Proc. Natl. Acad. Sci. USA
Vol. 77, No. 10, pp. 5794–5798, October 1980
Biochemistry

Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene

(molecular cloning/polymorphism/intervening sequence)

DAVID A. GOLDBERG

Division of Biology, California Institute of Technology, Pasadena, California 91125

Communicated by Norman Davidson, July 10, 1980

ABSTRACT **The alcohol dehydrogenase (ADH; alcohol: NAD^+ oxidoreductase, EC 1.1.1.1) gene (*Adh*) of *Drosophila melanogaster* was isolated by utilizing a mutant strain in which the *Adh* locus is deleted. Adult RNA from wild-type flies was enriched in ADH sequences by gel electrophoresis and then used to prepare labeled cDNA for screening a bacteriophage λ library of genomic *Drosophila* DNA. Of the clones that hybridized in the initial screen, one clone was identified that hybridized with labeled cDNA prepared from a wild-type *Drosophila* strain but did not hybridize with cDNA prepared from an *Adh* deletion strain. This clone was shown to contain ADH structural gene sequences by three criteria: *in situ* hybridization, *in vitro* translation of mRNA selected by hybridization to the cloned DNA, and comparison of the ADH protein sequence with a nucleotide sequence derived from the cloned DNA. Comparison of the restriction site maps from clones of three different wild-type *Drosophila* strains revealed the presence of a 200-nucleotide sequence in one strain that was absent from the other two strains. The ADH mRNA sequences were located within the cloned DNA by hybridization mapping experiments. Two intervening sequences were identified within *Adh* by S1 nuclease mapping experiments.**

The *Drosophila* alcohol dehydrogenase (ADH; alcohol:NAD^+ oxidoreductase, EC 1.1.1.1) locus (*Adh*) provides an interesting system for studying the molecular basis of differential gene expression during development because it is a developmentally regulated gene for which considerable genetic and biochemical data are available.

ADH activity varies during development and is restricted to certain tissues (1). Activity is first detected during the first larval instar, increases throughout larval development, and gradually declines during the pupal stage. It increases again just prior to adult emergence and peaks 5–10 days later. ADH activity is limited to the fat body, intestines, Malpighian tubules of larvae and adults and to the accessory genital structures of adults. The enzyme cannot be detected in the imaginal discs, larval brain, salivary glands, ovaries, or testes. Thus, ADH expression is developmentally regulated and those elements responsible for its tissue- and stage-specific expression may be accessible to genetic and biochemical analysis.

ADH is one of the best biochemically characterized enzymes in *Drosophila*. It is an abundant protein comprising approximately 1% of the total protein of the adult fly (2). The complete amino acid sequence for the wild-type protein has been determined, as well as the amino acid replacements in a number of genetically variant proteins (3).

Adh is among the best characterized genetic loci in *Drosophila*. The *Adh* locus has been mapped to within two bands (35B 2-3) on the second chromosome (4, 5). Flanking loci with visible and lethal alleles have been identified and their cytogenetic positions have been determined by using deletion mapping of chemically induced mutations (4–6). ADH protein from various strains may be distinguished by electrophoretic mobilities of the enzyme on starch gels. *cis*- and *trans*-acting elements affecting *Adh* expression have been identified (7, 8). The isolation of flies that express or lack ADH activity is possible because both positive and negative selections have been devised (9, 10). These selections are useful both in the isolation of *Adh* null mutants (10) and for the isolation of rare wild-type recombinants generated during fine structure genetic analysis (11).

This paper reports the molecular cloning and initial characterization of the *Drosophila* gene for ADH. The region surrounding this gene is present as a single copy within the *Drosophila* genome and the only adult transcript that is observed to originate from the region is the ADH mRNA. The presence of intervening sequences within the mRNA coding region of the *Adh* gene is demonstrated.

MATERIALS AND METHODS

Strains and Materials. The following bacterial strains were used: K802 for propagation of bacteriophage clones (12); HB101 for propagation of bacterial plasmids (13); JM101 for propagation of M13 clones (14). The following bacterial vectors were used: Charon 9 and Charon 10 bacteriophage (12); pBR322 (15); M13mp2 (14); pSV2 [constructed by placing a central *Eco*RI fragment from Charon 9 into pBR322 and subsequently removing the internal *Sac* I fragments (unpublished data)]. The *Drosophila* bacteriophage λ library was constructed by J. Lauer (see ref. 16 for a description of this library).

Adh^S and w;Adh^F *Drosophila* strains were obtained from W. Sofer. Strains carrying $Adh^{fn2}pr$ and Sco^{R+4} chromosomes were obtained from R. Woodruff (4).

Restriction endonucleases and T4 DNA ligase were purchased from commercial sources and used as directed. Avian myeloblastosis reverse transcriptase was obtained from J. Beard (Office of Program Resources and Logistics, Viral Cancer Program, National Institutes of Health).

Construction and Propagation of M13 Cloned DNA. mAC1-3 was constructed by transferring the *Drosophila* 4.75-kilobase (kb) *Eco*RI fragment of sAC1 to the M13 vector mp2 (see Fig. 3 for a description of *Adh* clones mentioned in this paper). Three independent clones with the *Drosophila* insert in the desired orientation were plaque-purified, and replicative form DNA isolated from 10-ml cultures was analyzed. Two clones deleted the insert with a high frequency; the third clone maintained the insert largely intact as shown by restriction analysis. This phage clone was grown directly in large scale. Single-stranded phage DNA was prepared from virions and fractionated on a 0.7% agarose gel. The intact clone DNA was electroeluted and used in S1 nuclease mapping experi-

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

Abbreviations: nt, nucleotide(s); kb, kilobase(s); *Adh*, alcohol dehydrogenase genetic locus; ADH, alcohol dehydrogenase.

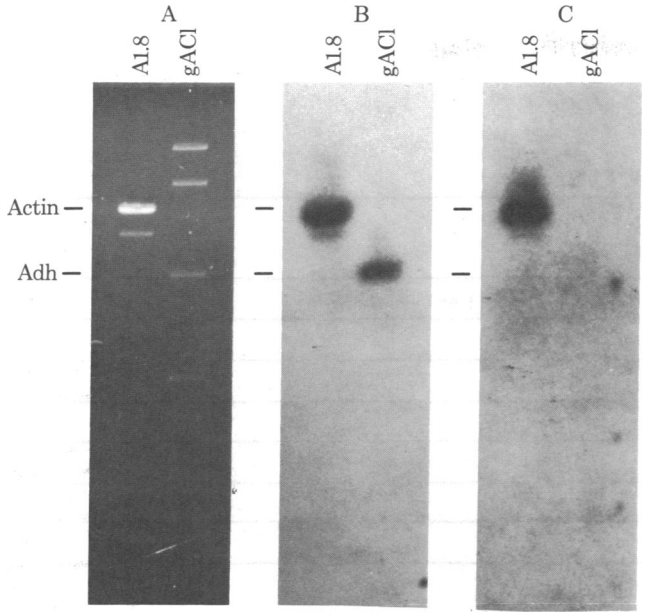

FIG. 1. Hybridization of wild-type and deficiency cDNAs to cloned *Drosophila* DNA. Duplicate samples of DNA prepared from gAC1 and a plasmid clone containing a *Drosophila* actin gene [A1.8, obtained from E. Fyrberg (26)] were digested with *Eco*RI, fractionated by agarose gel electrophoresis, and transferred to nitrocellulose (18). The pattern of the ethidium stained gel is shown in *A*. Separate blots were then hybridized with ^{32}P-labeled cDNA (23) prepared from total adult RNA from wild-type (Urbana S) flies (*B*) or total RNA obtained from flies genetically deleted for *Adh*[*Adh*^{fn2}/*Sco*$^{R+4}$ (4)] (*C*).

ments. Supercoiled replicative form DNA prepared from the large-scale culture was analyzed by restriction endonuclease digestion, and no deletion products could be detected.

S1 Nuclease Mapping Experiments. These experiments (17) were performed under different conditions depending on whether mAC1-3 or sAS1 DNA was used.

mAC1-3. mAC1-3 single-stranded phage DNA (100 ng) was hybridized with 2–10 μg of total adult *Drosophila* RNA for 1 hr at 37°C in 20 μl of 50% formamide/0.4 M NaCl/40 mM 1,4-piperazinediethanesulfonic acid, pH 6.8/1 mM EDTA. The reaction was diluted into 200 μl of S1 nuclease buffer (300 mM NaCl/30 mM NaOAc, pH 4.5/3 mM $ZnCl_2$) and digested with 2000–5000 units of S1 nuclease (Miles) at room temperature or 37°C for 1 hr. After digestion, the mixture was ethanol precipitated and run in two dimensions with a 2.0% or 2.5% agarose nondenaturing dimension and a 2.5% agarose alkaline dimension (17). The gel was transferred to a nitrocellulose filter (18) and probed with nick-translated sAS1 (19).

sAS1. Bgl II-digested sAS1 DNA (100 ng) was hybridized with 2–10 μg of total adult *Drosophila* RNA for 3 hr at 48 or 49°C in 20 μl of 80% formamide/0.4 M NaCl/40 mM 1,4-piperazinediethanesulfonic acid, pH 6.8/1 mM EDTA. The reaction was diluted into 200 μl of ice-cold S1 nuclease buffer and digested with 2000–5000 units of S1 nuclease at room temperature. The rest of the procedure was carried out as described above for mAC1-3 DNA.

RESULTS

Isolation of the ADH Gene. To obtain a hybridization probe enriched for ADH mRNA sequences, total polyadenylylated RNA from 1- to 3-day-old *Drosophila* adults was fractionated by agarose gel electrophoresis, and RNA in the size range 600–1500 nucleotides (nt) was recovered [ADH protein is encoded in approximately 760 nt (3)]. The fractions containing ADH mRNA were identified by their ability to stimulate incorporation of ^{3}H-labeled amino acids (20) into protein that could be specifically immunoadsorbed by an ADH antibody coupled to Sepharose (21). (We thank W. Sofer for providing ADH antibody and purified ADH protein.) The immunoadsorbed protein comigrated with purified ADH protein on a $NaDodSO_4$/polyacrylamide gel (22).

^{32}P-Labeled cDNA was prepared (23) from the ADH mRNA-enriched fraction and used to screen (24) a bacteriophage λ library of random, high molecular weight (>15 kb) *Drosophila* DNA (16). Of 50,000 plaques screened, 31 hybridized strongly.

To determine which of the phage clones carried *Adh* se-

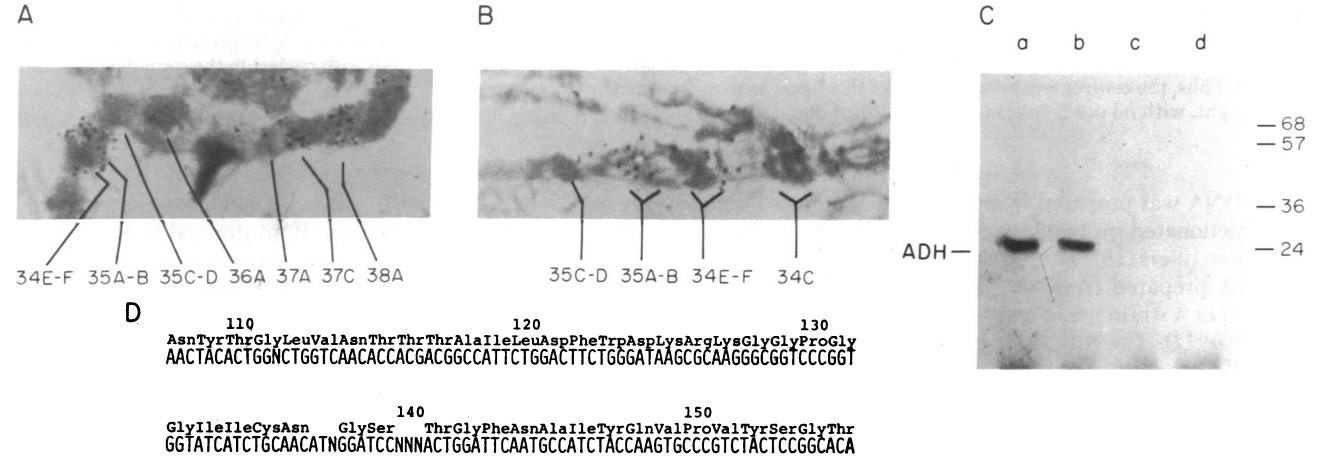

FIG. 2. Confirmation of the presence of *Adh* sequences in gAC1 DNA. (*A* and *B*) Salivary gland polytene chromosomes were prepared from wild-type (Oregon R) third-instar larvae. gAC1, 2E2, and 11E4 DNAs were ^{3}H-labeled by nick-translation and hybridized *in situ* to the chromosome preparations (25). 2E2 and 11E4 are cloned DNAs that hybridize only to regions 38A and 37C of *Drosophila* polytene chromosomes, respectively (J. Hirsch, personal communication). (*C*) A1.8, a *Drosophila* actin clone (25), sAC1, and gAC1 DNAs were cleaved with *Eco*RI and bound to nitrocellulose as described (26). Approximately 30 μg of total adult RNA was hybridized with each nitrocellulose filter. The filter was washed and the RNA was eluted as described (26) and translated in a rabbit reticulocyte *in vitro* translation system (20). The translation products were immunoadsorbed to ADH antibody immobilized on Sepharose (21) and fractionated on a $NaDodSO_4$/polyacrylamide gel (22). The filter hybridizations were performed with 26 μg of sAC1 DNA (lane a), 12 μg of gAC1 DNA (lane b), 10 μg of A1.8 DNA (lane c), and no DNA (lane d). Protein standard molecular weights ($\times 10^{-3}$) are listed at the right. (*D*) The nucleotide sequence was read in both directions from a *Bam*HI site in the sAS1 *Drosophila* DNA insert (27). The sequence was read from only one strand and only those sequences that were independently read from the autoradiogram by three people are presented. Nucleotides listed as N could not be read unambiguously. The amino acid residue numbers are shown.

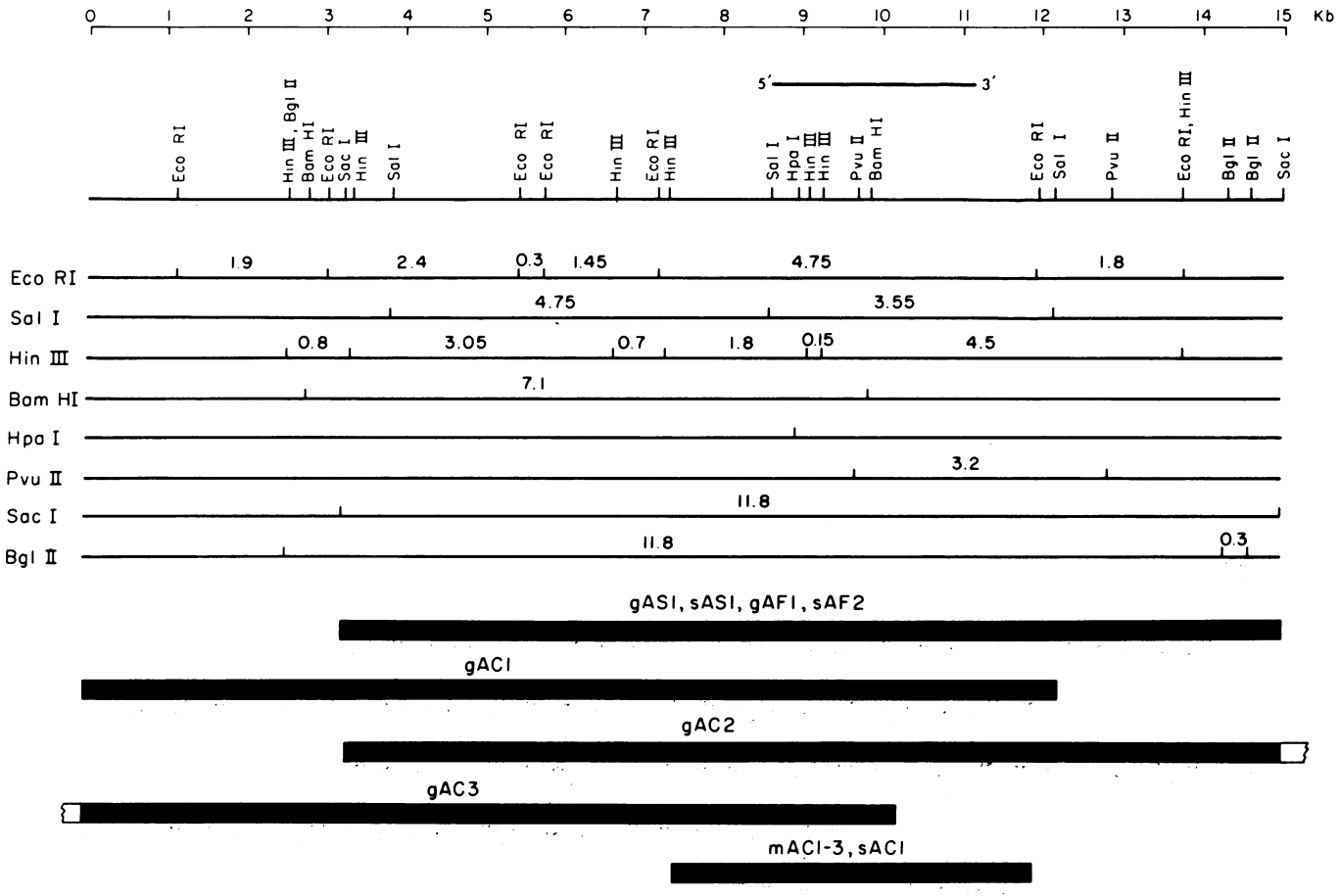

FIG. 3. Restriction map of the *Adh* region. A 4.7-kb *Eco*RI fragment of gAC1 that hybridizes to ADH mRNA was subcloned into pBR322 (sAC1). Two additional *Adh* clones (gAC2 and gAC3) were obtained by screening the *Drosophila* library with sAC1. Also, Adh^S (slow electrophoretic mobility/low specific activity) and Adh^F (fast electrophoretic mobility/high specific activity) alleles were cloned as 11.8-kb *Sac* I fragments in the bacteriophage vector Charon 10. These *Sac* I fragments were subcloned in a pBR322-derived *Sac* I vector (pSV2) to give clones sAS1 (from Adh^S) and sAF2 (from Adh^F). A restriction map was determined for the *Drosophila* inserts in gAC1 and sAS1. Some restriction site polymorphisms exist between the Canton S, Adh^S, and Adh^F strains from which gAC1, sAS1, and sAF2 were derived; however, these are not listed on the map. The restriction sites in the region overlapped by sAS1 and gAC1 are the same in both clones. Bars at the bottom of the figure denote the regions cloned in the respective recombinant molecules. gAC2 extends 2.0 kb to the right and gAC3 5.8 kb to the left of the region depicted in the figure. No cDNA hybridization to cloned restriction fragments was observed outside of the bar located above the composite restriction site map. The 5'- to -3' orientation of the ADH mRNA is from left to right. Fine structure genetic mapping within and around the *Adh* locus and protein sequencing of *Adh* null and electrophoretic variants has established the orientation of the ADH protein coding sequences on the *Drosophila* second chromosome (11). Together, the physical and genetic data orient the cloned DNA with respect to the second chromosome centromere. Thus, the centromere lies to the left of this figure with *outspread wings* (*osp*) being the closest proximal genetic locus. The telomere lies to the right, with *no ocelli* (*noc*) being the closest distal genetic locus. The physical distances of these adjacent genetic loci to *Adh* is not known.

quences, DNA was prepared from each clone, digested with *Eco*RI, fractionated on two agarose gels, and transferred to nitrocellulose filters (18). The duplicate filters were hybridized with cDNA prepared from total RNA of a wild-type strain (Oregon R) or a strain genetically deficient for ADH protein (Adh^{fn2}/Sco^{R+4}). The *Adh* deficiency strain is heterozygous for two different deletion chromosomes that overlap in the *Adh* region. By cytogenetic analysis these flies appear to be lacking the *Adh* locus (4). DNA from one phage, gAC1, hybridized to cDNA from the wild-type strain but not to cDNA from the *Adh* deficiency strain (Fig. 1).

Three experiments were carried out to confirm that gAC1 contains the *Adh* structural gene (Fig. 2). First, ^{3}H-labeled gAC1 DNA was hybridized *in situ* to salivary gland polytene chromosomes (25). Hybridization was observed at 35 A-B only, which agrees with the cytogenetic localization of *Adh* (4). Second, total adult *Drosophila* RNA was hybridized to nitrocellulose filters containing DNA from gAC1 or a *Drosophila* actin gene (26, 28). The RNA that hybridized was eluted, and translated *in vitro*, and the products were immunoadsorbed and fractionated by $NaDodSO_4$/polyacrylamide gel electrophoresis. ADH was translated only from RNA derived from the gAC1 DNA filter. Third, the nucleotide sequence of a region of the cloned DNA that hybridized to ADH-enriched mRNA was determined (27). As shown in Fig. 2, this DNA sequence agreed with the ADH protein sequence derived by Thatcher (3). In addition, C. Benyajati and W. Sofer (personal communication) have isolated and characterized an *Adh* cDNA clone and have determined that the cDNA clone hybridizes specifically to gAC1.

Characterization of *Adh* Genomic DNA. To characterize *Adh* genomic DNA, a number of different *Adh* clones and subclones were prepared and a map of their restriction sites was derived (Fig. 3). Southern blots (18) of various restriction enzyme digests of *Drosophila* DNA were probed with *in vitro*-labeled DNA from a number of different *Adh* clones. The *Adh* gene and its flanking sequences were present only once in the *Drosophila* genome, and no detectable rearrangements occurred during cloning.

Comparisons of the restriction maps of gAC 1, 2, 3 (from

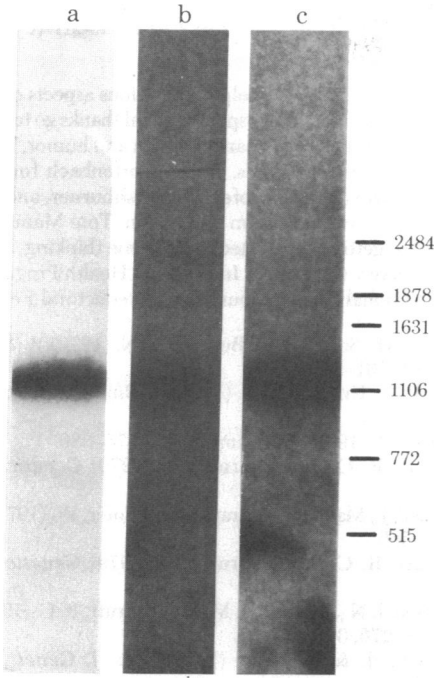

FIG. 4. Hybridization of cloned *Adh* DNA to total *Drosophila* adult RNA. Four micrograms of total adult *Drosophila* RNA was run in each lane of a 1% agarose formaldehyde gel (29). The electrophoresis buffer was 40 mM morpholinopropanesulfonic acid, pH 7.2, 10 mM Na acetate/1 mM EDTA. The gel contained 6% formaldehyde. Samples were heated at 55°C for 30 min in gel buffer containing 50% dimethyl sulfoxide and 6% formaldehyde before they were loaded onto the gel. After the run, the gel was washed three times with an excess of H_2O and then twice with 0.2 M ammonium acetate; the second wash contained ethidium bromide (0.5 μg/ml). After photography of unlabeled DNA size markers under short-wave UV illumination, the gel was soaked in 1.5 M NaCl/0.15 M Na citrate and the RNA was transferred to nitrocellulose filter paper. After transfer, the filter was baked for 3 hr at 80°C under reduced pressure and then hybridized with labeled DNA essentially as described (30). Lanes a, b, and c are total adult RNA hybridized with nick-translated gAS1, gAC2, and gAC3, respectively. (Transfer of RNA to nitrocellulose is an unpublished observation of E. Butler, B. Seed, and D. Goldberg.)

Canton S, an Adh^S strain), sAS1 (from Schwenk Forest, an Adh^S strain), and sAF2 (from an Adh^F strain) revealed very few differences. However, sAF2 contains a 200-base-pair insertion relative to both sAS1 and gAC2. It is not known whether this represents a deletion in the Adh^S strains or the insertion of DNA in Adh^F strains from elsewhere in the genome. This insertion has been mapped between the *Pvu* II and *Sal* I sites at 12.2 and 12.9 kb on the map of Fig. 3. This lies 3′ to the mRNA coding sequences, and its effects, if any, on ADH expression are unknown.

Transcription Mapping of the *Adh* Gene. There appears to be only one major adult RNA transcript originating from the region encompassed by the cloned *Adh* region. Three clones (gAC2, gAC3, sAS1) spanning approximately 21 kb were used to probe total adult *Drosophila* RNA which was run on denaturing formaldehyde agarose gels and transferred to nitrocellulose (Fig. 4). Only one hybridizing species, 1100–1150 nt, was detected. Because only 760 nt are required to code for the ADH protein, the ADH mRNA contains approximately 400 nt of noncoding sequences including poly(A). This experiment does not rule out low-abundance adult transcripts or transcripts produced at other stages of development.

To map the transcribed regions of the *Adh* clones, labeled cDNA prepared from adult poly(A)$^+$ RNA was used to probe blots of restriction digests of different cloned DNAs spanning the *Adh* region. The region of hybridization is shown in Fig. 3. The sequence of DNA surrounding the *Bam*HI site was determined (Fig. 2). This sequence corresponds to the amino acid sequence near the center of the protein and establishes the orientation of 5′- to -3′ transcription as left to right on the map in Fig. 3.

To determine whether the *Adh* gene is interrupted by intervening sequences, S1 nuclease mapping studies were performed using either linearized sAS1 plasmid DNA or a subclone (mAC1-3) of the 4.75-kb *Eco*RI fragment in the single-strand phage vector M13mp2 (14). The results of two mapping experiments are shown in Fig. 5. In both experiments RNA was hybridized to DNA, and the resultant hybrid was digested with S1 nuclease. The digested nucleic acid was run on a two-dimensional gel system, nondenaturing in the first dimension and alkaline denaturing in the second dimension. Migration in the nondenaturing dimension indicates the total hybrid length [mRNA less poly(A)]; migration in the denaturing dimension indicates the sizes of DNA fragments left intact by the digestion. Spots off the diagonal indicate the presence of intervening sequences in the genomic DNA. No significant hybridization was seen in control experiments with tRNA hybridized to mAC1-3 DNA or sAS1 DNA (data not shown). With adult RNA hybridized to mAC1-3 DNA (*A*) or sAS1 DNA (*B*), two major

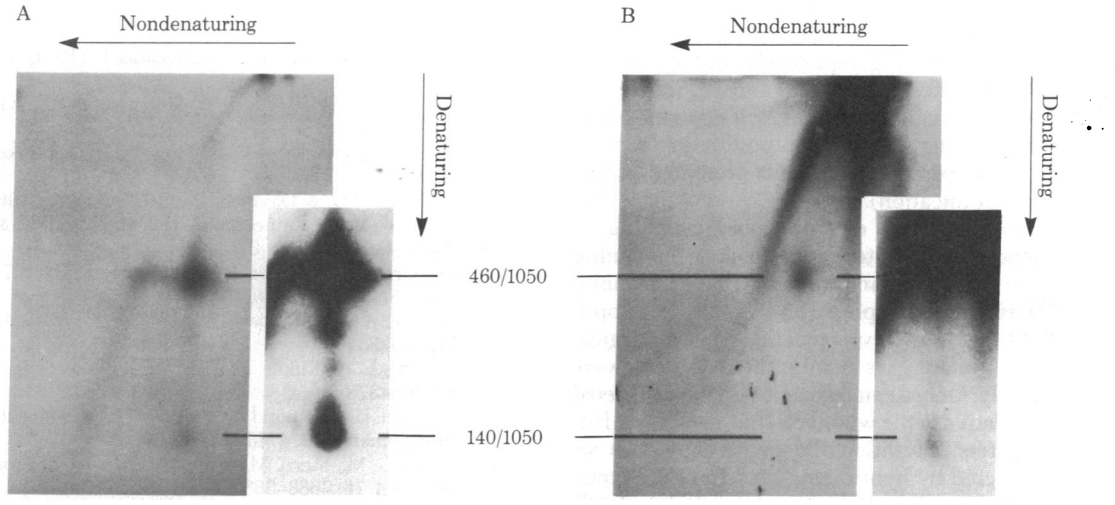

FIG. 5. S1 nuclease mapping analysis of ADH mRNA/genomic DNA hybrids. Each spot is identified by two numbers separated by a slash. The first number refers to the spot mobility in the alkaline denaturing dimension, and the second number refers to the mobility in the neutral dimension. (*A*) mAC1-3 DNA (100 ng) hybridized to 5 μg of polyadenylylated adult RNA. (*B*) sAS1 DNA (100 ng) cleaved with *Bgl* II and hybridized with 5 μg of polyadenylylated adult RNA. The minor spots in *A* have not been explained.

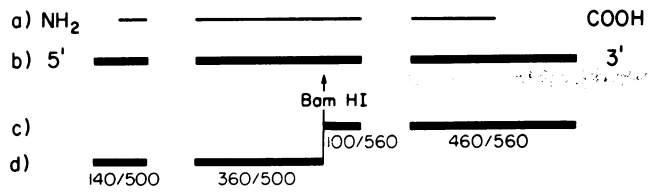

FIG. 6. Map of the *Adh* exons. Line a, approximate position of the regions of ADH mRNA coding for ADH protein. Line b, organization of exons within the genomic DNA. The relative positions of the intervening sequences are shown as blank spaces, although the size of the space is *not* indicative of the length of the intervening sequence. Lines c and d, interpretation of the results of the S1 nuclease mapping experiments performed with isolated *Bam*HI fragments of sAC1. These experiments were performed with the *Bam*HI fragment covering the 3′ end (line c) and the 5′ end (line d) of the *Adh* gene. Each DNA fragment is denoted by the spot it was associated with in the actual experiment according to the nomenclature of Fig. 5.

spots corresponding to 140 and 460 nt were observed in the second dimension. Both spots were derived from a 1050-nt RNA as determined from the first-dimension mobility. This size is the same as that determined for ADH RNA (Fig. 4). The 460- and 140-nt spots, which correspond to exons of ADH RNA, cannot account for the entire mRNA length of 1100–1150 nt.

In order to position the exon sequence on the restriction map of Fig. 3, S1 nuclease mapping experiments were performed with purified *Bam*HI fragments of sAC1 (Fig. 6). With the *Bam*HI fragment spanning the 5′ end of the RNA, spots corresponding to species 140/500 and 360/500 were observed (data not shown; see Fig. 5 for spot nomenclature). With the *Bam*HI fragment spanning the 3′ end of the RNA, spot 100/560 was observed. The expected complementary fragment 450/560 would be obscured in this experiment by hybridizing sequences along the diagonal. These data suggest that *Bam*HI cuts a 460-nt exon into 360- and 100-nt fragments (5′ and 3′, respectively). This interpretation assumes that the 460-nt spot found in S1 nuclease mapping experiments with mAC1-3 and sAS1 is composed of two fragments of similar size. In some experiments, two spots have been observed in the region of the gel corresponding to 460 nt.

The presence of an intervening sequence located 100 nucleotides 3′ to the *Bam*HI site was confirmed by comparing the location of restriction sites mapped on the cloned DNA with the predicted restriction sites determined from the amino acid sequence. Fine structure restriction endonuclease mapping studies revealed *Hpa* II and *Hae* III sites that are inconsistent with the amino acid sequence 165, 180, and 204 nt from the *Bam*HI site. Thus, an intervening sequence is located less than 165 nt from the *Bam*HI site. C. Benyajati and W. Sofer have determined sequences of regions from an *Adh* genomic clone and have found intervening sequences in positions consistent with the S1 nuclease mapping experiments presented in this paper (personal communication).

Hybridization with complex cDNA probes including or lacking *Adh* sequences was used to identify a clone containing the *Drosophila Adh* gene. Similar approaches have been used by others (31, 32) using cDNAs prepared from mRNA populations from cells of different developmental or physiological states. In this study, mutant flies lacking the *Adh* gene were constructed genetically. Only a small genetic region was deleted in the mutant flies relative to the wild-type flies, conferring fine selectivity to the screen. Additionally, single-copy *Adh* sequences can be detected by hybridizing nick-translated total *Drosophila* DNA to nitrocellulose filters containing cloned *Adh* DNA (unpublished observations). Thus, it may be possible to use differential screens to identify clones from various genetic loci at which homozygous viable deficiencies have been isolated, such as the *white* locus and the *dunce* locus (cyclic AMP phosphodiesterase).

There are more people who helped in various aspects of this work than can be acknowledged in this space. Special thanks go to Brian Seed and Randy Smith for their trenchant criticism and humor, Mark Silver for his inimitable technical skills, Jennie Cortenbach for preparing media, and Tip Benyajati, Bill Sofer, Mike Ashburner, and Bob Karp for making *Adh* a pleasant system to work on. Tom Maniatis, in particular, helped in getting me started, getting me thinking, and getting me space and money (a National Institutes of Health Program Project Grant and a National Science Foundation Predoctoral Fellowship).

1. Ursprung, H., Sofer, W. & Burroughs, N. (1970) *Wilhelm Roux Arch.* **164,** 201–208.
2. Sofer, W. & Ursprung, H. (1968) *J. Biol. Chem.* **243,** 3110–3115.
3. Thatcher, D. (1980) *Biochem. J.* **187,** 875–886.
4. Woodruff, R. C. & Ashburner, M. (1979) *Genetics* **92,** 117–132.
5. O'Donnell, J., Mandel, H., Krauss, M. & Sofer, W. (1977) *Genetics* **86,** 553–566.
6. Woodruff, R. C. & Ashburner, M. (1979) *Genetics* **92,** 133–149.
7. Thompson, J. N., Ashburner, M. & Woodruff, R. C. (1977) *Nature (London)* **270,** 363.
8. McDonald, J. & Ayala, F. (1978) *Can. J. Genet. Cytol.* **20,** 159–175.
9. Vigue, C. & Sofer, W. (1976) *Biochem. Genet.* **14,** 127–135.
10. O'Donnell, J., Gerace, L., Leister, F. & Sofer, W. (1975) *Genetics* **79,** 73–83.
11. Ashburner, M., Camfield, R., Clarke, B., Thatcher, D. & Woodruff, R. (1979) in *Eucaryotic Gene Regulation,* eds. Axel, R., Maniatis, T. & Fox, C. F. (Academic, New York), pp. 95–106.
12. Blattner, F. R., Williams, B. G., Blechl, A. E., Thompson, K. D., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196,** 161–169.
13. Boyer, H. W. & Roulland-Dussoix, D. (1969) *J. Mol. Biol.* **41,** 459–472.
14. Messing, J. (1979) *NIH Tech. Bull.* **2,** No. 2, 43–48.
15. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. & Boyer, H. W. (1977) *Gene* **2,** 95–113.
16. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15,** 687–701.
17. Favaloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65,** 718–749.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
19. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1184–1188.
20. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67,** 247–256.
21. March, S. C., Parikh, I. & Cuatrecasas, P. (1974) *Anal. Biochem.* **60,** 149–152.
22. Anderson, C. W., Baum, P. R. & Gesteland, R. F. (1973) *J. Virol.* **12,** 241–252.
23. Buell, G. N., Wickens, M. P., Payvar, F. & Schimke, R. T. (1978) *J. Biol. Chem.* **253,** 2471–2482.
24. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
25. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. (1974) *Cell* **3,** 315–325.
26. Ricciardi, R. P., Miller, J. S. & Roberts, B. E. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4927–4931.
27. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
28. Fyrberg, E. A., Kindle, K. L., Davidson, N. & Sodja, A. (1980) *Cell* **19,** 365–378.
29. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16,** 4743–4751.
30. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
31. St. John, T. & Davis, R. W. (1979) *Cell* **16,** 443–452.
32. Lepesant, J. A., Kejzlarova-Lepesant, J. & Garen, A. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5570–5574.