markers. Similar quantities of RNA were loaded in each lane as judged by hybridization to a 5C *actin* probe (data not shown).

Transcription of *singed* during development: From the phenotypes of *sn* mutants, expression of the gene is expected during metamorphosis and in adult females. A probe from −2.0 to 0.0, the region altered in sn^3, sn^{36a} and *P* element insertion alleles of *sn* (ROIHA, RUBIN and O'HARE 1988), was hybridized to an RNA blot of poly(A⁺) RNA made from various stages during Drosophila development. The results (Figure 2) show a 3.6-kb RNA, most abundant in pupae and present in embryos and adults of both sexes. RNAs of 3.0 and 3.3 kb are present in embryos and in adult females. Using strand specific probes we demonstrated that all three RNAs are transcribed from distal to proximal (left to right in Figure 1). The abundance of these RNAs is quite low, and we have generally used single-stranded probes to increase the sensitivity of detection. For example, compare Figure 2 (double-stranded probe) with Figures 5 and 6 (single-stranded probe).

With a more distal DNA fragment as a probe, −5.0 to −2.0, we detect a 5.4-kb RNA transcribed from proximal to distal (data not shown). This is the opposite direction to the 3.0-, 3.3- and 3.6-kb RNAs and is away from the region disrupted in *sn* mutants. This transcript has been shown to be from a gene with a role in olfaction, the *olfE* gene (HASAN 1990).

Extent of the *singed* transcription unit-RNA analysis: The approximate 5′ and 3′ ends of the three *sn* RNAs were mapped by blotting RNAs from male and female adults. For the 5′ end, no probe was found which discriminated between the three RNAs. They

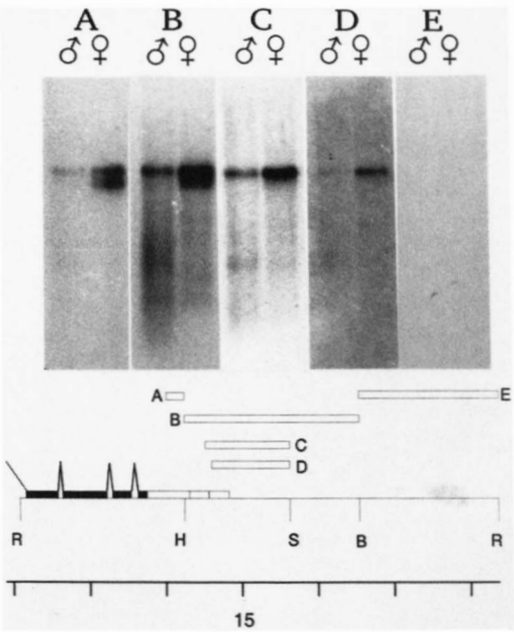

FIGURE 3.—Mapping the 3′ ends of *singed* transcripts. Panels A to E show RNA blots of male and female RNA hybridized to the corresponding strand specific probe shown in the diagram below.

either detected all three RNAs or none of them (data not shown). Two probes whose proximal limits differed by only 100 basepairs (bp) were most informative. The first probe (−1.8 to −0.6) detected all three RNAs, while the second probe (−2.7 to −0.7) did not detect any *sn* transcripts. This suggests that the 5′ ends of all three RNAs map between −0.7 and −0.6. If the 5′ ends of all three RNAs are within such a small interval, then it is most likely that they have the same 5′ end and differ elsewhere, at the 3′ end and/or internally due to differential splicing.

To map the 3′ ends of the RNAs, a series of more proximal probes were used. Probes from within the interval −0.6 to +14.3 hybridized to all three RNAs, or to none of them. These probes (*e.g.*, probe A from +14.0 to +14.3 in Figure 3A) showed approximately equal hybridization to the 3.0- and 3.6-kb RNAs and less hybridization to the 3.3-kb RNA. Different profiles of hybridization were detected using probes from +14.3 to +18.3 (Figure 3). A probe from +14.3 to +16.6 hybridizes most strongly to the 3.6-kb RNA, less strongly to the 3.3-kb RNA and least strongly to the 3.0-kb RNA (Figure 3B). Only the 3.6-kb RNA is clearly detected using probes from +14.5 to +15.6 or +14.6 to +15.6, while no RNA is detected with a more distal probe from +16.6 to +18.3 (Figure 3). From analysis of cDNAs (see below), we believe that the 3′ end of the 3.0-kb RNA maps to +14.3, the 3′ end of the 3.3-kb RNA to +14.55 and the 3′ end of the 3.6-kb RNA to +14.8. If this is so, then probe C from +14.5 to +15.6 should detect the 3.3-kb RNA in Figure 3C. Hybridization of this probe to the more abundant 3.6-kb RNA (and any degraded 3.6-kb RNA) may be obscuring detection of the 3.3-kb RNA. Furthermore, the short region of homology (57 bp) and the high A-T content (46/57) of the expected hybrid between probe C and the 3.3-kb RNA make it impossible to draw a definitive conclusion. Nonetheless, these results suggest that the 3′ ends of the RNAs map to different positions within the interval +14.3 to +16.6.

Extent of the *singed* transcription unit–cDNA analysis: To define more precisely the limits and structure of the *sn* transcription unit, cDNA clones were isolated from adult female, embryonic and pupal cDNA libraries (POOLE *et al.* 1985). Initially, a probe from the proximal region of the gene (−2.7 to +5.7) was used. A total of 22 clones were isolated from four different libraries. Two of the cDNAs were from *olfE*, and 4 of the 20 *sn* cDNA inserts were smaller than 1.5 kb. The remaining 16 were mapped and their maps can be aligned with that of the largest clone, the 2.9-kb pupal cDNA P5. The other cDNAs appear to be terminally deleted with respect to P5 at one or both ends.

The regions of genomic DNA corresponding to the ends of the P5 cDNA were mapped by probing DNA blots of the cloned genomic interval. This placed the 5′ end of P5 at −0.7 and the 3′ end at +14.4 (Figure 1). The genomic sequence around these end points was determined and compared to the terminal sequences of the different cDNAs. Four more independent cDNAs, 2 embryonic and 2 pupal, have the same 5′ end as P5 (Figure 4A), suggesting that transcription initiates at this site.

The sequence at the 3′ end of P5 maps to +14.4. However, sequences more proximal than +14.4 are present in the 3.6 kb RNA (Figure 3), so the 3′ end of this cDNA cannot correspond to the 3′ end of the RNA. The cDNA ends in a stretch of oligo-A (A_{16}) which is longer than the corresponding run of A residues encoded in the genome (A_{14}–Figure 4C). We believe that the 3′ end of P5 was generated during cDNA synthesis by oligo-dT priming on this internal A rich region of the mRNA rather than on the 3′ poly(A) tail.

To isolate cDNAs containing the 3′ sequences of *sn* RNAs, the cDNA libraries were rescreened with a probe from the genomic interval +14.3 to +15.7 and five clones were isolated. Compared to P5, these 3′ cDNAs are truncated at their 5′ ends, but have 3′ extensions. DNA sequence analysis of these extensions showed that they end in stretches of oligo(A) and correspond to three different sites in the genomic DNA sequence (Figure 4C). The separation of these three sites (about 250 bp) is consistent with their being used to generate the three sizes of *sn* RNA. We conclude that there is a single site for initiation of transcription of *sn* at −0.7, but three sites for poly-

Section A

```
                                                  50                                                 100
gtcgactttagaacggatatttcaaaataatgttaacctttcgaaaaatttaggtatatgcaagccgaattatttgatttggaaaataatacttagggcg

                                                 150                                                 200
tccatatcagaaaaaacgtttagagtaacataacgtattttaaaatttgtattaggccacgttatgtagtcaaatatatataaataaattattcatttta

                                                 250                                                 300
atttttacaaaagtgatatgctgatatgcttataaatgtaagattcttgtttatttgactttttttttaaatttcattatcattatttcttgcttgacaa

                                                 350                                                 400
ccctatgttaagtgcgacactgcgttgacaacagctaagaacatggctgtggggcgacaaacttagcagagcgcgcgctgctagagcttaaattgtattt

                                                 450                                                 500
gaaaacagtgcaggtgaatcgaaaatgtcttggcatgcagtcaaacattaatttgcgagcataattaactccgcttaagattgctggtagtcaaggagcc

                                                 550                                                 600
gtgaagagcggagacggccgctggtaatattcagtgattaacaactgtaaaatctaaaatcagtataaatattcatgtgtgttgatttaaaggattttgg

                                                 650                                                 700
tgaaatccggacccgaatcggagctgaggtccgaaactaacgtgcaacatccaataaaagcaagtgcagcttgggttattagagcataaagcatctccac

                                                 750                                                 800
cagagtctcggttgccagccaagcaaaacggccgagacgcgtctttcgggtggccagctgcagtgtcggatgcagtttttggctgcagctgcaattaaat

                                                 850                                                 900
atgcaatgcacttaagttatgcacatttcctttgcccagggaccgttacgatttcgtattttaaaagccgcgcgaaagctaacagcgctgttggctagca

                                                 950                                                1000
gacgtctgtttctgctgttgttagccaacagccaacaacaacaaactggctgttgtttcgaatgcactgagatcttagctctttcccgctaaatgtttc

                                                1050                                                1100
tttgtgtgtcatttttgtcaaccattcgatcgatatcgatattcctttgtgaaggctaaaaataaaattcttacatgctcatcttatacaataatggatcg

                                                1150                                                1200
cacatatagctactccttttcatgtatccaattcactaggtaagctacaaagatagccagacggctgtgctaacagagcgctgctatggtgagaagccg

                                                1250                                                1300
ccgcaatggcaacagtgccatctctggccggcgatacaaatagcgcaacttcgctcgcaacgggtttcATTCCCACTGGAGTGCAGTTCGTGAGCGGTCG
                                                                                  |----->  P3, P5, P7, E74, E76
                                                1350                                                1400
TTCTCTCCTCTCTCTCGCAAAAGTAAAACTTAAAGTTTTGTGCGCGCATGCGTCTCGTCTTAGAAAAACCTGCAATCAGTAATCTCGAATCGTGAAAATC

                                                1450                                                1500
CTTTTCTGCCATACCAAAAAAAGAGCCGGACAAAGCAAAGCGTTCGATTGCCTTCTGCCAATTTGAGCAGGAAATTACAAATTAATTGCCGCCCTGACCG

                                                1550                                                1600
TGATTTGTTTACGTTTTTTTCTTCACTTTTTTTTACATTGTGTGAGTGTGAGAAATTTGTTGGCCAAAAAATATTCCAGATATTCCCAGAGATATATATA

                                                1650                                                1700
TTCGAAATTCGAATTGGCGAAGTGAAGGGTGTGACGACAGACGCACACACACGCACGCACACGAAAACGCACTGAAAAAACACCGGCTAAAAGGCAGACA

                                                1750                                                1800
AAAAGGAGACGAGAGAGCGAGTGCTAAATAGAAAACGCGAAAACAGGATAATAATCAGCAAACTAATACTATAAAACAAGCAACTAAGTGACATCTCGAT

                                                1850                                                1900
CTCGCCGAGAAACAACAAAAACAACAGTTACAACTTCGGCAACAGCAACAAAAACAGCGCCAGCAGCAGCAACAACAACAAAATCAGCGAAGAGCATAGA

                                                1950
GCCAAAAAGAAGGAGCAGCAAAAGCAAAGAAAgtaagtggaattccttgcccaaagatggctctcctctcctttgcctttgctctctcatc
```

Section B

```
                                                  50                                                 100
acttaccatactttacccggaaaactaatcgtcccactctcacatccttcattgcagCACCATAAGTTCGATCTTCCTTACGCAAAGAAGCCGGATTCCGG

                                                 150                                                 200
AGACAGCCTCAAGATTGCTCCCATCAGCACCATGAACGGCCAGGGCTGCGAGCTGGGCCACAGCAATGGCGACATCATCTCACAGAATCAACAGAAGGGA
                                  M  N  G  Q  G  C  E  L  G  H  S  N  G  D  I  I  S  Q  N  Q  Q  K  G

                                 250                                                 300
TGGTGGACCATCGGCCTGATCAACGGCCAGCACAAGTACATGACCGCGGAGACCTTTGGGTTCAAGCTCAACGCCAATGGCGCCAGTCTGAAGAAGAAGC
 W  W  T  I  G  L  I  N  G  Q  H  K  Y  M  T  A  E  T  F  G  F  K  L  N  A  N  G  A  S  L  K  K  K

                                 350                                                 400
AACTGTGGACGCTGGAACCCTCGAACACCGGTGAAAgtgagtagaggttcttaaacagatgggagaaatgttagttagttgggttcttgaaagatttaag
Q  L  W  T  L  E  P  S  N  T  G  E
```

Section C

```
                                                  50                                                  100
gaattcttcgaaaaatgctttccttgcagcaatagctcgccaatgaagcgacttactaatccatggactttaccatatctgtttcagGTATTATCTACTT
                                                                                          S  I  I  Y  L

                                                 150                                                  200
ACGATCTCATCTGAACAAGTACCTTTCGGTCGATCAGTTTGGCAACGTGCTGTGCGAGAGCGATGAGAGGGACGCGGGCAGCCGCTTCCAGATCAGCATC
  R  S  H  L  N  K  Y  L  S  V  D  Q  F  G  N  V  L  C  E  S  D  E  R  D  A  G  S  R  F  Q  I  S  I

                                                 250                                                  300
AGTGAGGACGGCAGCGGACGTTGGGCGCTGAAGAACGAGTCGCGCGGCTACTTTCTGGGCGGCACTCCGGACAAACTGGTCTGCACGGCCAAGACGCCCG
 S  E  D  G  S  G  R  W  A  L  K  N  E  S  R  G  Y  F  L  G  G  T  P  D  K  L  V  C  T  A  K  T  P

                                                 350                                                  400
GTGCCAGTGAGTTTTGGACGGTCCATTTGGCTGCCCGGCCGCAGGTGAATCTGCGCTCCATTGGACGCAAGCGATTCGCCCATTTGTCGGAGTCGCAGGA
G  A  S  E  F  W  T  V  H  L  A  A  R  P  Q  V  N  L  R  S  I  G  R  K  R  F  A  H  L  S  E  S  Q  D

                                                 450                                                  500
CGAGATCCATGTGGACGCCAATATTCCTTGGGGCGAGGATACGCTCTTTACGCTGGAGTTCCGTGCCGAGGAGGGCGGTCGCTATGCTCTGCATACGTGC
  E  I  H  V  D  A  N  I  P  W  G  E  D  T  L  F  T  L  E  F  R  A  E  E  G  G  R  Y  A  L  H  T  C

                                                 550                                                  600
AACAACAAgtgagtggactcattgccgtaactctacaatgcatcaaacttttttttttccccttgcagATATCTGAACGCCAATGGAAAATTGCAGGTG
 N  N  K                                                                    Y  L  N  A  N  G  K  L  Q  V

                                                 650                                                  700
GTGTGCAACGAGGATTGCCTGTTCAGCGCCGAATATCATGGTGGCCATCTGGCGCTGCGCGATCGTCAGGGTCAGTACTTGTCGCCCATTGGCTCCAAGG
 V  C  N  E  D  C  L  F  S  A  E  Y  H  G  G  H  L  A  L  R  D  R  Q  G  Q  Y  L  S  P  I  G  S  K

                                                 750                                                  800
CGGTGTTGAAGTCCCGCTCGTCGTCAGTGACGCGGGATGAGCTCTTCTCGCTGGAGGATTCGCTGCCTCAGGCTTCCTTTATAGCCGGACACAATTTGCG
A  V  L  K  S  R  S  S  S  V  T  R  D  E  L  F  S  L  E  D  S  L  P  Q  A  S  F  I  A  G  L  N  L  R

                                                 850                                                  900
ATATGTGAGCGTTAAGCAGGGCGTCGATGTGACGGCCAACCAGGACGAGGTCGGTGAGAACGAGACGTTCCAGCTGGAGTACGATTGGTCGGCGCACCGT
  Y  V  S  V  K  Q  G  V  D  V  T  A  N  Q  D  E  V  G  E  N  E  T  F  Q  L  E  Y  D  W  S  A  H  R

                                                 950                                                 1000
TGGGCCCTACGCACCACCCAGGATCGCTACTGGTGTCTGTCGGCGGGCGGTGGCATCCAGGCCACCGGCAATCGTCGCTGTGCCGACGCTCTCTTCGAGC
 W  A  L  R  T  T  Q  D  R  Y  W  C  L  S  A  G  G  G  I  Q  A  T  G  N  R  R  C  A  D  A  L  F  E
    <---| P7
                                                1050                                                 1100
TGATCTGGCACGGCGATGGCTCGCTCTCGTTCCGGGCTAACAACGGCAAGTTCTTGGCCACCAAGCGCTCTGGTCATCTGTTTGCCACCTCGGAGTCGAT
L  I  W  H  G  D  G  S  L  S  F  R  A  N  N  G  K  F  L  A  T  K  R  S  G  H  L  F  A  T  S  E  S  I

                                                1150                                                 1200
CGAGGAGATAGCCAAGTTCTATTTCTACTTGATCAATAGgtaagggaattgtttacaccattctgtattttccgtttttaattgcattgtttttgttaaa
  E  E  I  A  K  F  Y  F  Y  L  I  N  R
<| P2
                                                1250                                                 1300
ttaacagACCAATTCTTGTACTGAAGTGCGAGCAGGGATTCGTGGGCTATCGCACGCCCGGTAACCTGAAGCTCGAGTGCAATAAGGCCACCTACGAAAC
        P  I  L  V  L  K  C  E  Q  G  F  V  G  Y  R  T  P  G  N  L  K  L  E  C  N  K  A  T  Y  E  T
          <---| E76
                                                1350                                                 1400
GATTCTGGTGGAGCGGGCCCAGAAGGGCCTGGTGCATCTGAAGGCGCATAGCGGCAAATACTGGCGCATCGAGGGCGAAAGCATCTCGGTGGATGCCGAT
  I  L  V  E  R  A  Q  K  G  L  V  H  L  K  A  H  S  G  K  Y  W  R  I  E  G  E  S  I  S  V  D  A  D

                                                1450                                                 1500
GCGCCCAGCGATGGTTTCTTCCTCGAGCTTCGCGAGCCGACCAGGATCTGCATAAGgttggaccgcatttgctgtcatttaaaaacggctgtaaattatg
 A  P  S  D  G  F  F  L  E  L  R  E  P  T  R  I  C  I  R

                                                1550                                                 1600
attgttaactaatgccgatcctctatattttagATCGCAGCAGGGCAAGTACCTGGGCGCTACGAAGAACGGGGCGTTCAAGCTGCTGGCTGATGGAACC
                                  S  Q  Q  G  K  Y  L  G  A  T  K  N  G  A  F  K  L  L  D  D  G  T

                                                1650                                                 1700
GATTCGGCCACACAGTGGGAGTTCTAGGGCGTGACCACGACCATCCCGAATGGGTTGCTGCACGCGTCATCAGTATCACAATCACACACATACTTCCCAA
 D  S  A  T  Q  W  E  F ***

                                                1750                                                 1800
CCAAGCGGGGATTCCCTAATTATTGATCCACGCAGCTGAAGCGCCAAGAAGAGTCGCACAAGAATTTGATATATCCTAAACAATTGCAAAATGAACGAAA

                                                1850                                                 1900
TTAAAATGATAGGAAGGAAACGTTGATAGACCCAAGGAGATACGAGTGAATGAAAAATGAAGAAACAAACTTATATATTTATATATAAAGATCTACGTAT

                                                1950                                                 2000
 ACATAGAGAGAGATTTTTTTTAAACAACCTATTTAGCTTTTTATACGTTTTATACGATGAGAATATCTACAAACGAGAACAAGAAATCTTTTAGCGTAGA
```

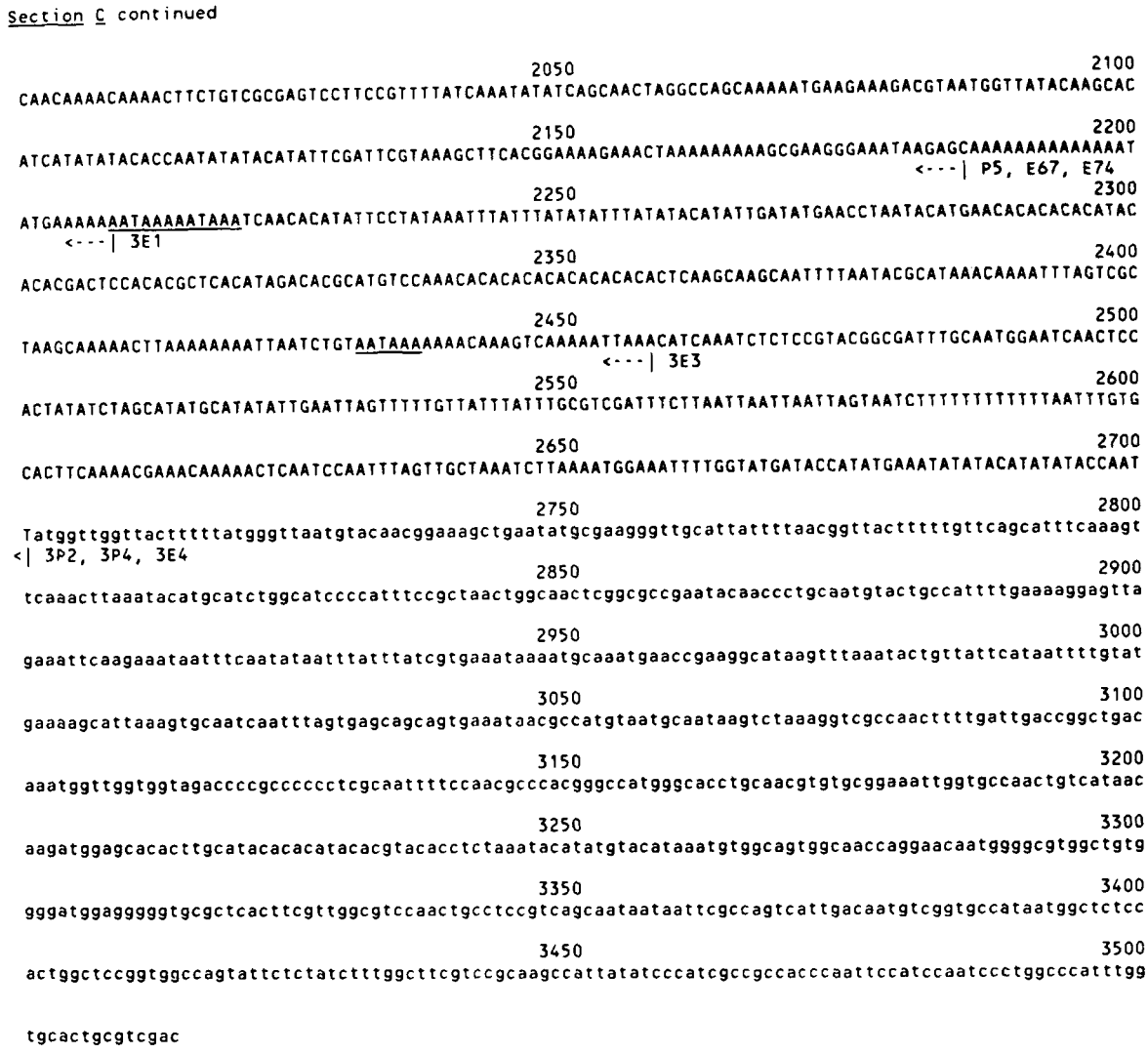

FIGURE 4.—Sequence of the *singed* gene. Exon sequences are in upper case. The positions corresponding to the ends of cDNA clones are shown below the DNA sequence. Pupal cDNAs are prefixed P, embryoic cDNAs are prefixed E. The proposed start and stop codons for the *sn* protein are underlined and the single letter code used to indicate the predicted amino acid sequence of the putative *sn* protein. Potential polyadenylation signals near the 3′ ends of RNAs are underlined. Section A is from the *Sal*I site at −2.0 to 47 bp beyond the *Eco*RI site at 0.0. Section B shows the sequence of the second exon and surrounding DNA, around +10.8. Section C is from the *Eco*RI site at +12.5 to the *Sal*I site +16.

adenylation at +14.3, +14.55 and +14.8, and that the RNAs do not otherwise differ in their intron-exon structures.

DNA sequence of *singed*: Analysis of RNA and cDNA defined the 5′ and 3′ ends of the *sn* transcription unit. An approximate intron-exon structure for *sn* was generated by hybridizing cDNA clones to DNA blots of cloned genomic DNA. The pattern of hybridization was consistent with that expected from the RNA blotting experiments (data not shown) and showed that the *sn* transcription unit covers approximately 15 kb. To precisely define the gene structure, the complete DNA sequence of the P5 cDNA was determined together with the DNA sequences of the corresponding genomic regions, and upstream and downstream sequences. This structural and sequence information is shown in Figures 1 and 4.

Our analysis of RNA and cDNA clones suggests that transcription initiates around position 1269 of Figure 4A. The sequence at this position is a minor variant of the Drosophila transcript initiation consensus sequence (BUSSLINGER *et al.* 1980). The region immediately upstream from this position has no TATA or CAT sequences and the absence of these eukaryotic promoter signals (BREATHNACH and CHAMBON 1981) may partly explain the low abundance of *sn* transcripts. The sequences at the exon-intron boundaries fit with consensus sequences of splice sites (MOUNT 1982). The first exon has a single ATG codon (position 1348 in Figure 4A) which is followed 15 codons later by a stop codon. The 5′ exon is followed by a 10.8-kb intron. We believe that translation starts within the second exon at position 132 of Figure 4B, where there is a good match with

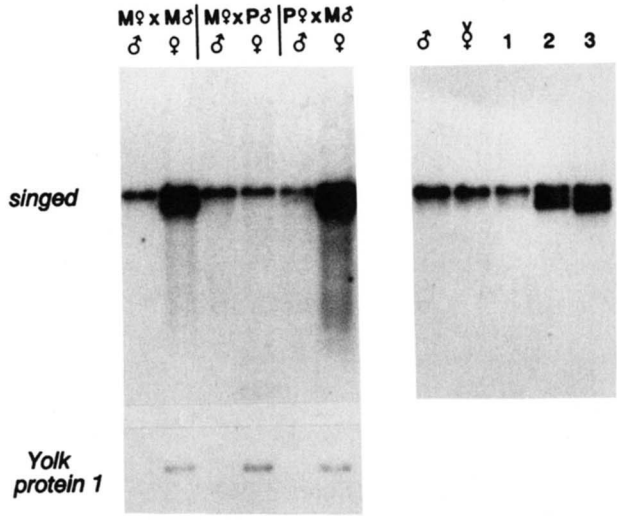

FIGURE 5.—The 3.0- and 3.3-kb RNAs are expressed in ovaries of mature egg-laying females. The left panel shows poly(A$^+$) RNAs from chromosomally identical females with (P ♀ × M ♂) or without (M ♀ × P ♂) ovaries hybridized to a probe from −2.0 to 0.0. The blot was then hybridized with a *yolk protein 1* probe to demonstrate the sex of each sample. The right panel shows poly(A$^+$) RNAs from males (♂), newly eclosed females (♀) and females allowed to mature for 1, 2, or 3 days hybridized to a probe from −2.0 to 0.0.

translation initiation sites (CAVENER 1987). This would be the second AUG in the RNAs, and there is a continuous open reading frame (ORF) from this position through successive exons to a stop codon in the 6th exon (position 1625 of Figure 4C). The 1536 nucleotide long ORF is present in all three transcripts and encodes a protein of 57 kilodaltons. The predicted polypeptide has no unusual sequence features and when compared to computer databases of protein sequence, no clearly homologous entries were found.

The final exon of *sn* RNAs varies in length according to the position of polyadenylation, and the use of the three polyadenylation sites is developmentally regulated. In pupae, only the site at 2701 of Figure 4C is used to produce a single 3.6-kb transcript. However, in adult females two additional sites around 2457 and 2209 of Figure 4C are also used to produce additional 3.3- and 3.0-kb mRNAs.

Adult female 3.0- and 3.3-kb RNAs are from middle to late stage ovaries: The arrest of oogenesis in female sterile *sn* mutants indicates that some or possibly all of the transcripts seen in adult females are from ovaries. To investigate this, flies lacking ovaries were made using the gonadal dysgenesis (GD) sterility which occurs in P-M hybrid dysgenesis. (See legend to Figure 5 and MATERIALS AND METHODS for details.) RNAs from flies without ovaries, from chromosomally identical flies with ovaries, and from the male siblings of these flies are analyzed in Figure 5. In dysgenic females lacking ovaries, the *sn* probe (−2.0 to 0.0) detects only the 3.6-kb RNA. The other samples show wild-type patterns of *sn* transcripts. This shows that the 3.0-kb and 3.3-kb RNA are at least ovary dependent, and probably are made in ovaries. The remaining 3.6-kb RNA present in dysgenic females is presumably from somatic tissue. As a control, the RNA blot was probed for RNA from the *yolk protein 1* gene (Figure 5). This gene is not expressed in males, but is expressed in both somatic and germ cells of adult females (HUNG and WENSINK 1981).

Only the 3.6-kb RNA is detected in late stage pupae (Figure 2). As female pupae have ovaries with previtellogenic egg chambers (stages 1–7; see KING 1970), this implied that the 3.3- and 3.0-kb RNAs might be restricted to later stages of oogenesis and only accumulate as the ovary matures. The time of appearance of the 3.3- and 3.0-kb RNAs was therefore determined. Females were harvested immediately after eclosion, or allowed to age before harvesting. Analysis of *sn* RNA in these females shows that the ovary specific RNAs do accumulate with time and significant amounts are not detectable until 2 days after eclosion (Figure 5).

Maternal inheritance of *singed* transcripts: The presence of all three *sn* transcripts in early embryos (Figure 2) suggested that they were being made in the ovary and maternally inherited by embryos. No role is known for *sn* during embryogenesis. Eggs laid by a mutant sn^- female sterile mother are not rescued by a paternal wild-type sn^+ gene, and sn^- eggs laid by heterozygous sn^+/sn^- mothers hatch and develop normally. We suspected that the presence of *sn* RNAs in early embryos reflected expression of *sn* in nurse cells during oogenesis and transfer of the RNA into the developing oocyte, rather than zygotic expression of *sn* in embryos. Analysis of transcription in sn^3 (see below) suggested a way to follow the maternal inheritance of *sn* RNAs. Although the deletion in this mutant from −0.8 to −0.5 has removed the position where transcription initiates in sn^+, RNAs are made (see below). Sequences upstream of −0.8 are incorporated into sn^3 RNAs, and a probe from −1.3 to −0.7 (probe A used in Figures 6 and 7) detects sn^3 RNAs, even in heterozygous sn^3/sn^+ adults (data not shown). Conversely, a probe from within the interval deleted in sn^3 (probe B in Figures 6 and 7) detects wild-type sn^+ RNAs in sn^3/sn^+ adults.

These two probes were used to follow the inheritance of *sn* RNAs from either sn^3 or sn^+ adult females to embryos. Figure 6 shows RNA from embryos produced by the two reciprocal crosses of sn^3 with sn^+. The embryonic RNA always corresponds to the maternal genotype indicating that in sn^+ (and in sn^3), the RNAs are maternally inherited. The 3.6-kb RNA is barely detectable in the RNA from embryos laid by sn^3 mothers. This suggests a very low level of zygotic expression from the paternal chromosome in embryos (see DISCUSSION), and confirms that the great bulk of the embryo RNA is maternally contributed.

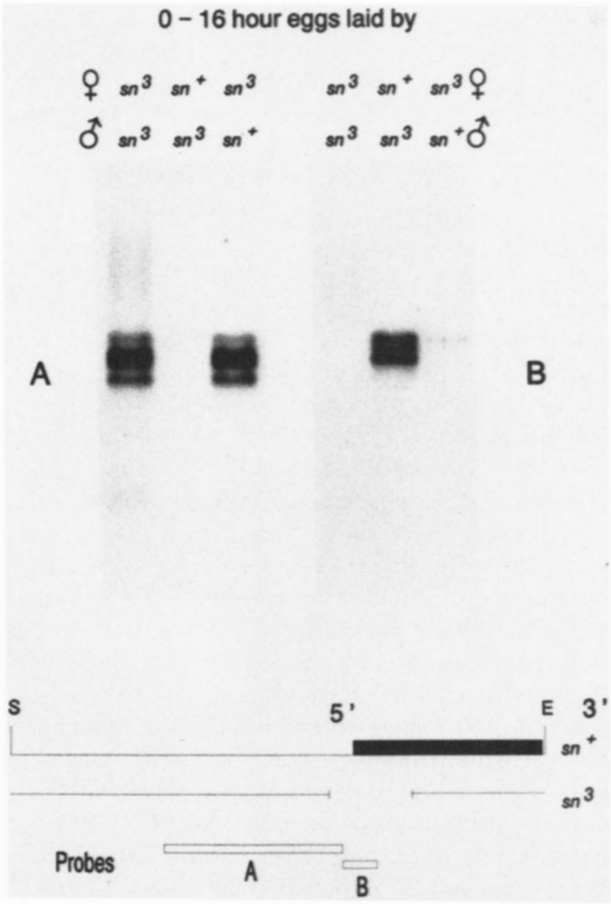

FIGURE 6.—Maternal inheritance of *singed* RNAs. Poly(A$^+$) RNAs were prepared from 0–16-hr embryos from the crosses indicated above each lane. In panel A, probe A from −1.3 to −0.7 only detects RNAs in embryos laid by sn^3 mothers, irrespective of the genotype of the fathers. In panel B the same RNA blot is reprobed with a sn^+ specific probe (probe B), and RNAs are only detected in embryos laid by sn^+ mothers irrespective of the genotype of the fathers.

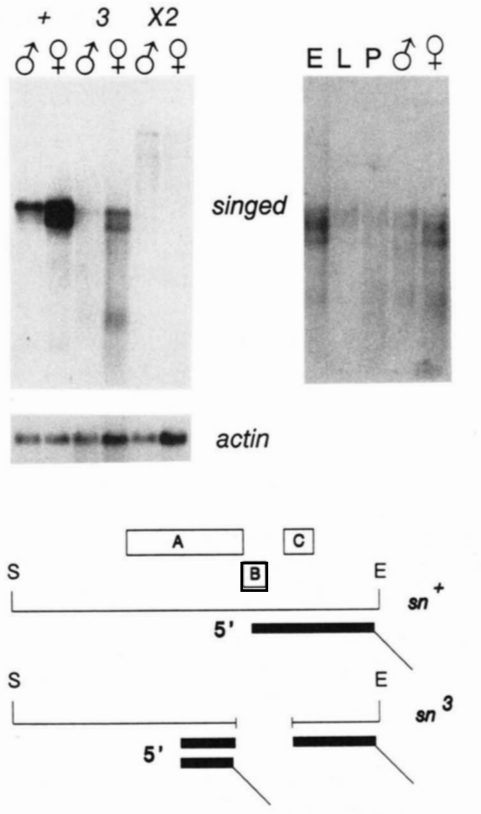

FIGURE 7.—Transcription in *sn* mutants. The left panel shows poly(A$^+$) RNAs from sn^+, sn^3 and sn^{X2} adults hybridized with a probe from −0.5 to −0.3 (probe C). The blot was then hybridized with a 5C *actin* probe to control for loading. The right panel shows the developmental profile of sn^3 RNAs. Poly(A$^+$) RNA from 0–16-hr embryos (E), larvae (L), pupae (P), adult males (♂) and adult females (♀) were hybridized with probe A from −1.3 to −0.7. The lower panel shows the position of probes A and C (and also probe B used in Figure 6) and our interpretation of the 5′ exons in sn^+ and sn^3.

Transcription in *sn* mutants: RNA from sn^3 and sn^{X2} was examined. These mutants have very similar extreme bristle phenotypes, but sn^{X2} is sterile while sn^3 is fertile. The RNAs detected in sn^3 pupae were the same as those detected in sn^3 adult males (data not shown). Compared with wild type, only trace amounts of some large RNAs are found in sn^{X2} males and females (Figure 7). As this mutant has no obvious lesion within the *sn* transcription unit, perhaps there is a point mutation which disrupts processing of the large *sn* precursor RNA.

Some probes do detect RNAs in sn^3, but they are less abundant than sn^+ RNAs (Figure 7). These sn^3 RNAs have been investigated in more detail and our interpretation of their 5′ ends is shown in Figure 7. With a sn^3 specific probe (probe A in Figures 6 and 7 from −1.3 to −0.7), RNAs of 3.5 and 3.0 kb are detected in larvae, pupae and adult males. Only the 3.5 kb sn^3 RNA hybridizes with a probe from −0.5 to −0.3 (probe C). The pattern of hybridization with these specific probes suggests that there is a single transcription start site around −0.9, with the larger RNA being spliced from the normal exon 1 donor site at 0.0 and the smaller RNA being spliced from a potential donor site at −0.8 (position 1141 of Figure 4A). This would make 5′ exons of about 600 and 100 bp for sn^3 compared to the 670 bp in sn^+. If these alternate 5′ exons were spliced to exon 2 and all other splices were as in sn^+, this would generate RNAs of the size detected. The patterns of hybridization seen in RNA from adult females and embryos is more complex, but is consistent with polyadenylation during oogenesis at the same three sites used in sn^+.

The *sn* coding region is not affected by the deletion in sn^3, and it appears to be incorporated intact into sn^3 RNAs. The phenotype of sn^3 suggests that there is *sn* function during oogenesis, but not during bristle development. We believe that in sn^3, the normal *sn* protein is made, but that the amounts produced are insufficient for normal bristle development. The abundance of sn^3 RNAs in adults is lower than that of sn^+ RNA (Figure 7). This may be due to differences

in RNA stability, or because the cryptic promoter revealed by the sn^3 deletion is weaker than the sn^+ promoter. The sn^3 promoter may also have a different developmental program than the sn^+ promoter. RNA is detected in sn^3 larvae but not in sn^+ larvae, and there is no obvious peak of RNA in sn^3 during metamorphosis (Figures 2 and 7). In comparison with sn^+, the transcription start site in sn^3 is about 200 bp closer to the *olfE* promoter. This may lead to the developmental profile of sn^3 transcription more closely resembling the profile of *olfE*, with little variation during development (data not shown; HASAN 1990), than that of *sn* (Figure 2). It is also possible that the sn^3 promoter is not active in the correct cells or at the correct time for the production of normal bristles. The maternal inheritance of sn^3 RNAs (Figure 6) suggest that it is active in the nurse cells during oogenesis. Although there has been no quantitative measure of fertility in sn^3, it is not obviously less fertile than sn^+. Perhaps less *sn* protein is required during oogenesis than during bristle development, and the sn^3 promoter may be sufficiently active in nurse cells, but not in bristle mother cells.

DISCUSSION

Gene structure: We originally used chromosome jumping to clone a 35-kb segment of DNA which included sequences from the *singed* locus (ROIHA, RUBIN and O'HARE 1988). We are confident that the transcription unit described here corresponds to *sn*. Its size, about 15 kb, is big enough for intra-allelic mapping of mutations. It is disrupted in many *sn* mutants. It is transcribed at the times and in the tissues (soma and germline) where *sn* gene products are required. Transcription is altered in the mutants sn^{X2} and sn^3.

Our physical map can be correlated with HEXTER'S (1957) genetic map of *singed* (Figure 1). We have detected DNA lesions in *sn* mutant alleles mapping to the distal and central intra-allelic recombination sites, but not in those which map to the proximal site. This is probably due to the location of the protein coding sequences at the proximal end of the *sn* gene (Figure 1). Point mutations in protein coding regions will often be deleterious while point lesions in noncoding regions are less likely to affect the phenotype. Mutations which map to noncoding regions are therefore more likely to be associated with large scale changes, such as insertions and deletions, and these alterations in gene structure are readily detected by DNA blotting. We suggest that mutations which map to the proximal recombination site (sn^1, sn^5 and sn^{50k}) are point mutations which alter the *sn* protein sequence. This uneven distribution of point and large-scale lesions between structural (coding) and regulatory (noncoding) sequences within a gene was first pointed out by ZACHAR and BINGHAM (1982) for the *white* locus of *D. melanogaster*.

In addition to characterizing the *sn* transcription unit, we have identified the neighboring distal transcription unit. Transcription of this gene initiates less than 1.7 kb upstream of the *sn* RNA initiation site and is away from *sn* to produce a 5.4-kb RNA present at all stages of development (data not shown). It corresponds to the *olfE* gene (HASAN 1990). DNA blotting experiments using cDNA clones for this transcript show that like *sn*, it has a short first exon followed by a large intron (our unpublished results). The profiles of transcription of *sn* and *olfE* during development are very different yet the intergenic gap is quite small. Since each gene has a large first intron relatively close to where transcription initiates, it is possible that some of the sequences involved in controlling transcription are in the nearby introns rather than in the upstream intergenic region. Although there is no requirement for zygotic *sn* expression during embryogenesis we did detect a very low level of zygotic *sn* transcription in embryos (Figure 6). Perhaps zygotic transcription of *olfE* results in the nearby *sn* transcription start site being occasionally used. A more sensitive developmental RNA blot than that in Figure 2 might reveal a low level of sn^+ RNA throughout development. The cryptic transcription start site uncovered by the sn^3 deletion is about 200 bp closer to *olfE* than is the start site in sn^+. This reduced distance may contribute to the generation of sn^3 RNAs during times when the sn^+ promoter is essentially inactive.

Only one *sn* gene product: From the genetics and the phenotypes of different *sn* mutants, we originally hypothesized that different proteins were made during metamorphosis and oogenesis. This could have explained why, for example, the mutation in sn^3 separated the bristle and fertility phenotypes. However, the results of our molecular analysis indicate that while different RNAs are transcribed from *sn* during metamorphosis and oogenesis, they differ only in the length of the 3′ untranslated region and so encode the same protein product. The phenotypes of point mutations of *sn* are entirely consistent with there being only one gene product; weak bristle mutants are fertile and extreme bristle mutants are sterile. Mutants such as sn^3 appear to affect how the gene is expressed rather than the structure of the gene product.

Computer analysis of the *sn* DNA sequence and of the putative *sn* gene product have not given any clues to the function of *sn*. There were no homologous sequences in databases, no unusual sequence features and predictions of the structure of the *sn* protein indicated no unusual structural features. To understand the biological role(s) of *sn* we must examine the

events which occur during metamorphosis and oogenesis, and how they are affected in *sn* mutants.

The role of *sn* during metamorphosis: The development of hairs and bristles in wild-type and *sn* mutants has been studied using the light and electron microscopes (LEES and WADDINGTON 1942; OVERTON 1967; PERRY 1968; REED, MURPHY and FRISTROM 1975; reviewed by POODRY 1980). Bristles such as the macrochaetes are made up of four cells; the trichogen which produces the bristle shaft, the tormogen which produces the socket, a sensory cell and a neurilemma cell. The bristle shaft develops as a cytoplasmic extension of the trichogen. The center of the shaft is filled with longitudinally oriented microtubules and there are bundles of similarly oriented microfibrils at regular intervals just beneath the cell membrane. The shaft is finally covered with cuticle, and the shape of the bristle therefore depends upon the form of this cytoplasmic extension of the trichogen. In *sn* mutants, electron microscope studies of the elongating shaft shows that there are far fewer bundles of microfibrils than in wild type. The shaft (and hence the bristle) in *sn* mutants is short and curved compared to wild type.

The role of *sn* during oogenesis: Oogenesis in wild-type flies has been extensively reviewed (KING 1970; MAHOWALD and KAMBYSELLIS 1980). The ovary is made up of 10–20 ovarioles along which egg chambers move as they develop. When a germline stem cell divides, one cell remains a stem cell while the other goes through four rounds of cell division with incomplete cytokinesis to produce a 16-cell cyst where the cells are linked through intercellular bridges. One of these germ cells develops as the oocyte while the remainder are nurse cells. The nurse cells become polyploid and increase in size until stage 11 when their cytoplasmic contents are transferred into the developing oocyte. In female sterile *sn* mutants, very few late stage egg chambers are produced (BENDER 1960). The eggs produced are flaccid, and shortened with abnormal filaments. The most severely affected mutant, sn^{36a} also showed reduced levels of ploidy in the nurse cells (KING and BURNETT 1957).

The time course for *sn* RNAs from late pupae through to mature egg-laying females (Figures 2 and 6) suggests that *sn* is not transcribed during the early previtellogenic stages of oogenesis (1–7). Transcription is initiated during vitellogenesis (stages 8–10) and the RNAs are present during late stages of oogenesis and in early embryos. The maternal contribution of *sn* RNAs presumably reflects transcription in nurse cell nuclei and transfer of the RNA along with other cytoplasmic components from the nurse cells into the oocyte at stage 11.

A single role for *sn*?: Several processes are common to both bristle development and oogenesis. Sister cells assume different fates (trichogen versus tormogen, nurse cell versus oocyte). However, these fates are not altered in *sn* mutants, so it is unlikely that *sn* has a role in determining the fates of these sister cells. Endoreplication occurs in the trichogen and tormogen during bristle development, and in nurse cells during oogenesis. One mutant, sn^{36a}, has been reported to show a defect in nurse cell endoreplication. This allele is associated with an insertion in the promoter region (ROIHA, RUBIN and O'HARE 1988), and it is possible that this aspect of the mutant phenotype does not reflect a simple absence of *sn* gene product. Furthermore, endoreplication occurs in many other tissues which are not affected in *sn* mutants, so it is unlikely that *sn* is required for endoreplication. Finally, there are asymmetric movements of cytoplasmic contents (up the developing bristle shaft of the trichogen, from nurse cells to the oocyte). Electron microscopic studies of bristle development in *sn* mutants show a reduction in the number and size of bundles of peripheral microfibrils in the elongating shaft. There have been no equivalent studies on the ovaries of *sn* mutants, but a study of the organization of microtubules and microfilaments in the nurse cell/oocyte complex during normal oogenesis and in *sn* mutants could be very informative. In both circumstances when *sn* is required, there is a reorganization and redistribution of cytoplasmic components, suggesting that *sn* may have a role in this process, perhaps in the function of microfilaments.

We thank DAVE HARTLEY for a critical reading of the manuscript, MARY BOWNES for the *yolk protein 1* probe and stock centers for *Drosophila melanogaster* strains. ALAN DRIVER helped with some of the DNA sequencing. J. P. was supported by a Science and Engineering Research Council studentship. K. O. H. was a Medical Research Council (MRC) Senior Fellow and work in the laboratory was supported by an MRC project grant.

LITERATURE CITED

BANKIER, A. T., and B. G. BARRELL, 1983 Shotgun DNA sequencing, pp. 1–34 in *Techniques in Life Sciences. Nucleic Acid Biochemistry*, edited by R. A. FLAVELL. Elsevier Scientific Publishers, New York.

BENDER, H. A., 1960 Studies on the various *singed* alleles in *Drosophila melanogaster*. Genetics **45:** 867–883.

BREATHNACH, R., and P. CHAMBON, 1981 Organisation and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:** 349–385.

BURKE, J. F., 1984 High sensitivity S1 mapping with single-stranded [P^{32}]DNA probes synthesised from bacteriophage M13mp templates. Gene **30:** 63–68.

BUSSLINGER, M., R. PORTMANN, J. C. IRMINGER and M. L. BIRNSTEIL, 1980 Ubiquitous and gene specific regulatory 5′ sequences in sea urchin histone DNA clones coding for histone variants. Nucleic Acids Res. **8:** 459–472.

CAVENER, D. R., 1987 Comparison of the consensus sequence flanking start sites in *Drosophila* and vertebrates. Nucleic Acids Res. **15:** 1353–1361.

CLANCY, C. W., and G. W. BEADLE, 1937 Ovary transplants in *Drosophila melanogaster*. Studies of the characters *singed*, *fused* and *female sterile*. Biol. Bull. **72:** 47–56.

FEINBERG, A. P., and B. VOGELSTEIN, 1984 A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **113:** 266–267.

FYRBERG, E. A., J. W. MAHAFFEY, B. J. BOND and N. DAVIDSON, 1983 Transcripts of six *Drosophila* actin genes accumulate in a stage and tissue specific manner. Cell **33:** 115–123.

GUO, L. H., R. C. A. YANG and R. WU, 1983 An improved strategy for rapid direct sequencing of both strands of long DNA molecules cloned in a plasmid. Nucleic Acids Res. **11:** 5521–5539.

HASAN, G. 1990 Molecular genetics of an olfactory gene of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **87:** 9037–9041.

HEXTER, W. M., 1957 Genetic resolution in Drosophila. Genetics **42:** 376.

HUNG, M. C., and P. C. WENSINK, 1981 The sequence of the *Drosophila melanogaster* gene for yolk protein 1. Nucleic Acids Res. **9:** 640–649.

KING, R. C., 1970 Ovarian development in *Drosophila melanogaster*. Academic Press, New York.

KING, R. C., and R. G. BURNETT, 1957 Oogenesis in adult *Drosophila melanogaster*. V. Mutations which affect nurse cell nuclei. Growth **21:** 263–280.

LEES, A. D., and C. H. WADDINGTON, 1942 The development of the bristles in normal and some mutant types of *Drosophila melanogaster*. Proc. R. Soc. Lond. Ser. B **131:** 87–110.

LEVIS, R. L., P. M. BINGHAM and G. M. RUBIN, 1982 Physical map of the *white* locus of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **79:** 564–568.

LEVIS, R. L., K. O'HARE and G. M. RUBIN, 1984 Effects of transposable element insertions on RNA encoded by the *white* gene of *Drosophila*. Cell **38:** 471–481.

MAHOWALD, A. P., and M. P. KAMBYSELLIS, 1980 Oogenesis, pp. 141–224 in *The Genetics and Biology of Drosophila*, Vol. 2d, edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, New York.

MARIANI, C., V. PIRROTTA and E. MANET, 1985 Isolation and characterisation of the *zeste* locus of *Drosophila*. EMBO J. **4:** 2045–2052.

MOUNT, S. M., 1980 A catalogue of splice sequences. Nucleic Acids Res. **10:** 459–472.

OVERTON, J., 1967 The fine structure of developing bristles in wildtype and mutant *Drosophila melanogaster*. J. Morph. **122:** 367–380.

PERRIMON, N., and M. GANS, 1983 Clonal analysis of tissue specific and recessive female-sterile mutations in *Drosophila melanogaster* using a dominant female sterile mutation Fs(1)K1237. Dev. Biol. **100:** 365–373.

PERRIMON, N., D. MOHLER, L. ENGSTROM and A. P. MAHOWALD, 1986 X-linked female sterile loci in *Drosophila melanogaster*. Genetics **113:** 695–712.

PERRY, M., 1968 Further studies on the development of the eye of *Drosophila melanogaster*. II. The interommatidial bristles. J. Morphol. **124:** 249–261.

PONCZ, M., D. SOLOWIEJCZYK, M. BALLANTINE, E. SCHWARTZ and S. SURREY, 1982 "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain termination method. Proc. Natl. Acad. Sci. USA **79:** 4298–4302.

POODRY, C. A., 1980 Epidermis: Morphology and development, pp. 443–497 in *The Genetics and Biology of Drosophila*, Vol. 2d, edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, New York.

POOLE, S. J., M. K. LAWRENCE, B. DREES and T. KORNBERG, 1985 The *engrailed* locus of *Drosophila*: Structural analysis of an embryonic transcript. Cell **40:** 37–43.

REED, C. T., C. MURPHY and D. FRISTROM, 1975 The ultrastructure of the differentiating pupal leg of *Drosophila melanogaster*. Wilhelm Roux's Arch. **178:** 285–302.

ROIHA, H., G. M. RUBIN and K. O'HARE, 1988 *P* element insertions and rearrangements at the *singed* locus of *Drosophila melanogaster*. Genetics **119:** 75–83.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

VALENCIA, G., 1966 Report of G. VALENCIA. Drosophila Inform. Ser. **41:** 58.

ZACHAR, Z., and P. M. BINGHAM, 1982 Regulation of *white* locus expression: the structure of mutant alleles at the *white* locus of *Drosophila melanogaster*. Cell **30:** 529–541.

Communicating editor: M. J. SIMMONS