Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 3259–3263, May 1990
Genetics

Characterization of two RNAs transcribed from the cis-regulatory region of the *abd-A* domain within the *Drosophila* bithorax complex

SUSAN CUMBERLEDGE, ANNABELLE ZARATZIAN*, AND SHIGERU SAKONJU†

Howard Hughes Medical Institute, Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT 84132

Communicated by E. B. Lewis, January 30, 1990

ABSTRACT **The infra-abdominal (*iab*) regions of the biothorax complex are thought to cis-regulate expression of the abdominal-A (*abd-A*) and Abdominal-B (*Abd-B*) transcripts. These large cis-regulatory regions are also actively transcribed. Here we present a detailed analysis of the transcription products of the *iab-4* region. During early embryogenesis a 6.8-kilobase (kb) RNA is transcribed and processed to yield 1.7- and 2.0-kb poly(A)$^+$ RNAs. These RNAs are expressed from the time of cellular blastoderm formation to germ-band shortening and are localized in parasegments 8–14 (the posterior second through the anterior ninth abdominal segments). Sequence analysis of the two RNAs suggests that they may not be translated. We discuss some possible functions of the 1.7/2.0-kb *iab* RNAs and the significance of their similarity to the bithoraxoid (*bxd*) RNAs and other transcripts from the *iab* regions of the *abd-A* and *Abd-B* domains.**

The bithorax complex (BX-C) is required for specifying the correct metameric identity of parasegments (PS) 5–14, which form the posterior second thoracic segment (T2p) through the anterior ninth abdominal segment (A9a) of *Drosophila melanogaster* (for review see refs. 1 and 2). The complex can be divided into three lethal complementation groups or domains: Ultrabiothorax (*Ubx*), abdominal-A (*abd-A*), and Abdominal-B (*Abd-B*), specifying PS5–13, PS7–13, and PS10–14, respectively (3, 4). Thus these three homeotic genes are thought to act combinatorially to determine the thoracic and abdominal identities.

One of the notable features of the BX-C is the presence of large cis-regulatory regions located in each of the three bithorax domains. The size and number of these regions may reflect the complex regulatory controls that act on the *Ubx*, *abd-A*, and *Abd-B* genes. The cis-regulatory region of the *Ubx* domain, the bithoraxoid (*bxd*)/postbithorax (*pbx*) region, has been most extensively characterized. This 40-kilobase (kb) region is thought to contain multiple enhancer-like elements (5). One unusual feature of the *bxd/pbx* region is the presence of multiple transcripts, which are expressed during early embryogenesis and the pupal stages (6, 7). Sequence analysis of the early *bxd* transcripts, however, revealed no significant open reading frames (ORFs) that may code for proteins (7). Irish *et al.* (8) have found that *Ubx* and *bxd* transcription is coordinately regulated in *hb*$^-$ and *osk*$^-$ embryos, suggesting that activation of the *bxd* promoter depends on developmental cues that also regulate *Ubx* transcription.

In this paper, we describe an analysis of transcription from one of the putative cis-regulatory regions of the *abd-A* domain. The *abd-A* domain contains a 23-kb protein-coding transcription unit, which is expressed in PS7–13 (ref. 9; S.S., B. Appel, and M. Lamka, unpublished work). Unlike the *Ubx* and *Abd-B* protein-coding transcripts, the *abd-A* transcript appears to encode a single homeodomain-containing protein (S.S., B. Appel, and M. Lamka, unpublished work). In addition to the lethal *abd-A* mutations, which disrupt the protein-coding transcripts, Lewis (10) has identified nonlethal mutations that affect the development of subsets of the parasegments affected by the lethal mutations. Many of the nonlethal lesions map to regions outside of the *abd-A* transcription unit; these are termed the *iab* regions (1, 11, 12). The *iab-2* region is located downstream and the *iab-3* and *iab-4* regions extend ≈15 kb upstream of the *abd-A* transcription unit (12). Genetic analysis has suggested that the *iab-2*, *-3*, and *-4* regions are cis-regulatory regions of *abd-A* (2, 11, 12). In this paper we describe the structure, sequence, and temporal and spatial expression of two poly(A)$^+$ RNAs transcribed from the *iab-4* region.‡ We speculate about the function of these RNAs, and about why transcriptional activity within cis-regulatory regions is a general feature throughout the BX-C.

MATERIALS AND METHODS

RNA Preparation and Northern Hybridization. Preparation of RNA and Northern hybridizations were as described (12). Single-stranded hybridization probes (probes 1–4; Fig. 1) were prepared from pφ3905, pφ3906, pφ3907, and pφ3913. These plasmids were constructed by inserting a 5.3-kb *Eco*RI fragment [corresponding to map positions +72.5 to +78 of Karch *et al.* (12)], a 3.5-kb *Hin*dIII fragment (+71.5 to +75), a 5.5-kb *Eco*RI fragment (+79.5 to +85), or a 2-kb *Eco*RI fragment (+85 to +87), respectively, into pφX. The pφX vector and the conditions for the viral strand synthesis have been described in detail (13).

Analysis of cDNA Clones. The cDNA c15 clone was isolated from a 1.5- to 5-hr Oregon-R embryonic cDNA library constructed by M. Goldschmidt-Clermont (13). A 5.3-kb *Eco*RI genomic DNA fragment (+72.5 to +78) was nick-translated and used to probe the library. Four other cDNA clones (including cDNA 7-2) were isolated from a 4- to 8-hr cDNA library constructed by Brown and Kafatos (14). These cDNA clones were generously provided by E. B. Lewis's laboratory. To localize the exons within the BX-C, we used the cDNA c15 and 7-2 clones to probe Southern blots containing *Eco*RI- or *Eco*RI/*Sal* I-digested DNA isolated from λ phage clones 2226, 2235, 2255, 2265, 2279, 8004, and 8019 (12). DNA sequencing was performed as described (15).

S1 Nuclease Protection and Primer Extension Studies. These studies were performed as described by Sharp *et al.* (16), with modifications (15). The sequence of the 25-base oligodeoxynucleotide primer was 5′-GTAATTAAATAAACACTGGGCACAC-3′.

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: BX-C, bithorax complex; ORF, open reading frame; PS*n*, parasegment *n*; A*n*, abdominal segment *n*.

*Present address: Albert Einstein College of Medicine, Bronx, NY 10461.
†To whom reprint requests should be addressed.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M32109).

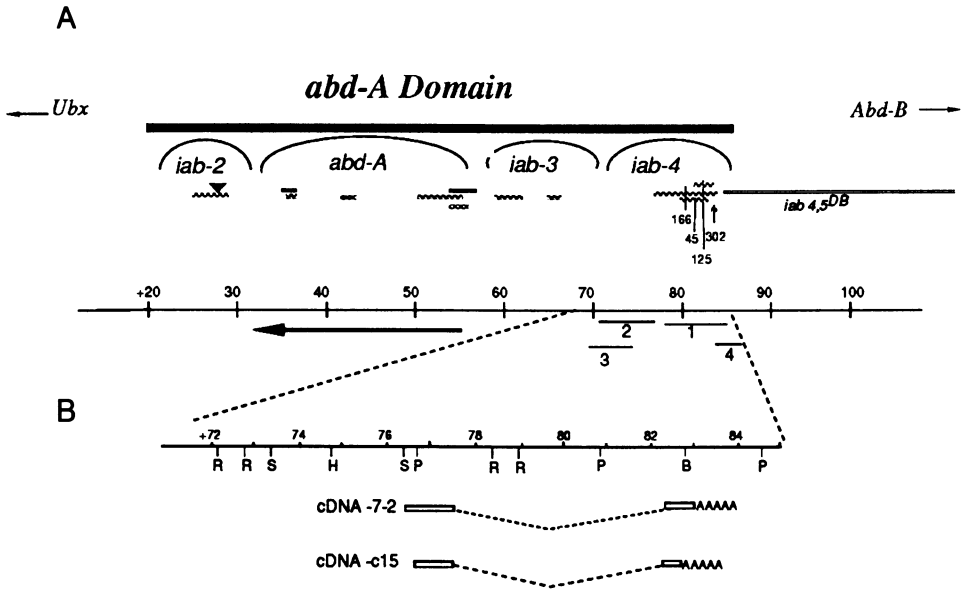

FIG. 1. Molecular map of the *abd-A* domain. DNA scale is in kilobases. Numbering is that of Karch *et al.* (12). (*A*) Positions of rearrangement breakpoints are shown by wavy lines. Vertical lines point to breakpoints associated with *iab-4* alleles. Deletions in *abd-A* are indicated by solid bars, and the *iab-4,5*^{DB} deletion is shown by an open bar. Insertion in *iab-2* is indicated by filled triangle. The *abd-A* transcription unit is shown as a heavy line; arrow denotes the direction of transcription. Locations of the DNA probes used for Northern analysis are indicated by numbered lines. (*B*) Localization of the *iab-4* exons within the *abd-A* domain. Open boxes represent restriction fragments that hybridized to the indicated cDNA. The 5′-3′ orientation of the cDNAs is left to right. Restriction sites: R, *Eco*RI; S, *Sal* I; H, *Hin*dIII; P, *Pst* I.

***In Situ* Hybridization.** Embryos were collected, fixed, and hybridized as described (17). Digoxigenin-labeled probes were prepared by the random-primer labeling technique, using a kit according to the protocol recommended by the manufacturer (Boehringer Mannheim), with a 1-kb *Eco*RI–*Cla* I fragment isolated from the cDNA c15 clone as a template. These probes hybridize to both the 1.7- and 2.0-kb transcript. We identified the parasegments expressing the 1.7/2.0-kb *iab* transcripts by comparing embryos of the same developmental stages that were hybridized with an engrailed probe in parallel experiments (data not shown).

RESULTS

Transcripts Encoded Within the *iab-4* Region. The five known *iab-4* mutations map within the abdominal region of the BX-C, from +76 kb to +85 kb on the genomic map (Fig. 1*A*). Transcripts encoded within the *iab-4* region were detected by probing developmental Northern blots with four genomic DNA fragments spanning this region (probes 1–4; Fig. 1*A*). Probe 1, spanning from +79.5 kb to +85 kb, detected two poly(A)$^+$ RNAs, 1.7 and 2.0 kb in length (Fig. 2). The 1.7-kb RNA was observed only during the first 8 hr of embryogenesis; however, the 2.0-kb RNA could still be detected at very low levels in later-stage embryos. Neither transcript was detected in larval, pupal, or adult stages. By using strand-specific probes, we determined the direction of transcription for both RNAs to be from left to right, opposite the direction of *abd-A* transcription (Fig. 1). Both transcripts also hybridized with probe 2 (+72.5 to +78), but not with probe 3 (+71.5 to +75) or 4 (+85 to +87) (data not shown). Thus both RNAs are apparently transcribed from sequences between +74 and +87. The minor RNA band at 6.8 kb most likely corresponds to unprocessed primary transcript (see below), while the low molecular weight band may be a breakdown product (Fig. 2). Since the 1.7- and 2.0-kb RNAs are transcribed from the *iab-4* region of the *abd-A* domain, we will refer to them as the 1.7/2.0-kb *iab* transcripts. A third poly(A)$^+$ RNA, transcribed in the opposite direction was detected by probes 2 and 3; however, it extends beyond the region defined by the *iab-4* mutations, spanning from the *iab-3* region to *iab-7* (S.S., unpublished data).

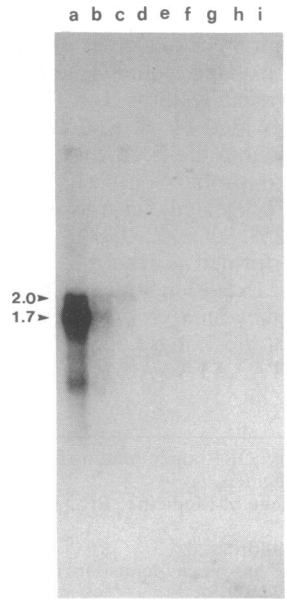

FIG. 2. Developmental Northern analysis. A 5.3-kb *Eco*RI fragment (corresponding to map position +72.5 to +78) was used to probe a Northern blot of poly(A)$^+$ RNA (5 μg per lane) isolated from flies at various developmental stages. Lanes: a, 0- to 8-hr embryos; b, 8- to 16-hr embryos; c, 16- to 24-hr embryos; d, first-instar larvae; e, second-instar larvae; f, third-instar larvae; g, first- and second-day pupae; h, third- and fourth-day pupae; i, adults. Unlabeled arrowhead denotes 6.8 kb. The RNA blot was rehybridized with a *Ubx*-specific probe to confirm the presence of poly(A)$^+$ RNA in all lanes (data not shown).

Molecular Characterization of the 1.7/2.0-kb *iab* Transcripts. Characterization of the five cDNA clones corresponding to the 1.7/2.0-kb *iab* RNAs has shown that they fall into two classes (Fig. 1*B*; S. Smolik-Utlaut, personal communication). The longest clone from each class (cDNA c15 and cDNA 7-2) is shown in Fig. 1*B*. Southern blot hybridization of the cDNAs to genomic DNA encompassing the chromosomal walk through the *abd-A* domain revealed that both cDNAs are derived from the same 7.0-kb region, from +76 kb to +83 kb on the genomic map. The two cDNAs were also used to probe Northern blots; both clones hybridized to the 1.7- and 2.0-kb RNAs (data not shown; S. Smolik-Utlaut, personal communication).

The 5′ ends of the RNAs were determined by S1 analysis and primer extension experiments. A single start site was found 76 base pairs (bp) upstream from the 5′ end of cDNA 7-2 and 169 bp upstream from cDNA c15 (Fig. 3). The complete DNA sequence of cDNA c15 was determined. The 5′ and 3′ regions and the exon borders of cDNA 7-2, as well as the corresponding genomic DNA regions, were also sequenced. The combined results of the sequence analyses are shown in Fig. 4. Comparison of the cDNA and genomic sequences indicates that both cDNAs are derived from the same two exons. Both cDNAs contain a run of adenine residues at their 3′ ends. cDNA 7-2 is polyadenylylated 303 bp farther 3′ than cDNA c15. This 303 bp corresponds to the size difference between the two transcripts detected on Northern blots (Fig. 2).

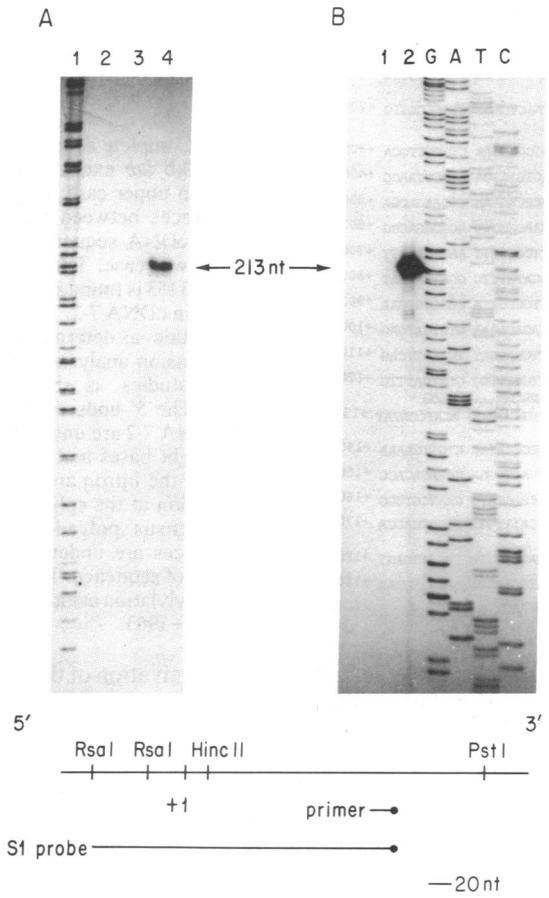

FIG. 3. Identification of the 1.7/2.0-kb *iab* transcription start site. (*A*) Primer extension analysis. The location of the 25-base oligonucleotide primer used for extension is shown below the autoradiogram. Lanes: 1, molecular size markers; 2, no-RNA control; 3, 2 μg of yeast poly(A)$^+$ RNA; 4, 20 μg of total RNA isolated from 3- to 7-hr embryos. (*B*) S1 nuclease protection analysis. The location of the 313-base single-stranded probe is shown below the autoradiogram. Lanes: 1, no-RNA control; 2, 2 μg of poly(A)$^+$ RNA isolated from 3- to 7-hr embryos; G, A, T, and C, products of dideoxynucleotide chain-termination reactions showing the sequence of the corresponding genomic DNA. The 3′ ends of the 25-base oligonucleotide and the S1 probe are identical. nt, Nucleotides.

Taken together, these results indicate that the primary transcript, ≈6.8 kb in length, contains one large intron (4.8 kb) and two polyadenylylation sites, giving rise to the 1.7/2.0-kb *iab* RNAs. There is no consensus TATA sequence at the appropriate distance from the transcription start site, the sequence TATCT being the closest fit (Fig. 4, nucleotides −28 to −24). The exon/intron boundaries include consensus splice sequences, except for a TG replacing an AG at the first exon/intron border. This TG/GTAAGT sequence is identical to one of two alternative splice sequences for exon 7 of the *Antp* gene (18). The first exon is 1284 bp long, and the second exon is 380 or 683 bp long depending on which polyadenylylation site is used. The polyadenylylation sites at nucleotide positions +1662 and +1963 (Fig. 4) are preceded by AATAAA consensus polyadenylylation signals at +1641 and +1943 (19).

We have scanned the 1.7/2.0-kb *iab* cDNA sequences for ORFs. Both the 1.7- and 2.0-kb RNAs contain multiple short ORFs; the longest AUG-initiated ORFs found in cDNA c15 and cDNA 7-2 are 216 and 453 bp, respectively. These ORFs are found at the 3′ ends of the RNAs and are preceded by shorter AUG-initiated ORFs. For example, the 453-bp ORF from nucleotides +1252 to +1705 is preceded by five short AUG-initiated ORFs. Since the 1.7-kb RNA is polyadenylylated at position +1662, this ORF is disrupted in the shorter transcript. Translation of the 453-bp ORF would yield a protein of ≈17 kDa. Codon usage in the 1.7/2.0-kb *iab* transcripts was examined by matching to a collection of ORF sequences in *Drosophila* regulatory genes, with the use of a Staden program (ref. 20; K. Burtis, personal communication). The *iab* ORFs show poor *Drosophila* codon-usage bias. The presence of multiple short ORFs and poor *Drosophila* codon usage suggests that these sequences are probably not translated. We also searched the National Institutes of Health and EMBL sequence data libraries (August 1989) for DNA sequences that show similarity to the 1.7/2.0-kb *iab* cDNA sequences and found no significant matches.

Localization of 1.7/2.0-kb *iab* Transcripts in Developing Embryos. Using *in situ* hybridization techniques, we determined the spatial distribution of the 1.7/2.0-kb *iab* transcripts during embryogenesis (Fig. 5). Both transcripts can first be detected at the time of cellularization [stage 5 embryos, ≈2.5 hr into development (21)]. The RNAs accumulate between 14% and 40% egg lengths, as measured from the posterior end (Fig. 5*A*). Hybridization appears to be concentrated over nuclei and is relatively uniform throughout this domain. This region contains the primordia for PS8–14 (21). The same pattern of hybridization was observed when embryo sections were hybridized with antisense RNA probes complementary to the 1.7/2.0-kb *iab* RNAs (data not shown). Therefore, the observed pattern is due to the 1.7/2.0-kb *iab* RNAs and not to other RNAs that might be transcribed from the opposite strand.

By the germ-band-extended stage (Fig. 5*B*) hybridization is observed from PS8 through PS14. Within each parasegment, the anterior portion shows the most intense hybridization (Fig. 5 *B* and *C*). This corresponds to the region within the abdominal parasegments where the *abd-A* transcripts are also expressed at high levels (S.S., B. Appel, and M. Lamka, unpublished work). The parasegmental nature of the expression is clearly seen: the hybridization bands straddle the tracheal pits that mark segmental borders (Fig. 5 *C* and *D*). The intensity of hybridization becomes considerably weaker during the germ-band-shortening stage; however, hybridization can still be observed in the ventral nerve cord from PS9 to PS14 and very faintly over PS8 (data not shown).

DISCUSSION

We have shown that a region of DNA thought to contain cis-regulatory elements for the *abd-A* gene is transcribed. Transcriptional analysis of this region detected 2.0- and 1.7-kb poly(A)$^+$ RNAs, both of which contain multiple short ORFs, with poor codon-usage bias. Our results suggest, but do not prove, that the RNAs from the *abd-A* cis-regulatory region are not translated.

How well does the chromosomal location of these transcripts correlate with the functions expected in the region upstream of *abd-A*? Lewis (10, 11) has isolated two classes of *iab* mutations in this region. The *iab-3* mutations cause a transformation of abdominal segments A3 through A6 toward A2; the *iab-4* mutations lead to a transformation of A4 and A5 to A3 and also cause a loss of gonads. All four *iab-4* breakpoints (*iab-4*^{166}, *iab-4*^{45}, *iab-4*^{125}, and *iab-4*^{302}) map within the 6.8-kb transcription unit (15). The breakpoint of the *iab-4*^{302} lesion, originally thought to lie distal to the *Bam*HI site at +83 kb, has been cloned and found to be proximal to the *Bam*HI site (M. Lamka and S.S., unpublished data). The fifth allele, *iab-4,5*^{DB}, has a 28-kb deletion beginning ≈2 kb downstream from the 3′ end of the transcribed region (12).

Lewis (10, 11) has proposed a model in which individual BX-C functions would be expressed in overlapping domains

```
                                                                                                    TTAGTTCGAA GCACGTGCGT GTGAGTTCCG CATGAACGGT -101
ACACCAACAC TTTCGAACAG AAACCCCTCA CAGATACTCC TCACACAGAT ACTCTACTGT ACTGGGCTAC GATATCTGTG AGATACGCAC AGATACCGGC  -1
+1                                                                                                  > cDNA7-2
AGGCCGCCGG CGTCAAAGTC GACGTGCTGC CGGCGTTTGA GAGACCGAGC GGCGTGCAAG AAATCGTCGG GCAACGGCTT CAAATGAACG CAAAACGTTG +100
                                                                                  > cDNA-c15                    +
GAAAGCAAAC AACCGGCAAA AGCCATCAGC AGCTAAAAAT AAAACTGCAA CTATCAAAGT GAATCTGTAG AGAAAGAGTG ATAGCTCAGT GTGCCCAGTG +200
                                                                                  cag attcactt

TTTATTTAAT TACTTATTTA AATTTGATTT TCAAAATAAG TTGTGATTTT GGTTCGTTTG TTTGTTTCCC CATTAAATCG CATCGCGGAA CTGGTTTGCA +300
TTTTGTATCT TTTATCGCTG CAGTTCGTTT CTGTGTGGTT TCATGTATTC CCCATCTATT CCATCTCTTT TTGCAGATAC AAGTGAATGA AACCCTAACG +400
TTTAAACGCC ACCATAAACA CCAAATAACA AATTTTCTGG CAAACATGGC ATGTCACGGT CCCATTTAAC GTGGAAAATC GGCCCAGAAA TTCAACACGA +500
CGACAACTTT TTCGACTGTC CCAGAAAGCA AATATGTAAA GTGTATTTGT GTGCGTGCCA TTGACATTGA CTTATGGTCA TGAAGGGGTT AGGGGGACGG +600
GTCTAAGGGT GGGGGTTGAT GGCTAAAGGG GGGTCTCGCC GCTACGACAA AAACTTGGAT TCTCTCGTTT CTTTTTTGTC TTTCGTCTTC AACGCACATT +700
TATGGTTTGT TTTGCTCACA CACACACTCA AAAAGCCTAC GCAGCGGCAG AGAGGGCACA AGCAGCAGAG AAAGAGACGA AACATGTCTC GGATGATGAT +800
AAACGCCTCT CGTCTATTTT GGCGGGTCCA TGAATGAAAC AATTGCCTCG AATGACTGAC TGCCAGACAG ATAGCTAGAA ACTGCCCTTA CGGCGGGCAA +900
ATTGGGGTAT TTTGGGTTTC GGTGACTTTG AGGGACCGGG ACCTGAGGGC CCGGAAAGCC GACAGCAATA AACACTGTGC TCAGACAAAA GTCTGATGAG +1000
ACAGACCATC ATGATGCTCA TCATCATCGG CGATAATAAT GATTTTCATT TGCACATTCC ACCCGTTTCG CTGTTGCCGG CTTCCAGCAT TCAGATTCAG +1100
TTTGCATTCA GTGGGGCAGG TTATTTATGG CACTTTTCCT ATCGCCACTC ACTGGGGTAA TCGATTTTTA TTGGATTAAT ATTGCTTTCC GCATGATTTG +1200
                                                a                   GTAAGTCT...intron...TTCCAG
GCCTTCTTTT CGTTATCACT TAGCAGGTTT CTTGACGCTG AGGAGTGGAC AATGACTAAA ACTCAGTCAT CTGGAAAAGC TGTGACCAAGG ACATCCATAT +1300
                                                g
ACAAGTGCTG GCTAAACGTC TGTGGTTATC GCAATCATTT ATGGCGGAAT AGAAGGACAC GTCTTATGTG TCGATTTTCG GGCCCCATCC CTGTTGAAAA +1400
ATCAGCCGGA CAGCAGCAAC CACCGAACAC AAGACACAGG ACGACCGGAA GTGAGGAAAA CGCTTGTAAA ATCGGTGCAC ACGAGTGCGA GCGAGACACG +1500
AGCGAACGGG CCGAAAGGAA ACGACAATTA CGAGGACACG CCGGCTTGCT TTTCTGCTTT TCCTGTCTGC TCCTTTCAGT TGCGGCCGTG CAATCGTCGG +1600
GTCAACAACG CCGCGTGTCC GATTTCAATT TACAAATTAG AATATACATC TTTACATCTG GATTTGGTTT CATCTTTATC TGCATTTGGA TCCGCGACGA +1700
                                                aaaaa
CGCATAATTT TCAGCACTTT TGAATAATGG CGATTCTGTG CGACCTCGTA AACGTATACT GAATGTATCC TGAGTGTATC CTATCCGGTA TACCTTCAGT +1800
ATACGTAACA CGAGAGAGCA CTAGCAACGT GGGATGTGAA ACACTTTTTA AGACACCTTA ATACACGTAG CACTTTGTTA TATGGTCTTA CACTTTTTAA +1900
ATATCTCGAC TTATAAATAT ATAATTTAAA TATAAATACT TTAATAATGC GATATAAGTT AAATACTTCG ATTTACTGGC TGGAAGTTCA T          +2000
                                                aaaaa
```

FIG. 4. Complete sequence of the 1.7/2.0-kb *iab* exons. Exons are shown in upper case; nucleotide differences between the genomic and cDNA sequences are shown in lower case. The T to A change at +1153 is found in cDNA c15 but not in cDNA 7-2. The first (+1) nucleotide, as determined by primer extension analysis and S1 protection studies, is shown in bold face. The 5′ ends of cDNA c15 and cDNA 7-2 are underlined. The first eight bases and the last six bases of the intron are shown as an insertion at the splice junction. Consensus polyadenylylation sequences are underlined in the last line of sequence. The sites of polyadenylylation are located at +1662 and +1963.

and the combinatorial signals would determine unique segmental identities. According to this model, the *iab-3*$^+$ function is expressed from A3 to A8 and the *iab-4*$^+$ function from A4 to A8. Here, we have shown that the 1.7/2.0-kb *iab* RNAs are expressed in the primordia of A3 to A9 from the time of cellular blastoderm formation through germ-band shortening. Therefore, although the RNAs are transcribed from the *iab-4* region, their spatial localization correlates more closely with the domain of *iab-3* function(s).

Recently, Sánchez-Herrero and Akam (22) have shown that the cis-regulatory regions between the *abd-A* and *Abd-B* genes are transcribed in three overlapping domains. The 1.7/2.0-kb *iab* RNAs described here appear to belong to their domain I RNAs, which are transcribed from the *iab-3* and *iab-4* regions. They have suggested that transcription of the domain I RNAs may reflect a regional activation of the *abd-A* domain during early embryogenesis. However, developmental Northern analysis has shown that the temporal patterns of expression of the 1.7/2.0-kb *iab* RNAs and the RNAs detected by probes from the *iab-3* region are different (S.S., unpublished results). Thus the RNAs transcribed from the *iab-3* and *iab-4* regions appear to be a response to different developmental cues rather than a reflection of a generalized activation of the domain.

Apparently, transcription through the large cis-regulatory regions is a general feature of the entire BX-C. Genetic and molecular characterization of the 1.7/2.0-kb *iab* RNAs as well as other *iab* RNAs (refs. 3, 4, and 22; S.S., unpublished data) suggest that they share a striking resemblance to the early *bxd* RNAs described by Lipshitz *et al.* (7). Neither the

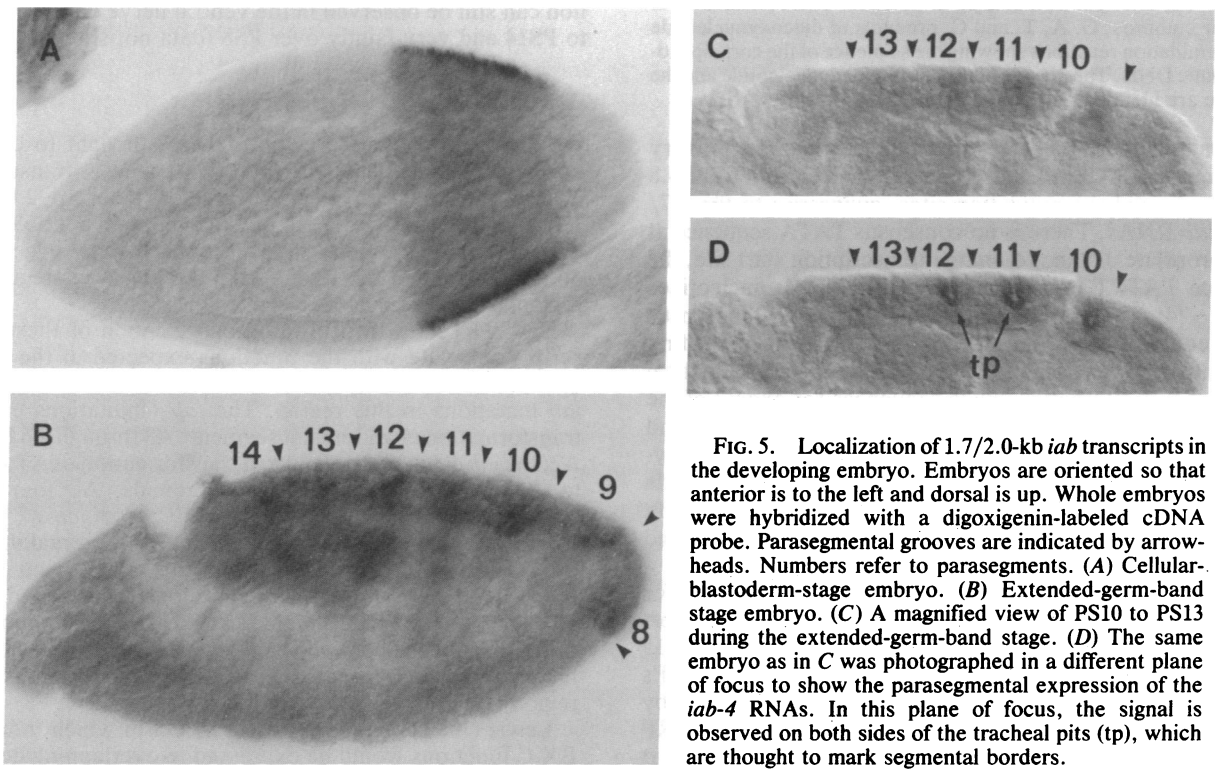

FIG. 5. Localization of 1.7/2.0-kb *iab* transcripts in the developing embryo. Embryos are oriented so that anterior is to the left and dorsal is up. Whole embryos were hybridized with a digoxigenin-labeled cDNA probe. Parasegmental grooves are indicated by arrowheads. Numbers refer to parasegments. (*A*) Cellular-blastoderm-stage embryo. (*B*) Extended-germ-band stage embryo. (*C*) A magnified view of PS10 to PS13 during the extended-germ-band stage. (*D*) The same embryo as in *C* was photographed in a different plane of focus to show the parasegmental expression of the *iab-4* RNAs. In this plane of focus, the signal is observed on both sides of the tracheal pits (tp), which are thought to mark segmental borders.

1.7/2.0-kb *iab* RNAs nor the early *bxd* RNAs appear to code for proteins. Hogness *et al.* (6) and Lipshitz *et al.* (7) have proposed three possible explanations for the presence of the early *bxd* transcripts: (*i*) the RNAs may provide trans-regulatory functions, (*ii*) transcription itself may be necessary for appropriate cis-regulation of the *Ubx* unit, or (*iii*) the chromatin state of the cis-regulatory region may allow fortuitous transcription from this region.

While the existence of independent trans-acting functions has not been disproven, the temporal and spatial expression of the *bxd* transcripts does not support such models (6). Hogness *et al.* (6) have also argued against the second possibility, whereby transcription through the *bxd/pbx* region would regulate cis-acting elements, thus indirectly regulating *Ubx*. They note that the morphogenetic function (*pbx*) assigned to this region is required in PS6, yet *bxd* transcripts are expressed at very low levels in PS6 but at high levels in PS7 and more posterior parasegments (6, 8, 23).

Given that specific RNAs are transcribed from cis-regulatory regions in each of the three bithorax domains, it may be worthwhile to reconsider the possibility that transcription alters or regulates the activity of neighboring enhancer-like sequences. In the simplest case, transcription through the cis-regulatory regions during early embryogenesis could inhibit the activity of some nearby enhancer/repressor elements simply by occluding the interaction of other trans-acting factors; thus trans-regulation of these elements would not occur until later in embryogenesis. This "enhancer occlusion" model would be similar to transcriptional regulation by promoter occlusion, as described in both prokaryotes (24) and eukaryotes (25). *bxd* and *iab* transcription would indirectly cis-regulate *Ubx* and *abd-A* transcription. In the case of the *Ubx* domain, many PS6 cells do not express *bxd* RNAs, or express them at very low levels. In these cells, the *pbx* enhancer elements would be active, whereas in PS7–13 cells and some PS6 cells, *bxd* transcription would shut off the enhancer elements. Thus *pbx* enhancer elements would be active in a unique mosaic pattern in PS6, thereby producing *pbx* function. What about regulation of the *iab-4* enhancer elements? If we assume that the mosaic pattern of *iab* expression in A4 is different than in other abdominal segments, then $iab\text{-}4^+$ function would be expressed in a unique mosaic pattern, thus defining the segmental identity of A4. Our *in situ* analysis indicates that the pattern of *abd-A* transcription in $iab\text{-}4^-$ embryos is very similar to the pattern observed in wild-type embryos (S.C. and S.S., unpublished results). Apparently, deleting the *iab-4* region causes only subtle alteration in *abd-A* transcription in A2–7. These subtle differences in *abd-A* expression could be influenced by transcription of the *iab* RNAs.

To critically evaluate this model, it is necessary to know when, and in which cells, the bxd^+ or $iab\text{-}4^+$ functions are required. In the case of *abd-A*, we do not know when during embryogenesis most *iab* functions are required; but at least one, the formation of the gonadal primordia, occurs shortly after the peak of 1.7/2.0-kb *iab* transcription (S.C., Y. Szabad, and S.S., unpublished results). Analysis of the model is also complicated because many of the mutations in the *pbx/bxd* and *iab-3/iab-4* regions are deletions or chromosomal rearrangements that affect both cis-regulatory elements and the *bxd* and *iab* promoters. For example, pbx^1 flies, in which a 17-kb region including both *pbx* DNA and the *bxd* promoter has been deleted, show a mutant phenotype only in PS6. However, deletion of the *pbx* regulatory element may be phenotypically dominant over the loss of *bxd* transcription in PS7–13, even if *bxd* transcription is normally required to repress pbx^+ function in PS7–13.

A third possibility discussed by Hogness *et al.* (6) is that transcription through *bxd*, *iab-4*, and other *iab* regions reflects a change in localized chromatin structure (7). For example, as enhancer-like elements in the *iab-3/iab-4* region become activated, the chromatin state may be altered, fortuitously activating the 1.7/2.0-kb *iab* promoter. This model does not propose a function for either the transcripts or transcription *per se*; rather, the temporal and spatial patterns of expression may reflect the complex regulation that occurs in this region. At least in the case of the *Ubx* domain, both *Ubx* and *bxd* transcription appear to be co-regulated by common upstream factors (8). This co-regulation could occur by a general activation of the chromatin domain; alternatively, these factors could specifically regulate both *bxd* and *Ubx* transcription, thus changing the spatial domains in which the *pbx* enhancers are occluded.

If transcription of the *iab* and *bxd* RNAs prior to the expression of the *Ubx*, *abd-A*, or *Abd-B* transcripts (22, 23) mirrors the initial activation state of the cis-regulatory regions, then cessation of 1.7/2.0-kb *iab* and early *bxd* transcription during midembryogenesis, while *Ubx* and *abd-A* continue to be transcribed, may also indicate alterations in the chromatin state. Based on the temporal profile of *bxd* expression, Lipshitz *et al.* (7) have suggested that there may be both an early and a late period of regulation of *Ubx* transcription. Our results, and those of Sánchez-Herrero and Akam (22), indicate that this model could be extended to both the *abd-A* and *Abd-B* domains as well.

We thank Chris Rushlow for help with the *in situ* hybridizations and Lynn Boyd for assistance in DNA sequencing. We thank Michael Akam, Ernesto Sánchez-Herrero, and Howard Lipshitz for helpful discussions; Sarah Smolik-Utlaut for sharing unpublished results of experiments carried out in E. B. Lewis's laboratory; and Anne Boulet, Tulle Hazelrigg, Carl Thummel, and Sarah Smolik-Utlaut for critical reading of the manuscript. S.S. thanks David Hogness, in whose lab a portion of this work was carried out.

1. Duncan, I. (1987) *Annu. Rev. Genet.* **21,** 285–319.
2. Peifer, M., Karch, F. & Bender, W. (1987) *Genes Dev.* **1,** 891–898.
3. Sánchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. (1985) *Nature (London)* **313,** 108–113.
4. Tiong, S., Bone, L. M. & Whittle, J. R. S. (1985) *Mol. Gen. Genet.* **200,** 335–342.
5. Bender, W., Weiffenbach, B., Karch, F. & Peiffer, M. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 173–180.
6. Hogness, D. S., Lipshitz, H. D., Beachy, P. A., Peattie, D. A., Saint, R. B., Goldschmidt-Clermont, M., Harte, P. J., Gavis, E. R. & Helfand, S. L. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 181–194.
7. Lipshitz, H. D., Peattie, D. A. & Hogness, D. S. (1987) *Genes Dev.* **1,** 307–322.
8. Irish, V. F., Martinez-Arias, A. & Akam, M. E. (1989) *EMBO J.* **8,** 1527–1537.
9. Rowe, A. & Akam, M. (1988) *EMBO J.* **7,** 1107–1114.
10. Lewis, E. B. (1978) *Nature (London)* **276,** 565–570.
11. Lewis, E. B. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 155–164.
12. Karch, F., Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M. & Lewis, E. B. (1985) *Cell* **43,** 81–96.
13. Kornfeld, K., Saint, R. B., Beachy, P. A., Harte, P. J., Peattie, D. A. & Hogness, D. S. (1989) *Genes Dev.* **3,** 243–258.
14. Brown, N. M. & Kafatos, F. C. (1988) *J. Mol. Biol.* **203,** 425–437.
15. Zavortink, M. & Sakonju, S. (1989) *Genes Dev.* **3,** 1969–1981.
16. Sharp, P. A., Berk, A. J. & Berget, A. (1980) *Methods Enzymol.* **65,** 751–768.
17. Tautz, D. & Pfeifle, C. (1989) *Chromosoma* **98,** 81–85.
18. Schneuwly, S., Kuroiwa, A., Baumgartner, P. & Gehring, W. J. (1987) *EMBO J.* **5,** 733–739.
19. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
20. Staden, R. (1984) *Nucleic Acids Res.* **12,** 551–567.
21. Campos-Ortega, J. A. & Hartenstein, V. (1985) *The Embryonic Development of Drosophila melanogaster* (Springer, Berlin).
22. Sánchez-Herrero, E. & Akam, M. (1989) *Development* **107,** 321–330.
23. Akam, M. E., Martinez-Arias, A., Weinzierl, R. & Wilde, C. D. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 195–200.
24. Adhya, S. & Gottesman, M. (1982) *Cell* **29,** 939–944.
25. Corbin, V. & Maniatis, T. (1989) *Nature (London)* **337,** 279–282.