Detailed Structure of the Drosophila melanogaster Stellate Genes and Their Transcripts

Kenneth J. Livak

Central Research & Development Department, E. I. du Pont de Nemours & Co., Inc., Wilmington, Delaware 19880-0328

Manuscript received July 3, 1989
Accepted for publication October 14, 1989

ABSTRACT

The X-linked Stellate locus contains two major size classes of a tandemly repeated gene. An example of each class has been sequenced. The steady-state level of Stellate RNA is much higher in XO testis than in XY testis. Sequencing of six cDNA clones derived from XO testis RNA shows that there are two major introns in the Stellate genes. Primer extension and RNase protection analyses show that these introns are spliced much more efficiently in XO than in XY testis. These results also indicate the major transcriptional start site for Stellate RNA. P element transformation results with a marked Stellate gene demonstrate that at least one of the genes sequenced contains a functional promoter, which generates low levels of RNA in XY testis and high levels of RNA in XO testis. This promoter does not contain a TATA element in the −30 region relative to the transcriptional start. Previous results had implicated a specific region of the Y chromosome, designated here as the Su(Ste) locus, in the control of the Stellate genes on the X. Analysis using segmental Y deficiencies shows that the Su(Ste) region suppresses both the high levels and efficient splicing of Stellate RNA.

In Drosophila melanogaster males lacking a Y chromosome, primary spermatocytes contain either needle- or star-shaped proteinaceous crystals. Crystal morphology is determined by the Stellate (Ste) locus, which maps to position 45.7 on the X chromosome (Hardy et al. 1984). This locus has been cloned (Lovett, Kaufman and Mahowald 1980) and studies of its sequence organization and expression have uncovered a number of remarkable features (Lovett 1983; Livak 1984). The Stellate locus contains a tandemly repeated gene whose transcription has only been detected in testes. Formation of crystals appears to be a direct consequence of overexpression of the tandemly repeated Stellate gene. Copy number of the Stellate gene varies among strains, with the highest being approximately 200 copies found in Oregon R. Low copy number corresponds to the Ste⁺ allele which produces needle-shaped crystals, and high copy number corresponds to the Ste allele and formation of star-shaped crystals.

The presence of a specific region of the Y chromosome prevents the appearance of crystals in spermatocytes. This defines a Y-linked locus that will be called Su(Ste) because it suppresses the Stellate phenotype. Su(Ste) maps to a region on the long arm of the Y just proximal to the fertility factor kl-2 (Hardy et al. 1984). Previous results showed that the X-linked Stellate gene is homologous to a moderately repeated sequence that maps to the Su(Ste) region (Livak 1984). This prompted the speculation that the Su(Ste) genes are Stellate homologs that regulate not only their own activity but the activity of the Stellate genes as well. Thus, when the Y-linked Su(Ste) genes are removed, transcription of the X-linked Stellate genes is uncontrolled leading to overproduction of Stellate RNA and the formation of crystals.

In order to investigate regulation of the Stellate genes, it is necessary to define a Stellate transcriptional unit. This paper reports structural studies that map Stellate RNA to the Stellate gene sequence. An unexpected finding is that the Su(Ste) locus affects splicing of Stellate RNA in addition to its effect on Stellate RNA levels. Results from transformation studies are also reported that begin to define the functional limits of the Stellate gene.

MATERIALS AND METHODS

Preparation of DNA and RNA: Drosophila DNA and RNA were prepared as described previously (Livak 1984). Except where noted, XY testis RNA was prepared from ry⁺ males and XO testis RNA was prepared from the males generated by crossing ry⁺ males to C(1) RM, F² Su(ste) w⁻ b/b O females. Plasmid DNA was prepared from 40-ml overnight cultures by the following modification of the alkaline lysis procedure (Birnbom and Doly 1979). Cells were pelleted by centrifugation at 5000 rpm for 10 min, resuspended in 2.5 ml 50 mM glucose; 25 mM Tris-HCl (pH 8.0); 10 mM EDTA; 4 mg/ml lysozyme, and incubated on ice for 15 min. After the addition of 5 ml 0.2 M NaOH; 1% (w/v) sodium dodecyl sulfate (SDS), the sample was mixed by gentle inversions and incubated on ice for 5 min. After the addition of 3.75 ml K-acetate (pH 4.8) (3 M with respect to K⁺; 5 M with respect to acetate), the sample was mixed by

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

Genetics 124: 303–316 (February, 1990)
inversion, incubated on ice for 15 min, and centrifuged at 10,000 rpm for 10 min. The supernatant was transferred to a fresh tube and 11.5 ml isopropanol were added. After a 10-min incubation on ice, precipitated nucleic acid was recovered by centrifugation. The pellet was resuspended in 3 ml 50 mM NaCl in TE (10 mM Tris-HCl, pH 8.0; 1 mM EDTA). Following the addition of 30 μl 10 mg/ml RNase A, the sample was incubated for 30 min. Then, 1.5 ml 100 mM Tris-HCl (pH 7.8); 1 mM EDTA was added and the sample was loaded on a NENORB PRSP column (NEN) that had been activated with 5 ml methanol. The column was washed three times with 5 ml 100 mM Tris-HCl (pH 7.8); 1 mM EDTA, one time with 5 ml H2O, and five times with 1 ml 3% (v/v) ethanol. The plasmid DNA was eluted from the column in five 1-ml fractions of 20% (v/v) ethanol. The fractions were dried under vacuum in a SpeedVac concentrator (Savant) for 2 h and each residue was resuspended in 100 μl TE. Following analysis of 1-μl aliquots on an agarose gel, peak fractions were combined, ethanol precipitated, and resuspended in 50–100 μl TE. For phagemid clones, secretion of packaged single-stranded DNA was induced by infecting 2-ml cultures grown for 4–6 hr with 5–10 × 109 helper M13 phage (Russe1, Kidd and Kelley 1986) and continuing incubation at 37°C with vigorous shaking, overnight. For phagemids or M13 phage, single-stranded DNA was purified as described in the dye sequencing manual from N. E. BioLabs. For subcloning, vector fragments were dephosphorylated by treatment with 4 units calf intestinal alkaline phosphatase (Boehringer) during the last 15 min of restriction enzyme digestion. Both vector and insert fragments were fractionated on 0.8% agarose gels run in TBE (Maniatis, Fritsch and Sambrook 1982), eluted into 7.5 mM NH4-acetate using the IBI Electroeluter, and ethanol precipitated. Following standard hybridization and processing steps were the same as for DNA blots and plaque lifts, pre-hybridization and hybridization were at 65°C in 25% (v/v) formamide (Fluka); 50 mM Tris-HCl (pH 7.5); 1% (w/v) SDS with 10% (v/v) dextran sulfate (Pharmacia) included in the hybridization buffer. Following overnight hybridization, membranes were washed four times at 65°C in 2X SSC; 0.5% (w/v) SDS, 15 min per wash, briefly dried, and exposed to X-ray film (Kodak). RNA samples were treated with 2.2 M formaldehyde/50% (v/v) formamide at 60°C for 5 min and electrophoresed in 1.2% agarose/2.2 M formaldehyde gels (Lehrach et al. 1977) using an electrophoresis buffer containing 20 mM MOPS (Sigma), (pH 7.0); 5 mM Na-acetate; 1 mM EDTA. Electrophoresis conditions were 10 mA constant current for 18–24 hr at 4°C with buffer recirculation. RNA was transferred from gels to GeneScreenPlus membranes (NEN) according to the manufacturer’s directions, which concludes with a 30-min vacuum baking. Hybridization and processing steps were the same as for DNA blots except the prehybridization and hybridization buffers contained 40% (v/v) formamide and the wash solution contained 0.3X SSC. For hybridization to blots and plaque lifts, 32P-labeled SP6 RNA transcripts were synthesized on plasmid templates linearized with EcoRI following a procedure provided by Promega. From 107 to 108 cpm of 32P-labeled RNA probe were used per hybridization.

**Sequencing:** DNA sequencing reactions were performed by the dye deoxy chain termination method (Sanger, Nicklen and Coulsdon 1977) using reverse transcriptase (NEN) or Sequenase (U. S. Biochemical). The XbaI insert of pSX1.3 was transferred to the M13 vector mBW2348 (Barnes and Bevan 1983) in both orientations in order to prepare single-stranded template for sequencing. For pSX83.4, double-stranded plasmid DNA was used directly as template (Zagursky et al. 1985). For the cDNA clones in phagemid vectors, single-stranded templates were generated by infection with helper phage. After initial sequencing using a vector primer, oligonucleotide primers were synthesized to continue sequencing along the same template. Sequencing gels were 6% acrylamide/8 M urea in TBE electrophoresed at 65 watts constant power. For pSX1.3 and pSX83.4, both strands were sequenced; whereas, only one strand was sequenced for the cDNA clones.

**Genomic library:** Approximately 5 μg male Ste+ Drosophila DNA was partially digested by incubation with 0.08 unit Sau3A in a 40 μl reaction at 37°C for 1 hr. After electrophoresis in a 0.5% agarose gel in TBE, DNA fragments in the size range 15–25 kb were electroeluted out of the gel, ethanol precipitated, and resuspended in 10 μl TE. A λ library was constructed with 2 μl of the size-fractionated Drosophila DNA using the λEMBLY/λamiHI vector kit and Gigapack packaging extract from Stratagene. Phage from the library were plated using the E.coli host CES200 (Nader et al., 1985), which is recE recF. The library contains approximately 75,000 clones with Drosophila inserts. For hybridization with a Stellate RNA probe prepared from pSX1.3, plaques were transferred to Colony/PlaqueScreen membrane discs (NEN) following the manufacturer’s protocol. Hybridization to five filters with approximately 2000 plaques per filter identified 83 plaques with homology to the Stellate probe. After plaque purification, DNA was prepared from plate lysates of each positive clone using the Stellate probe. Clones that contain 1150 or 950 bp of Stellate fragments are derived from the X chromosome; whereas, clones containing 800-bp fragments are derived from the Y (Livak 1984). Of the 83 Stellate-positive clones, 12 are from the X chromosome, 58 are from the Y chromosome, 10 clones are of unknown chromosomal origin, and 3 clones were lost during workup.

**cDNA libraries:** Total RNA was isolated from testes dissected from ry+ XO males generated by crossing P(Small, ry)50.2 females to Y; ry+ X ry+. RNA preparations were combined to make a sample of RNA derived from approximately 400 testes. Poly(A)+ RNA was selected by fractionation on 20 mg oligo-dT cellulose using the RNA Isolation Kit from Pharmacia and a 10 μl sample was used as template for the synthesis of double-stranded cDNA following a protocol obtained from Pharmacia that is an adaptation of the RNase H method of Helms et al. (1985). To identify chromosomal origin, CfoI digested were fractionated on agarose gels, blotted, and hybridized with the pSX1.3 RNA probe. Clones that contain 150 or 950 bp of Stellate fragments are derived from the X chromosome; whereas, clones containing 800-bp fragments are derived from the Y (Livak 1984). Of the 83 Stellate-positive clones, 12 are from the X chromosome, 58 are from the Y chromosome, 10 clones are of unknown chromosomal origin, and 3 clones were lost during workup.
extensive digestion with EcoRI, the sample was extracted with 100 μl phenol, passed over a spin G-50 column (5' → 3'); ethanol precipitated, dissolved in 10 μl TE, and stored at 4°C. Libraries, each constructed using 1 μl of the double-stranded cDNA, were made with the Agt10 or XZAP EcoRI CIAP treated vector kits and Gigapack packaging extract (Stratagene). Following the procedure used with the genomic library, the cDNA libraries were screened with Stellate RNA probe prepared from pSX1.3. The EcoRI Stellate fragment of one positive clone from the λgt10 library was subcloned in the phagemid vector pIBI30 (IBI) to create the cDNA plasmid pOG1.24. For 5 positive clones from the genomic library, the cDNA plasmids were subcloned in the phagemid vector pBluescript SK(M13). The EcoRI Stellate fragments from these plasmids were then subcloned in pIBI30 to create the cDNA plasmids pZ18.4, pZ216.4, pZ44.1, pZ613.4, and pZ724.3. Subcloning into pIBI30 was done because it gave a better yield of single-stranded DNA than the pBluescript SK(M13) vector.

**Primer extension:** Primers were labeled by mixing 7 μl [γ-32P]ATP (70 μCi; 6000 Ci/mmol; NEN), 1 μl (50 ng) oligonucleotide, 1 μl 10X kinase buffer [0.35 M Tris-HCl (pH 7.0); 0.05 M MgCl2; 25 mM dithiothreitol], and 1 μl (10 units) T4 polynucleotide kinase (N. E. BioLabs), and incubating at 37°C for 30 min. After the addition of 190 μl 100 mM Tris-HCl (pH 7.8); 1 mM EDTA, the oligonucleotide was purified on a NENSORB 20 column (NEN) following the manufacturer’s protocol. The 20% (v/v) ethanol eluate from the column was dried under vacuum in a Speed-Vac concentrator (Savant) for 2 hr and the residue was resuspended in 50–100 μl 0.1 mM EDTA (pH 7.2), to make a solution containing approximately 106 cpm/μl. Each RNA sample was dissolved in 100 μl 0.1 mM EDTA (pH 7.2), passed over a spin G-50 column equilibrated with 0.1 mM EDTA (pH 7.2), and ethanol precipitated. The RNA pellet was dissolved in 6 μl of the 32P-labeled primer solution, placed in a boiling water bath for 2 min, then placed on ice. The primer extension reaction was performed by adding 2 μl 5X RT buffer (250 mM Tris-HCl (pH 8.3 at 42°C); 40 mM MgCl2; 150 mM KCl; 5 mM dithiothreitol), 1 μl 20 mM each dNTP, and 1 μl (17 units) reverse transcriptase, and incubating at 42°C for 1 hr. After the addition of 2 μl 0.25 mM EDTA (pH 8.0) and 4 μl 1 M NaOH, incubation at 42°C was continued for 15 min. The reaction was stopped by adding 8 μl 1 M Tris-HCl (pH 7.4), 0.4 μl 25 mg/ml yeast tRNA, 25 μl H2O, 50 μl 5 M NH4-acetate, and 250 μl ethanol, then storing at −20°C. The ethanol precipitate was collected by centrifugation, dissolved in 5 μl formamide-dye solution [95% (v/v) formamide; 12.5 mM EDTA; 0.3% (w/v) xylene cyanol; 0.5% (w/v) bromophenol blue], and electrophoresed on a sequencing gel.

**RNase protection:** For preparing RNA probes for RNase protection, subclones were constructed in the vector pSP6-1AT (NEE-155; NEN). This vector has a poly A tract and BglII site for linearization placed downstream from the SP6 promoter and multiple cloning sites. 32P-labeled SP6 transcripts with specific activity >106 cpm/μg were synthesized as described in the NEN instruction manual for its SP6/77 RNA Polymerase Labeling System and purified on oligo-dT cellulose as described in the cDNA section. The poly(A) selection eliminates any contamination with opposite strand transcripts. For each Drosophila RNA sample to be analyzed, DNA was degraded in 100 μl reaction containing 50 mM Tris-HCl (pH 8.3); 8 mM MgCl2; 30 mM KCl; 1 mM dithiothreitol; 25 μg yeast tRNA; 25 units RNasin (Promega); and 1 unit RNase-free DNase (Promega), which was incubated at 37°C for 15 min. After the addition of 5 μl 0.2 mM EDTA (pH 7.2), extraction with 100 μl phenol: chloroform:isomyl alcohol (50:49:1), and ethanol precipitation, the RNA sample was dissolved in 30 μl hybridization buffer [80% (v/v) formamide; 40 mM PIPES (pH 6.8); 0.4 M NaCl; 1 mM EDTA] containing 105 cpm of a 32P-labeled RNA probe. This mixture was placed in a boiling water bath for 1 min, then transferred immediately to 50°C and allowed to hybridize overnight. RNA digestion was accomplished by adding 250 μl RNase solution [10 mM Tris-HCl (pH 7.4); 5 mM EDTA; 0.3 M NaCl; 40 μg/ml RNase A (Sigma); 950 units/ml RNase T1 (BRL); 120 U/ml RNase T2 (BRL)] and incubating at 30°C for 1 hr. After the addition of 20 μl 10% (w/v) SDS and 50 μl proteinase K solution [10 mM Tris-HCl (pH 7.4); 5 mM EDTA; 0.3 M NaCl; 400 μg/ml yeast tRNA; 1 mg/ml proteinase K (Boehringer)], the sample was incubated at 37°C for 15 min, extracted once with 400 μl phenol:chloroform:isomyl alcohol (50:49:1), and extracted once with 400 μl chloroform. The RNA was precipitated by the addition of 1 ml ethanol at −20°C, collected by centrifugation, dissolved in 5 μl formamide-dye solution, and electrophoresed on a sequencing gel.

**Drosophila transformation:** The tandemly repeated

![Probe: pSX1.3](image-url)

**Figure 1.—DNA blot showing plasmid and genomic fragments homologous to a Stellate probe. XbaI digests of 1 ng pSX1.3, 1 ng pSX83.4, 1 μg male Ste+ DNA, 1 μg female Ste- DNA, the DNA from six ry50 females, and the DNA from four ry50 females were electrophoresed, blotted, and hybridized with pSX1.3 RNA probe. Plasmids pSX1.3 and pSX83.4 contain XbaI fragments isolated from Dms2L1 and XST83, respectively, and inserted in the vector pSP64 (Promega). The Ste+ male DNA was used to construct the genomic library from which pSX83.4 was derived. Tissue RNA for the analyses in Figures 3, 4 and 5 was prepared from males carrying the X chromosome of the ry50 stock. The large fragment observed in the pSX1.3 lane is due to partial digestion. The small amounts of the 2.6-, 1.8- and 1.65-kb fragments observed in ry50 female DNA may be there because the DNA was extracted from nonvirgin females which would have stored sperm containing the Y chromosome.
Figure 2.—Sequence of two Stellate genes and six Stellate cDNAs. The pSX1.3 sequence shown is a permutation of the sequence actually in pSX1.3. The actual insert in pSX1.3 starts at the XbaI site at position 272, proceeds to 1269, then continues at position 1, and ends at the XbaI site at 272. The other sequences are shown in reference to the master pSX1.3 sequence. A hyphen means that the base is the same as in the pSX1.3 sequence and an asterisk means that that base is missing. The numbering system refers to the pSX1.3 sequence. The cDNA sequences were derived from the following plasmids: pOG1.24 (cDNA1), pZ613.4 (cDNA2), pZ724.3 (cDNA3), pZ18.4 (cDNA4), pZ44.1 (cDNA5), and pZ216.4 (cDNA6). The GT of splice donors (D) and the AG of splice acceptors (A) are circled. Polyadenylation signals are boxed. The locations of the probable transcriptional start (357–360), the Ste RNA primer (487–483) used in primer extension analyses, and the site (between 617 and 618) of the yeast acetolactate synthase (ALS) insert in the Ste-ALS fusion gene are indicated. The location of key restriction sites used in the creation of subclones are marked. The predicted amino acid sequence of the Stellate protein is shown above the DNA sequence. Amino acid variations predicted by the pSX83.4 or cDNA sequences are shown above amino acids predicted by the pSX1.3 sequence.
structure of the *Stellate* genes was exploited to first insert a BglII site into the protein coding sequence and then to make a permutation of the *Stellate* sequence that would be more likely to be functional in transformation experiments. Using DNA purified from a dam' host, the 1269-bp XbaI insert of pSX1.3 was gel purified then ligated with itself to create long concatemers. Half the ligations should be in a head-to-tail fashion that regenerates a proper *Stellate* repeat. This material was digested with HphI, which cleaves once per *Stellate* repeat in the protein coding region. (DNA from a dam' host was used because the HphI site in the *Stellate* sequence overlaps a dam methylation site.) After removing 3' overhangs by treatment with T4 DNA polymerase, phosphorylated BglII linkers (CAGATCTG) were ligated to the HphI ends. Following digestion with BglII, a fragment of approximately 1300 bp was gel purified and subcloned in a pBR322 derivative with a BglII site to create pBG44. The BglII insert of pBG44 was gel purified, ligated with itself, and digested with DraI. DraI cleaves once per *Stellate* repeat in a region that appears to be an AT-rich spacer. Following ligation with phosphorylated XhoI linkers (CCTCGAGG) and digestion with XhoI, a fragment of approximately 1300 bp was gel purified and subcloned in a pBR322 derivative with a XhoI site to create pXH58.

The plasmid pSA71 was made by inserting a fragment of the yeast acetolactate synthase (ALS) gene into the BglII site of the *Stellate* insert in pXH58 to create a fusion gene designated Ste-ALS. The yeast segment inserted was a 450 bp BglII-BglII protein-coding fragment (base positions 631–1081 in the sequence reported in Falco, Dumas and Livak 1985) for which linkers were used to convert the BglII site to a BglII site. Portions of the *Stellate* insert of pSA71 were sequenced in order to assess the structure of the new junctions formed in the series of plasmid constructions. The 1750-bp Ste-ALS XhoI fragment of pSA71 was gel purified and inserted in the SalI site of the *P* element transformation vector Carnegie-20 (Rubin and Spradling 1985) to create the plasmid pCA80. A mixture of pCA80 and helper plasmid pr25.7wc (Karels and Rubin 1984) was injected into y^506 embryos (Spradling 1986). After mating surviving adults to y^506 flies, three y^+ progeny (from a single injected female) were recovered and used to establish homozygous transformant stocks designated P[(Ste-ALS, ry')30.1], P[(Ste-ALS, ry')30.2], and P[(Ste-ALS, ry')30.3].

**RESULTS**

**Structure of Steellate gene:** Previous results showed that the clone λDm2L1 contains at least eight copies of a tandemly repeated sequence. Analysis using various restriction enzymes indicated that each repeat contains a single XbaI fragment, a single copy of which was subcloned to create plasmid pSX1.3. Figure 1 shows a blot of plasmid and Drosophila genomic DNA hybridized with an RNA probe prepared from pSX1.3. Fragments found in both female and male DNA are derived from the *X* chromosome; whereas, male-specific fragments are derived from the *Y* chromosome (Livak 1984). Figure 1 shows that there are two major types of fragments on the *X* chromosome with approximate sizes of 1250 bp and 1150 bp. RNA probe prepared from pSX1.3 was hybridized to a λ library containing inserts of Drosophila DNA. Restriction analysis of positive clones showed that one clone, AST83, contains a repeated 1150-bp XbaI fragment, a single copy of which was subcloned to create plasmid pSX83.4. Figure 1
shows that the XbaI inserts of pSX1.3 and pSX83.4 correspond to the two major types of Stellate fragments found on the X chromosome. Figure 2 shows the sequences of the inserts of pSX1.3 and pSX83.4, presented as permutations of the sequences of the XbaI fragments. The permuted sequences are presented because this is the arrangement shown to be functional in transformation experiments (see below).

**Structure of Stellate RNA:** Poly(A⁺) testis RNA isolated from XO males was used as template to synthesize double-stranded cDNA. Following the addition of EcoRI linkers, this cDNA was cloned in the vectors λgt10 and λZAP. The cDNA libraries were hybridized with RNA probe made from pSX1.3 and EcoRI inserts from positive clones were transferred to phagemid vectors. The Stellate sequences of six XO testis cDNA clones are presented in Figure 2. Comparison of the cDNA and gene sequences shows that Stellate RNA isolated from XO testis contains two major splices. The sequence of cDNA3 indicates a third splice that does not cause a shift in reading frame, but this splice is not observed in three other cDNA clones. Translation of the cDNA sequences predicts that Stellate RNA codes for the protein shown in Figure 2. The size of this predicted protein is 19,500 daltons, which is in good agreement with the observed size of the putative Stellate protein (Lovett 1983). The sequence of cDNA1 suggests that the sequence AAUCAA is used as a polyadenylation signal. The location and shortness of the A tail of cDNA2 indicates that this cDNA was probably generated by internal oligo dT priming. For the other four cDNAs, polyadenylation occurred shortly after the normal AAUAAA signal.

Figure 3 shows a blot of total testis RNA prepared from XO and XY males, fractionated on a formaldehyde-agarose gel, and hybridized with RNA probe prepared from pSX1.3. As previously observed, XO testis contains an abundant Stellate RNA, approximately 750 bases in size. This size agrees well with the size of the largest cDNA isolated (650 bp) plus a polyA tail. The band of XO Stellate RNA is somewhat diffuse. This probably reflects heterogeneity due to polymorphisms among the Stellate repeats. In XY testis RNA there is an RNA species that appears to be slightly larger than the XO Stellate RNA plus a smear of RNA that extends to very high molecular weight. Furthermore, RNA homologous to the Stellate probe is much less abundant in XY testis than in XO testis. The mixed sample in lane 2 of Figure 3 confirms that the smallest XY Stellate RNA is indeed slightly larger than the XO Stellate RNA and emphasizes the greater abundance of XO Stellate RNA.

Figure 4 shows a primer extension experiment comparing XO and XY Stellate RNA. The primer used is located just 3' to the first intron in the Stellate gene and its position is shown on the sequence in Figure 2. The most prominent primer extension product in the XO testis sample is at 74 nucleotides, indicating that the 5' end of spliced Stellate RNA is at about position 357 marked in Figure 2. The other bands observed in the vicinity of the major band at 74 nucleotides might indicate multiple initiation points or might be due to polymorphisms among the many copies of the Stellate gene. There is a primer extension product at 128 nucleotides which is the size expected for an unspliced Stellate RNA that begins at position 360 shown in Figure 2. In the XY testis sample the most prominent primer extension product is at 128 nucleotides and there is only a minor band at 74 nucleotides. This indicates that the first intron of Stellate RNA is largely unspliced in XY testis.

The structure of Stellate RNA in XO and XY testis was also investigated by the RNase protection experiments shown in Figure 5. The RNA probes used in Figure 5a span the first intron of Stellate RNA and extend beyond the initiation point indicated by the
Structure of the Stellate Genes

Figure 4.—Primer extension analysis of XO and XY Stellate RNA. XO and XY total testis RNA, each prepared from 40 hand-dissected testes, were analyzed by primer extension as described in MATERIALS AND METHODS. The 32P-labeled primer used was the Ste RNA primer whose sequence is AACCAATCGATCCAG and whose location relative to the pSX1.3 sequence is shown in Figure 2. The sequencing ladder shown was prepared by using the same labeled primer in dideoxy reactions on a pSX1.3 template. The sizes indicated are determined from the sequencing ladder. The 128-b fragment should be the most accurately sized because the primer extension product should be identical or nearly identical in sequence to the co-migrating fragment in the sequencing ladder. Because of splicing, the primer extension product migrating at 74 b has a different sequence than the fragment in the sequencing ladder and thus its size may be off by a base or two. A fairly short exposure is shown in order to facilitate the comparison of the primer extension products to the sequencing ladder. A longer exposure of the primer extension gel in Figure 11 shows that XY samples do contain a small amount of the 74-b fragment.

In the XO testis sample major protected fragments are observed at approximately 40 and 69 nucleotides. Based on the initiation points and intron positions shown in Figure 2, the first exon would be expected to protect a 38–41 nucleotide fragment and the 5' portion of the second exon would be expected to protect a 66-nucleotide fragment. The largest protected fragment observed in the XO testis sample is at approximately 161 nucleotides, which corresponds well to the 161–164 nucleotide fragment expected from unspliced Stellate RNA. In the XY testis sample the protected fragment at 161 nucleotides is almost as prominent as the XO tests, but the relative intensities of the 40- and 69-nucleotide protected fragments are much reduced. This again indicates that Stellate RNA is spliced much less efficiently in XY tests than in XO tests. The other bands observed in Figure 5a might be due to polymorphisms among the Stellate genes, protection by homologous transcripts from the Y chromosome, or overdigestion with RNase. 

The RNase experiment shown in Figure 5b examines splicing of the second major intron of the Stellate gene. Based on the position of this intron relative to the RNA probe, the expected protected fragments are at 153 nucleotides for unspliced RNA and at 86 nucleotides for spliced RNA. The results show that XO testis contains predominantly spliced RNA; whereas, in XY tests, the amount of spliced RNA relative to unspliced RNA is much less.
**Definition of Stellate promoter:** A segment of the yeast acetolactate synthase gene was placed in a permuted version of the *Stellate* sequence cloned in pSX1.3 in order to serve as a hybridization marker. Figure 6 shows the sequence inserted. The point of insertion in the *Stellate* gene is in the protein-coding region, at the position shown in Figure 2. The marked *Stellate* gene was inserted in the vector Carnegie-20 and used to transform ry506 embryos. Blots of genomic DNA from three transformant stocks 30.1, 30.2, and 30.31 were used to analyze restriction fragments homologous to a yeast acetolactate synthase probe (data not shown). The results indicate that each transformant stock carries a single copy of the marked *Stellate* gene and that the marked gene is inserted at a different position in each of the three stocks. Inheritance of the *ry*+ marker linked to the marked *Stellate* gene indicates that all three insertions are autosomal. In situ hybridization done by GOEFF JOSLYN indicates that the *P* element insert in 30.2 is at the tip of 2R in region 60E on the salivary gland map.

Figure 7 shows a blot of total testis RNA isolated from *XY* and *XO* males carrying the marked *Stellate* gene of the three transformant stocks. The RNA blot was hybridized with a yeast acetolactate synthase probe so that only expression of the marked *Stellate* gene is being analyzed. The results show that the marked *Stellate* gene produces a 525-nucleotide RNA and that this RNA is expressed to a much greater extent in *XO* testis than in *XY* testis. This demonstrates that the marked *Stellate* gene has a functional promoter and that this gene seems to be controlled in a manner that is similar to the endogenous *Stellate* genes. The size of the chimeric RNA is smaller than is expected for a *Stellate* RNA containing a 500-nucleotide insert. The relatively minor RNAs observed at 1500 and 1350 nucleotides are about the size expected for unspliced and spliced Ste-ALS transcripts. Examination of the sequence in Figure 6 discovers a polyadenylation signal (AATAAA) embedded in the yeast acetolactate synthase gene. Although this signal is not used in yeast (FALCO, DUMAS and LIVAK 1984), its utilization in Drosophila testis would produce an RNA species of about 500 nucleotides, assuming that the RNA start and intron processing is the same as for the endogenous *Stellate* genes.

Total testis RNA isolated from *XY* and *XO* males carrying the marked *Stellate* gene of the 30.2 transformant stock was analyzed in the primer extension experiment shown in Figure 8. The primer used hybridizes to the yeast portion of the chimeric transcript. Comparison to the DNA sequencing ladder shows that the *XO* sample contains a primer extension product that corresponds exactly to the most prominent primer extension product observed in the *XO* sample in Figure 4. This indicates that the RNA start and the processing of the first *Stellate* intron is the same for the marked *Stellate* gene and the endogenous *Stellate* genes, at least in *XO* testis.

**Effect of Y deletions on expression of Stellate RNA:** Figure 9 shows the location of *T(X,Y)* breakpoints in the vicinity of the *Su(Ste)* region on the long arm of the *Y* chromosome. By making the appropriate crosses, the *T(X,Y)* stocks can be used to generate segmental *Y* deficiencies. Figure 10 presents a blot of total testis RNA isolated from various *Y*-deficient
Structure of the *Stellate* Genes

![Figure 7](image)

**Figure 7.**-Blot of XO and XY testis RNA showing the transcripts from the Ste-ALS fusion gene in three transformants. Each of the XY and XO total RNA samples were prepared from 40 hand-dissected testes. For the transformants 30.1, 30.2 and 30.31, XO males were generated by crossing transformant males to C(1)RM, y w su(w) w/bb/O females. Thus, the XO samples are from males hemizygous for the P element insert; whereas, the XY samples are from males homozygous for the P element insert. The RNA samples were electrophoresed, blotted, and hybridized with pALS84 RNA probe. Plasmid pALS84 contains an ALS-specific, 452-bp BgII fragment (nucleotides 3-454 in Figure 6) inserted in the vector pSP64 (Promega). The sizes indicated are based on comparison to the 0.24-9.5-kb RNA ladder (BRL) electrophoresed in a parallel lane and stained with ethidium bromide.

![Figure 8](image)

**Figure 8.**—Primer extension analysis of Ste-ALS RNA. Total RNA samples were prepared from 40 hand-dissected testes from P(Ste-ALS, y+)30.2 XY and XO males. The XO males were generated as in Figure 7. The RNA samples were analyzed by primer extension as described in MATERIALS AND METHODS. The 32P-labeled primer used was the ALS RNA primer whose sequence is GA-CATTTCCATTTCTAC and whose location is shown in Figure 6. The sequencing ladder shown was prepared by using the same labeled primer in dideoxy reactions on a pSAX2 template. Phagemid pSAX2, created by inserting a XbaI-XhoI fragment from pSA71 in the vector pIB130, contains the following sequences: nucleotides 272-617 in Figure 2, then nucleotides 1-509 in Figure 6, followed by nucleotides 618-1269 in Figure 2. The size indicated is determined from the sequencing ladder. Note that the major band in the XO samples co-migrates with the second A in a string of 5 A's in the sequencing ladder just as the major band in the XO sample of Figure 4 co-migrates with the second A in a string of 5 A's in the sequencing ladder.

males and hybridized with a *Stellate* probe. The results show that high levels of the 750-nucleotide *Stellate* RNA are observed only when the *Su(Ste)* region of the Y chromosome is deleted. This demonstrates that correlation between the appearance of crystals and high levels of *Stellate* RNA. An unexpected finding is that the high molecular weight smear of RNA observed in XY tests (see Figure 3) is detected only when the portion of the Y chromosome proximal to the *E15* breakpoint is present. This suggests that there is a gene (or genes) located on the Y chromosome proximal to the *E15* breakpoint that produces RNA homologous to the *Stellate* probe.

Figure 11 shows total testis RNA isolated from Y-deficient males analyzed in a primer extension experiment using the same primer as in Figure 4. The 74-nucleotide product is diagnostic for spliced *Stellate* RNA and the 128-nucleotide product is diagnostic for unspliced *Stellate* RNA. This analysis is complicated by the fact that the *T(X,Y)* stocks differ in the *Stellate* allele carried on their *X*<sup>0</sup>*<sup>y</sup>* elements. The *T(X,Y)W27*, *P7*, and *E15* stocks carry the Ste<sup>+</sup> allele (low copy number); whereas, *T(X,Y)E1* carries the Ste<sup>+</sup> allele (high copy number). Figure 11 shows that *T(X,Y)E1* males have a higher amount of spliced *Stellate* RNA than observed in *W27*, *P7*, or *E15* males. First, deficiencies constructed using just the *W27*, *P7*, and *E15* stocks will be considered. The results indicate that high levels of spliced *Stellate* RNA are found in ΔW27-E15 and ΔP7-E15 males in which all or most of the *Su(Ste)* region is deleted. This suggests that the *Su(Ste)* locus not only suppresses the level of *Stellate* RNA, but also suppresses splicing of *Stellate* RNA. For deficiencies carrying the *E1 X*<sup>0</sup>*<sup>y</sup>* element (ΔP7-E1 and ΔW27-E1), the amount of spliced *Stellate* RNA must be compared to the elevated levels observed in E1 males. The amount of spliced *Stellate* RNA in ΔP7-E1 males does not seem to be much higher than in E1 males. The larger deficiency ΔW27-E1 seems to have somewhat more spliced *Stellate* RNA, but still not as much as in the ΔW27-E15 males. The ΔE1-E15 males, which carry the Ste<sup>+</sup> allele, have a very low amount of spliced *Stellate* RNA. Thus, partial deletions that do not remove most of the *Su(Ste)* region do not seem to have much effect on either the level or splicing of *Stellate* RNA.

**DISCUSSION**

The sequence data, RNA analyses, and transformation results presented here give a detailed picture of the structure of the *Stellate* gene and its transcripts. The gene sequences show that the two major classes of *Stellate* repeats differ mainly by a 150-bp deletion in the 3′ end of the gene that does not affect the
coding sequence or the promoter region. The danger in obtaining sequence data for a repeated gene like *Stellate* is that the particular copy sequenced may not be functional. This illustrates the importance of being able to perform *P* element transformation. The transformation results with a marked *Stellate* gene demonstrate that the pSX1.3 sequence, when arranged as shown in Figure 2, contains a functional promoter and generates low levels of RNA in XY testis and high levels of RNA in XO testis. In XO testis, it appears that the transformant gene uses the same major transcriptional start site as the endogenous genes.

The analysis of *Stellate* RNA is complicated by the fact that there are *Stellate*-homologous sequences on the *Y* chromosome. For example, the possibility that some of the primer extension and RNase protection products observed in Figures 4 and 5 are derived from the *Y* chromosome must be considered. The screen that was used to identify AST83 from the *X* chromosome also identified at least 58 *Stellate*-homologous clones derived from the *Y* chromosome. A complete analysis of these *Y* clones will be published at a later date, but a few results are relevant to the present discussion of *Stellate* RNA. First, the *Stellate* probe is homologous to a repeated sequence on the *Y* chromosome, but the repeat length of the *Y* sequence is 2.6–2.8 kb rather than the 1.15–1.25 kb observed on the *X* chromosome. The *Y* repeat is a combination of segments that are very homologous to *Stellate* plus segments that are totally unrelated. The presence of these unrelated segments means that the transcription and protein products of the *Y* repeats could be very different from those produced by the *Stellate* genes on the *X*. Second, the genomic sequence of two *Y* repeats has been determined. Figure 12 shows portions of these *Y* sequences in the vicinity homologous to the transcriptional start site in the *Stellate* sequence. Assuming that these two sequences are representative of the bulk of the *Y* repeats, the sequence differences would have an effect on both primer extension and RNase protection experiments. The numerous single nucleotide differences between the *Y* sequences and the *Stellate* sequence would probably affect the size of RNase protection products. More strikingly, the *Y* sequences contain a variable number of AAC repeats, which would definitely affect the size of both primer extension and RNase protection products. The single nucleotide difference where the *Ste* RNA primer hybridizes could decrease the efficiency of primer extension on *Y* transcripts. Finally, in both *Y* sequences, the AG of the splice acceptor in the *Stellate* sequence is not present, indicating that *Y* transcripts will have a different splicing pattern than *Stellate* transcripts. Of course, any band observed in XO testis must be derived from the *X* chromosome. For the analysis of Figures 4 and 5, it is assumed that if the same band is observed in XY testis, it is also derived from the *X* chromosome. The sequence differences shown in Figure 12 provide experimental support for this assumption.

The cDNA, primer extension, and RNase protection data indicate a major transcriptional start at position 357–360 of the *Stellate* sequence in Figure 2. The longest cDNAs stop 14 nucleotides short of position 360, which is consistent with a study showing that 8–21 nucleotides are lost from the 5′ terminus.
AUG 33

FIGURE 11.—Primer extension analysis showing the effect of Y deletions on Stellate RNA. Total RNA samples, each prepared from 40 hand-dissected testes, were subjected to primer extension analysis as in Figure 4 using the Ste RNA primer. The W27, P7, E15, and E1 samples were prepared from stock T(X,Y) males (which carry a complete Y complement). 

When cDNA is prepared by the RNase H method (D’Alessio and Gerard 1988). The numerous other bands observed in the RNase protection and primer extension analysis are most likely due to polymorphisms among the repeated Stellate genes or overdigestion with RNase, but might indicate some minor transcriptional start points. Examination of the sequence upstream of the transcriptional start shows that there is no TATA element associated with the Stellate promoter. Thus, Stellate is one of the growing number of eukaryotic genes that is expressed, sometimes at high levels, in the absence of an apparent TATA element. In the -30 region where a TATA is usually found, the Stellate sequence does have an oligopurine tract that is reminiscent of the GAGA element found in the engrailed (Soeller, Poole and Kornberg 1988) and Ultrabithorax (Biggin and Cremers 1988) promoters, but the homology is not compelling in the absence of factor binding data.

The cDNA sequences definitively show that the Stellate gene has two introns. The unexpected finding is that splicing of these introns occurs much more efficiently in XO than in XY testes. The primer extension and RNase protection data do not agree on the amount of splicing that is occurring in XY testes. For the RNase protection data in Figure 5, the ratio of major spliced bands to unspliced bands would suggest that 50% or more of the Stellate transcripts are spliced in XY testes. This number is certainly an underestimate because hybrids involving unspliced RNA are longer and thus are more of a target for internal cleavage due to polymorphisms or helix breathing. The primer extension data for W27, P7, and E15 in Figure 11 would suggest that 10% or less of the Stellate transcripts are spliced in XY testes. Although there is no obvious bias, primer extension results do not necessarily accurately reflect the extent of splicing. If, for example, the secondary structures of spliced and unspliced transcripts differ such that the primer binding site is more accessible in unspliced RNA, then the primer extension data could underestimate the degree of splicing. The fact that very little, if any, 750 b Stellate RNA is observed in XY testes (Figure 3) would tend to support the notion that the amount of splicing is fairly low. This observation, though, might also reflect differential use of polyadenylation signals (see below). Thus, it is clear that Stellate transcripts are spliced less efficiently in XY than in XO testes, but the data cannot be reliably used to quantitate the magnitude of this effect.

Attempts to use RNase protection to assess the use of different polyadenylation signals in XO and XY testes were not successful because too many bands were generated. The multitude of bands were probably due to the heterogeneity and AT-rich nature of the 3′ end of the Stellate genes. For the XO cDNAs it appears that the first polyadenylation signal encountered determined the location of the poly(A) tail. The interesting observation is in the polyadenylation of the Ste-ALS transcripts in the transformants. Figure 7 shows that the relative use of the cryptic yeast polyadenylation signal is much greater in XO than in XY testes. As with splicing, this indicates that the presence of the Y chromosome decreases the efficiency of processing for the Stellate transcripts. The apparent use of AAUCAA as a polyadenylation signal in cDNA1 and the inadvertent splice made in cDNA3 may be indications of superefficient processing in XO testes.

Previous studies correlated deletions of the Su(Ste) region on the Y with the appearance of crystals in primary spermatocytes (Hardy et al. 1984). Figures 10 and 11 confirm that deletions of the Su(Ste) region are also correlated with high levels of Stellate RNA. In addition, Figure 11 shows that the effect of the Su(Ste) region on splicing parallels the effect on RNA abundance. Thus, deletion of the entire Su(Ste) region (AW27-E15) shows the highest level of Stellate RNA and the highest extent of splicing. Partial deletions of the Su(Ste) region, such as AP7-E15, show an intermediate level of spliced Stellate RNA. The analysis of partial deletions of the Su(Ste) region is complicated by the fact that the E1 X0Y+ element (present in the E1, AW27-E1, and AP7-E1 flies) carries the Ste allele...
which corresponds to high copy number of the Stellate gene. Thus, E1 flies contain the entire Su(Ste) region but still show an elevated level of spliced Stellate RNA. This illustrates the competitive relationship that exists between the Stellate genes on the X and the Su(Ste) genes on the Y where the same effect can be observed by increasing copy number on the X or decreasing copy number on the Y.

Previously it had been postulated that the Su(Ste) genes somehow regulate not only their own activity but the activity of Stellate genes on the X as well (Livak 1984). The fact that deletion of the Su(Ste) region simultaneously affects RNA levels, splicing, and possibly polyadenylation suggests an alternative model. Perhaps there is a specific site or type of complex responsible for the synthesis and processing of Su(Ste) transcripts. This could be thought of as a transcription organelle analogous to the nucleolus where rRNA is synthesized and processed. The Su(Ste) complex would have a high concentration of RNA polymerase II, the proper transcription factors, spliceosomes, and other processing factors. It is proposed that the Su(Ste) and Stellate genes compete for a limited number of sites at the complex with preference given to the Su(Ste) genes. In normal XY testis with low Stellate copy number, the Stellate genes are essentially excluded from the complex and only background transcription and processing are observed. In the absence of Su(Ste) genes, the sites of the complex are fully occupied by the Stellate genes resulting in high levels of transcription and processing.

Another model suggested by the splicing results is that there is a sequence in one or both Stellate introns that makes the RNA unstable. Furthermore, it can be postulated that the Su(Ste) and Stellate transcripts compete for some critical splicing factor with preference given to Su(Ste) RNA. In XY testis, very little spliced Stellate RNA accumulates because the splicing factors are occupied with the processing of Su(Ste) transcripts. A steady state level of unspliced Stellate RNA is attained, determined by the rates of transcription and degradation. In XO testis, appreciable splicing occurs so that a large amount of stable, spliced Stellate RNA can accumulate. There would still be a steady state level of unspliced Stellate RNA in XO testis that now depends also on the rate of processing. This model can be tested by making constructs with and without the Stellate introns and analyzing with transformation experiments.

Searching the DNA databases looking for protein similarities (Henikoff and Wallace 1988), it was found that the predicted Stellate protein is homologous to the β subunit of casein kinase II of Drosophila and cow. These homologies are shown in Figure 13. Casein kinase II has been implicated in the phosphorylation of DNA topoisomerase II in Drosophila Kc cells (Ackerman, Glover and Osheroff 1988), but it is the α subunit that possesses the catalytic activity. The function of the β subunit is unknown although a regulatory role has been suggested (Takio et al. 1987). The fact that the Stellate sequence is as similar to the bovine casein kinase sequence as it is to the Drosophila sequence indicates that the divergence of the Stellate and casein kinase sequences was probably not a recent evolutionary event.

In the protein coding region of the two Stellate gene sequences and the six cDNA sequences, there are 19 single nucleotide differences predicting 14 amino acid differences. No frame shift or nonsense mutations have been detected. The absence of the Stellate se-
quency in related species (Livak 1984) raised the possibility that the Stellate gene may serve no function. The homology to the highly conserved casein kinase II β-subunit and the maintenance of an open reading frame in eight Stellate sequences might be taken as indications that the Stellate protein is functional. Yet, it is the spliced Stellate RNA that is translated to give the Stellate protein. In normal XY testis, the level of Stellate protein may vary, but it is low in Ste+ (low copy number) flies and much higher in Ste (high copy number) flies. Thus, it would appear that a wide range of concentration of Stellate protein is tolerated without having a discernible effect on spermatogenesis.

Perhaps, the Stellate genes are derivatives of a functional gene and arose so recently in evolution (Livak 1984) that there has not been time to collect gene-disrupting mutations. Also, unequal crossing-over or some other mechanism may be operating to keep the tandem repeats homogeneous and with an intact open reading frame.

The RNA blot in Figure 10 shows that there is RNA homologous to the Stellate probe that is found in testis only when Y material proximal to the E15 breakpoint is present. This indication that there may be another Y-linked gene homologous to Stellate would seem to contradict earlier evidence that all Stellate homology on the Y chromosome is localized to between the E15 and W27 breakpoints (Livak 1984). The earlier result, though, is based on DNA blots hybridized with a DNA probe; whereas, the present result is an RNA blot hybridized with an RNA probe. Because of the greater stability of RNA-RNA hybrids, the experiment in Figure 10 may be showing a homology that is too weak to be detected in the previous experiments. Thus, in addition to the genes in the Su(Ste) region, the Stellate sequence may be able to detect another testis-specific gene on the Y chromosome that is proximal to the E15 breakpoint.

I am grateful to Terri McGugan for providing technical assistance; Allan Spradling for providing facilities and training for the injection of embryos; Andy Coceburn for providing the Ste+ male and female DNA; Groff Joslyn for the in situ results with the 30.2 transformant; Steve Henikoff for performing the protein homology search; and Rick Scott and Nancy Shepherd for carefully reading the manuscript. The Stellate sequence of pSX1.3 will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number X15899. This contribution number 5196 from the Central Research & Development Department.

LITERATURE CITED


BIRNBAUM, H. C., and J. DOLY, 1979 A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513–1523.


GATTI, M., and S. PIMPINELLI, 1983 Cytological and genetic analy-


Communicating editor: T. SCHÜPBACH