The bithorax complex is regulated by trithorax earlier during Drosophila embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early trithorax transcripts

Yuri Sedkov*, Sergei Tillib*, Lev Mizrokhi and Alexander Mazo†

Department of Microbiology and Immunology and Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, PA 19107, USA

*These authors contributed equally to the work
†Author for correspondence

SUMMARY

The trithorax gene is required throughout development to maintain expression of homeotic genes of the bithorax and Antennapedia complexes. We determined complete structures of maternal and zygotic alternatively spliced trithorax transcripts, and found that two RNA isoforms are expressed in a surprising manner in the early embryo. At syncytial blastoderm their expression is confined to the ventral region fated to become mesoderm. An additional broad domain of trithorax expression arises during cellularization and is quickly resolved into four pair-rule-like stripes in the posterior half of the embryo. This early expression pattern suggested the possibility that trithorax is involved in the very early regulation of homeotic genes expressed only in the posterior region of the embryo. Indeed, transcription of bithorax complex genes in the mesoderm and ectoderm is altered in strong trithorax mutants during germ band elongation, while the anteriorly expressed Antennapedia complex genes are affected only at late stages of embryonic development. In addition, in another mutant allele (trxE3), expression of bithorax complex genes is normal, while expression of Antennapedia complex genes is reduced. These results suggest that proper expression of genes in the two homeotic complexes is maintained by products of different trithorax RNAs at different times of embryogenesis.

Key words: Drosophila, trithorax, mRNA structure, homeotic gene regulation

INTRODUCTION

Differential expression of the selector homeotic genes of the bithorax (BX-C) and Antennapedia (ANT-C) complexes determine developmental pathways along the body axis of the Drosophila embryo (Scott and Carroll, 1987; Ingham, 1988; Kaufman et al., 1990). Mutations in members of this group of genes result in transformation of one body segment into a different one. Expression of homeotic genes is confined to unique domains each corresponding precisely to the region transformed in mutant embryos, and are maintained throughout development (reviewed by Akam, 1987). The BX-C genes Ultrabithorax (Ubx), abdominal-A (abd-A), and Abdominal-B (Abd-B) are expressed in a region posterior to the second thoracic segment and extending through the eighth abdominal segment into part of the telson, while genes of the ANT-C are mainly expressed anterior to the third thoracic segment. Expression of homeotic genes is initiated at the blastoderm stage and relies on positional information provided by segmentation gene products (reviewed in McGinnis and Krumlauf, 1992). Although many segmentation genes are only transiently expressed, their products fading away by the completion of gastrulation, the homeotic genes maintain their initially established patterns of expression through later development, both among and within segments. Therefore, additional mechanisms and factors are required to preserve the homeotic gene expression patterns initiated at blastoderm.

This complicated function can be subdivided into two parts: prevention of ectopic homeotic gene expression and the maintenance of expression where it is appropriate. Posterior repression of each homeotic gene outside its normal expression domain is mediated by cross-regulatory interactions among the homeotic genes themselves (Struhl, 1982; Hafen et al., 1984; Harding et al., 1985), while anterior boundaries are maintained by the products of the Polycomb group (Pc-G) genes. Pc-G genes are not required to establish the initial patterns of expression, but in the absence (or haplo-insufficiency) of one or more members of the group, strong anterior toward posterior segmental transformations occur, resulting from anterior expansion of homeotic gene expression (reviewed by Paro, 1990). Protein products of several members of the Pc-G bind to multiple sites on salivary gland chromosomes, and two of them appear to be constituents of a multimeric protein complex (Zink et al., 1991; DeCamillis et al., 1992; Rastelli et al., 1993;
Mutations in ***trithorax*** (**trx**) mimic the phenotypes of loss-of-function mutations of multiple ANT-C and BX-C genes (**Lewis, 1968; Ingham and Whittle, 1980; Castelli-Gair and Garcia-Bellido, 1990**). Accordingly, examination of the expression of homeotic genes in stage 15-17 **trx** mutant embryos revealed reduced levels of proteins expressed by all genes of the BX-C (**Mazo et al., 1990; Breen and Harte, 1993**), and of the Deformed (**Dfd**), **Sex combs reduced** (**Scr**), and **Antennapedia** (**Antp**) genes of the ANT-C (**Breen and Harte, 1993**). Moreover, each of these genes exhibits a complex pattern of tissue-, parasegment- and promoter-specific requirements for **trx**. Clonal analysis revealed that **trx** function is required continuously throughout development (**Ingham, 1981, 1985**). The requirement for **trx** function exists even at very early stages of embryonic development (0-4 hours) as was demonstrated by temperature-shift experiments; moreover, the lack of **trx** product at this stage cannot be redeemed by its presence later on (**Ingham and Whittle, 1980**). These results imply that **trx** is required during the early phase of homeotic gene expression.

Previous studies have shown that the **trx** gene encodes several 10-15 kb mRNAs with different developmental profiles (**Mozer and Dawid, 1989; Breen and Harte, 1991**). A set of overlapping cDNAs corresponding to this region was obtained (**Mozer and Dawid, 1989**) and sequenced, revealing the existence of a huge open reading frame (ORF) of 3759 amino acids (**Mazo et al., 1990**). The predicted protein contains several zinc-finger–like structures, and bacterial **trx**-*lacZ* fusion proteins bind zinc in vitro, indicating that **trx** might be a nucleic acid-binding protein. Several other genes have been cloned that encode products with homologies to **trx**.

We report here a detailed analysis of the structure and expression of alternatively spliced **trx** transcripts. In the course of this work, we discovered an unexpected posteriorly restricted pattern of two **trx** RNAs during cellularization. Our results suggest that existence of this unique expression domain is likely to be associated with a distinct **trx** function required for proper expression of only BX-C genes in the mesoderm and ectoderm as early as at embryonic stages 10-11. Expression of the anteriorly expressed homeotic genes **Scr** and **Antp** was unaffected at this stage. We further show that normal levels of ANT-C expression require another **trx** function in the end of embryogenesis, and that this function is possibly provided by the products of the late embryonic **trx** transcript.

**MATERIALS AND METHODS**

**Northern blot analysis**

Total RNA was extracted from embryos (staged as indicated in Fig. 3), larvae, pupae and adults by the guanidinium thiocyanate/phenol procedure. Poly(A)* RNA was fractionated by electrophoresis through 0.8% formaldehyde agarose gel at 4 V/cm for 14-18 hours and transferred by capillary blotting to a nylona membrane (Magna, MSI). RNA isolation and hybridization was performed as described in Mozer and Dawid (**1989**). Radioautographs were scanned on the ULTROSCAN XL Laser Densitometer (LKB).

**Determination of the 5′ ends of **trx** RNAs**

RNase protection and primer extension assays were performed with 1 µg of poly(A)* RNA as described in Current Protocols in Molecular Biology (**Ausubel et al., 1993**). For RNase protection RNA probes were synthesized from a pBluescript plasmid containing the XhoI-D fragment of the genomic DNA from +2932 to –302 (1 is the position of the first nucleotide of common exon 1). An oligonucleotide corresponding to nt 100-132 of exon 1 (Fig. 1) was used in primer extension experiments and for sequencing of genomic DNA.

**RT-PCR experiments and sequencing**

Oligonucleotides specific for ends of exons 1-4 were used in the analysis of **trx** mRNA by the reverse transcriptase (RT)-PCR method as described in Current Protocols in Molecular Biology (**Ausubel et al., 1993**). DNA fragments obtained after PCR amplification were sequenced directly by a PCR-based Cycle Sequencing System according to protocol of GIBCO BRL for dsDNA.

**In situ hybridization**

RNA in situ hybridization on whole-mount embryos using digoxi-
genin-labeled antisense RNA probes was performed according to a modification of the protocol of Tautz and Pf vie (1989). 0-18 hour embryos from trxB1, red, eTM6, trxE3/ TM3 and Oregon-R stocks were collected, fixed in formaldehyde and prepared for hybridization. Generally 1 µg of plasmid DNA was used to synthesize antisense RNA probes with the Genius 4 kit (Boehringer-Mannheim). Spatially restricted staining was seen with all trx-specific probes in dilutions 1:10,000-1:40,000. The following cDNA sequences were synthesized by PCR and subcloned into the Bluescript KSII+ vectors (Stratagene). Five forms of trx mRNA, generated by alternative splicing and

RESULTS

The 5' and 3' termini of the trx transcription unit

We undertook a detailed structural analysis of trx RNAs the results of which are summarized in Fig. 1. Probes derived from the first exon detected all three major 10, 12 and 15 kb trx RNAs (not shown) which had been identified previously (Mozer and Dawid, 1989; Breen and Harte, 1991). This indicates that trx transcripts share the same 5' exon. No trx transcripts were detected in northern blot experiments using probes covering a 3 kb region upstream of the first exon. The 5' end of trx mRNAs was determined by a combination of RNase protection and primer extension analyses. A 3 kb Xbal-NdeI fragment (Fig. 1) including a part of the first exon was used to synthesize a single-stranded RNA probe for RNase protection experiments. These detected a single protection product of ~300 nt, roughly corresponding to the 5' terminus of the first exon (Fig. 2A). The same result was obtained when RNA isolated from differently staged embryos was used (stages as in Fig. 3A, not shown). To confirm that this site represents the true 5' end of trx mRNA, primer extension experiments were performed, which showed that transcription commences at a single nucleotide (Fig. 2B) at position 1 of the previously published sequence (Mazo et al., 1990). This position coincides with the 5' ends of the two 5'-most trx cDNAs obtained in a previous study (Mozer and Dawid, 1989).

To identify 3' termini of trx mRNAs, we used 3' RACE-PCR technique (Frohman, 1990). Prior northern hybridizations suggested that the 3' end of both the 12 and 14 kb mRNAs is within a 2.2 kb Xbal-XhoI fragment (Fig. 1) downstream of the previously sequenced cDNA (Moser and Dawid, 1989; Mazo et al., 1990). We sequenced this DNA fragment and designed several specific oligonucleotide primers for PCR experiments. Sequencing of several PCR products detected in these experiments (not shown) revealed two 3' termini. The proximal 3' end was found to be at position 12,275 of the previously published cDNA sequence (Mazo et al., 1990), leaving 154 nt of 3' noncoding sequence, while the distal 3' end was 1578 nt further downstream, suggesting that some trx mRNAs have a 1732 nt 3' untranslated region. Those 3' ends terminate 16 and 32 nt, respectively, downstream of a consensus polyadenylation signal AATAAA.

Alternative splicing yields five mRNA forms

Five forms of trx mRNA, generated by alternative splicing and

Fig. 1. Molecular structure of trx RNAs. Top line represents trx genomic DNA; Xb, Xbal; E, EcoRI; N, NdeI; H, HindIII; X, XhoI. trxB1 and trxE3 are deletions in the trx ORF which were sequenced previously (Mazo et al., 1990). trxB1 is a frameshift deletion after amino acid 658 (17.5% of the protein length), and is, therefore, a probable null allele. trxE3 is an in frame deletion of amino acids 2164-2443. Arrow under the top line indicates the common transcription start site detected by RNase protection and primer extension as described in Fig. 2. (A)n indicates two polyadenylation sites detected in PCR RACE experiments (see Materials and methods). Five trx RNAs were detected and analyzed in this study: M, 10,941 nt; ME, 12,519 nt; E1, 13,756 nt; E2, 12,968 nt; L, 14,205 nt. Positions of exons in the cDNA sequence and the lengths of RNAs are estimated based on results of the previous cDNA sequencing (Mazo et al., 1990) and on sequencing of the RT-PCR products obtained in this work. Initiation and termination codons in each RNA are indicated; microintrons are shown by vertical lines; cDNA probes used for northern blot and in situ hybridizations are shown on the bottom. Diagrams of embryo staining patterns, described in the text, appear to the left.
encoding distinct protein isoforms, have been identified in this study. The alternate splicing routes of *trx* mRNAs were determined from a combination of PCR amplification, RNase protection, and northern analysis. cDNAs were obtained that correspond to the *Xbal-NdeI* fragment in Fig. 1 (*NdeI* is at position 302 of the cDNA sequenced previously, Mazo et al., 1990). The lane marked ‘M’ contains end-labeled *MspI* restriction fragments from pBR322 as size markers; the lane marked ‘−’ is a negative control containing 20 µg of yeast tRNA, and the lane marked ‘+’ contains 1 µg of wild-type embryonic (Oregon-R, 2-17 hours) poly(A)*+* RNA. The protected product has an estimated size of 300 nt. The same size protected products were obtained when other RNA preparations from embryos staged as in Fig. 3 were used.

B. Primer extension analysis of the *trx* 5′ terminus. An oligonucleotide primer 33 nt in length (positions 100-132 in the cDNA sequence, Mazo et al., 1990) was annealed to 20 µg of yeast tRNA (lane ‘−’). To 1 µg of poly(A)*+* wild-type 2-4 hours embryonic RNA (‘+’). The lanes C, T, A and G are sequence data obtained with the oligonucleotide and subclone used to generate the primer, and yield directly the initiation site sequence coinciding with the first nucleotide of the cDNA sequenced previously (Mazo et al., 1990). No products were detected in experiments when oligonucleotides specific to exons 2, 3 and 4 in Fig. 1 were used.

![Fig. 2](image)

**Fig. 2** Identification of the 5′ terminus of the *trx* transcription unit. A, RNase protection for the 5′ end of exon 1. The probe is the RNA complementary to the *Xbal-NdeI* fragment in Fig. 1 (*NdeI* is at position 302 of the cDNA sequenced previously, Mazo et al., 1990). The lane marked ‘M’ contains end-labeled *MspI* restriction fragments from pBR322 as size markers; the lane marked ‘−’ is a negative control containing 20 µg of yeast tRNA, and the lane marked ‘+’ contains 1 µg of wild-type embryonic (Oregon-R, 2-17 hours) poly(A)*+* RNA. The protected product has an estimated size of 300 nt. The same size protected products were obtained when other RNA preparations from embryos staged as in Fig. 3 were used. B, Primer extension analysis of the *trx* 5′ terminus. An oligonucleotide primer 33 nt in length (positions 100-132 in the cDNA sequence, Mazo et al., 1990) was annealed to 20 µg of yeast tRNA (lane ‘−’). To 1 µg of poly(A)*+* wild-type 2-4 hours embryonic RNA (‘+’). The lanes C, T, A and G are sequence data obtained with the oligonucleotide and subclone used to generate the primer, and yield directly the initiation site sequence coinciding with the first nucleotide of the cDNA sequenced previously (Mazo et al., 1990). No products were detected in experiments when oligonucleotides specific to exons 2, 3 and 4 in Fig. 1 were used.

In our previous study, we reported finding of a 70 nt microintron at position 3104 (Mazo et al., 1990). Sequencing of genomic DNA in this region revealed an additional microtron of 64 nt 207 nt downstream of the first (Fig. 1). We have also found three 60-77 nt microintrons near the C-terminal end of the *trx* ORF (nt 11538-11948). It seemed possible that some RNA species contain one or more of these microintrons. Another possibility was that splice sites of the first two microintrons are employed for junction with each other or with exons 2 and 3, which would produce mRNA forms lacking 5′ parts of exon 4. To test for possible alternative splicing combinations, multiple PCR-based experiments were performed. We could not find evidence for the existence of mRNAs with any of these potential structures. Therefore, we conclude that the five microintrons are spliced out of all *trx* mRNAs as shown in Fig. 1.

A great deal of effort was devoted to testing the possibility of splicing within the long ORF encoded mainly by the 10.5 kb exon 4 (Fig. 1). However, we were not able to find any evidence of alternative splicing in this region. Thus, the five structures presented in Fig. 1 appear to represent the predominant, if not the only, products of the *trx* locus.

**Developmental profile of alternatively spliced *trx* mRNAs**

Expression of the different RNA forms during development was analyzed by northern blot hybridization. Interestingly, in all experiments, we observed a very narrow peak of expression of the 12 kb RNA at 2-4 hours of embryonic development (Fig. 3, top). This result is consistent with our finding of strong patterned expression of *trx* RNA during cellularization (see below). The 14 kb RNA is also expressed transiently, in 11-15 hour embryos. It again becomes very abundant in early pupae and is the major species in adult males. The 14 kb band was seen with about equal intensity when probes specific to exons 2 and 3 were used (Fig. 3A, middle), suggesting that it corresponds to RNA L, which contains both these exons (Fig. 1). However, the 12 kb band seen at 2-4 hours was very weak with the exon 2 probe, suggesting that the RNA form designated E2 in Fig. 1 is a minor species. A stronger signal was seen at 12 kb with the probe specific to exon 3, although its ratio to the 14 kb signal at 11-15 hours was significantly lower than when a common probe was used. We conclude, therefore, that three different RNA forms, ME, E1 and E2, in the approximate proportions of 11:3:1 (see Materials and Methods), are expressed at 2-4 hours.

Two bands of 10 and 12 kb were detected only in adult females and in 0-1 hour embryos, suggesting that they are deposited maternally (Fig. 3A,B). Since we see a 12 kb RNA in 0-1 hour embryos, and since this signal is absent when probes specific to exons 2 and 3 are used, we conclude that RNA ME (Fig. 1) is both a maternal and an early zygotic species. RNA ME also appeared to be present in 11-15 hour embryos as a minor component (Fig. 3A). When a probe specific to the long 3′ untranslated region was used, all species except the 10 kb RNA were seen (Fig. 3A, bottom). Both the 12 and 14 kb major RNA species were detected in these experiments (Fig. 3A); essentially the same result was obtained when only the 3′-most 267 nt fragment was used as a probe (not shown). These results show that all *trx* mRNAs with the exception of a small maternal species designated M in Fig. 1 contain the additional ~1.6 kb of 3′ noncoding sequence. Predicted sizes of all *trx* RNA forms based on their exon composition are given in the legend to Fig. 1.
Regulation of bithorax by trithorax in Drosophila

To determine the expression pattern of the individual \(\text{trx} \) RNA forms, we used digoxigenin-labeled RNA probes, which were titrated to ensure the specificity of staining (see Materials and methods). Using these probes, we find an extremely interesting temporally and spatially restricted expression pattern. Zygotic \(\text{trx} \) expression is first apparent at early stage 4 of the syncytial blastoderm (Fig. 4C). It is limited to the posterior half of the embryo and is initially confined to the ventral region fated to become mesoderm; at the cellular blastoderm stage, this domain of \(\text{trx} \) expression is 18 cells wide (Fig. 4H). An additional pattern of \(\text{trx} \) expression arises at the beginning of cellularization (Fig. 4D). It first appears as a broad domain and is rapidly resolved into four pair-rule-like stripes in the posterior half of the embryo (Fig. 4G, H). Unlike fully refined pair-rule gene stripes, \(\text{trx} \) stripes 1, 2 and 4 are stronger then stripe 3. At cellular blastoderm, stripes 1, 2 and 4 are four cells wide, while stripe 3 is two cells wide. These stripes can be seen until early stage 8 (Fig. 4I). Double staining experiments with \(\text{ftz} \) antibodies demonstrated that \(\text{trx} \) stripes when completely resolved, coincide with the \(\text{ftz} \) stripes 3-6, which correspond to parasegments (PS) 6, 8, 10 and 12 (Fig. 4F). At the initiation of gastrulation when \(\text{ftz} \) stripes narrow to about three nuclei, the \(\text{trx} \) stripes do not. At cellular blastoderm the anterior boundary of \(\text{trx} \) expression only slightly extends beyond \(\text{ftz} \) stripe 3, corresponding to the posterior of PS5. The posterior boundary of the \(\text{trx} \) expression domain is between \(\text{ftz} \) stripes 6 and 7, i.e. in PS13. Early zygotic \(\text{trx} \) expression persists until the end of germ band elongation (late stage 8; Fig. 4J).

The different intensity and duration of two components of the \(\text{trx} \) pattern, the ventral domain and the pair-rule-like stripes, suggested that they might correspond to the patterns of two RNAs with different structures. Therefore, we examined the expression pattern of specific alternatively spliced exons. Staining with a probe specific to exon 3 is in general similar to that seen with the common probe (Fig. 4E). This suggests that mRNA E1 is expressed both in the stripes and in the presumptive mesoderm. There is no probe specific to mRNA ME, which is the predominant species at this stage. It is, however, likely that mRNA ME is expressed predominantly in the presumptive mesoderm, since staining with the exon 3 probe (i.e.
specific to E1 at this stage) appears substantially decreased in the ventral domain (Fig. 4E) when compared with staining using the common probe. We could not detect any restricted expression pattern with a probe specific to exon 2 in early embryos. Therefore, we propose that a minor RNA corresponding to structure E2 in Fig. 1A is uniformly distributed throughout the embryo.

The ‘late’ embryonic *trx* expression may be defined by the appearance of a new domain of expression during germ band retraction, and by appearance of the 14 kb RNA L on northerns (Fig. 3A). During germ band elongation, the most intense staining still corresponds to the early RNAs, being restricted to the posterior part of developing mesodermal tissues. As development proceeds, the last traces of the early pattern...
Regulation of bithorax by *trithorax* in *Drosophila*

become obscured by weak staining throughout the embryo, with slightly higher expression in the CNS, including the procephalic lobe neuroblasts. Later on in development, staining in the CNS intensifies; after segmentation is complete strong staining is seen in the supraoesophageal and suboesophageal ganglia and in the ventral nerve cord. Weak staining is also seen in the head, the clypeolabrum and the antennal-maxillary complex (Fig. 4K). Thus, following germ band retraction,
expression of trx becomes progressively less restricted in the course of development. The RNA isoform L is the main component of this trx expression pattern.

Expression of BX-C, but not ANT-C genes, is altered in trx mutants at the beginning of segmentation

During early stages of embryogenesis, trx RNAs are expressed predominantly within PS 6-13, which coincides with the major domains of expression of Ubx, abd-A and the α class transcripts produced from the P4 promoter of Abd-B. It was, therefore, interesting to examine whether expression of early trx RNAs is associated with a requirement for trx function for normal expression levels of only these BX-C genes. To test this, we analyzed expression of Ubx, abd-A, Antp, Scr and products of the P4 and P3 promoters of Abd-B in trx- embryos using in situ hybridization with specific RNA probes. In these experiments, we used the trx\textsuperscript{B11} allele, which is presumably a null allele since it is a frameshift deletion near the beginning of common exon 4 (Fig. 1). It would produce truncated proteins consisting of the N-terminal 7.7% (RNA isoforms ME and E2) and 17.5% (L and E1) of the predicted trx proteins.

We did not detect changes in the expression pattern of any of the tested genes in trx\textsuperscript{B11} embryos from cellular blastoderm through embryonic stage 9. This suggests that zygotic trx products are not essential for initiation of homeotic gene expression. However, at stage 10 and more clearly at early stage 11, expression of Ubx, abd-A, and the α class Abd-B RNAs is significantly altered. At this stage, boundaries of Ubx expression in the mesoderm of PS6-12 are in register with those in the ectoderm (Akam and Martinez-Arias, 1985). Although in trx embryos Ubx expression is reduced in all tissues, the reduction is greatest in the mesoderm (Fig. 5B). The consequence of the lack of trx is most prominent in the mesoderm and ectoderm of PS6, where Ubx is normally expressed at its highest level. Consistent with the low level of trx RNA in PS5, Ubx expression does not appear to be changed significantly in this region. Beginning from stages 10-11, expression of abd-A is also strongly reduced in trx embryos in all but the most posterior portion of its domain in the epidermis of PS7-13 (Fig. 5D). At the extended germ band stage, the α class transcript of Abd-B is normally expressed at low levels in the mesoderm of PS11, at higher levels in the mesoderm of PS12, and at highest levels in the mesoderm and ectoderm of PS13 (Fig. 5E, Kuziora and McGinnis, 1988b). In trx embryos, its expression is almost completely abolished in PS11 in the ectoderm and is significantly reduced in the rest of the domain (Fig. 5F). Promoter P3 of Abd-B produces β class transcripts which are expressed in the mesoderm and ectoderm of PS14/15 (Kuziora and McGinnis, 1988b). We did not find any significant changes in its expression level in trx embryos at any stage. We were also unable to detect any consistent changes in the expression of Scr and Antp RNAs in trx\textsuperscript{B11} embryos until embryonic stages 16-17.

Expression of Antp and Scr, but not the BX-C, is altered in trx\textsuperscript{E3} mutant embryos

The results of our experiments show that the requirement for trx function for normal levels of ANT-C transcription occurs much later than for BX-C transcription. This temporally distinct ANT-C-related trx function is likely to be supplied by protein products of the major late zygotic RNA (isoform L), which is expressed in a broad domain in 11-15 hour embryos, i.e. at embryonic stages 15-16 (Fig. 3A). Since the predicted protein products of this isoform are different from those of the major early transcript ME, this late function may require different protein domains from the early function.

To substantiate this idea, we have examined whether other trx alleles have differential effects on early versus late embryonic function. We have found that one allele (trx\textsuperscript{E3}), which is associated with a deletion in the C-terminal half of the ORF (Fig. 1), has such a differential effect. Previously we had shown that the distribution of the Ubx protein in the embryonic nervous system is not affected in trx\textsuperscript{E3} embryos (Mazo et al., 1990). In addition, in this work we had not detected changes in the distribution of either Ubx (Fig. 5G,H) versus abd-A, or both classes of Abd-B transcripts (β class shown in Fig. 5I,J) at any embryonic stage. In late homozygous trx\textsuperscript{B11} embryos, the first and third midgut constrictions are shifted posteriorly (Breen and Harte, 1993). We could not detect any substantial changes in gut morphology in trx\textsuperscript{E3} embryos. This is consistent with the fact that expression of the BX-C genes is not altered in these mutants. Therefore, trx\textsuperscript{E3} does not affect the early BX-C-related trx function.

We were also unable to detect any consistent changes in the expression of antennapedia (Antp) RNAs, which are expressed in the mesoderm and ectoderm of PS6-12, in register with those in the mesoderm (Akam and Martinez-Arias, 1985). Although in trx embryos Antp expression is reduced in all tissues, the reduction is greatest in the mesoderm (Fig. 5B). The consequence of the lack of trx is most prominent in the mesoderm and ectoderm of PS6, where Ubx is normally expressed at its highest level. Consistent with the low level of trx RNA in PS5, Ubx expression does not appear to be changed significantly in this region. Beginning from stages 10-11, expression of abd-A is also strongly reduced in trx embryos in all but the most posterior portion of its domain in the epidermis of PS7-13 (Fig. 5D). At the extended germ band stage, the α class transcript of Abd-B is normally expressed at low levels in the mesoderm of PS11, at higher levels in the mesoderm of PS12, and at highest levels in the mesoderm and ectoderm of PS13 (Fig. 5E, Kuziora and McGinnis, 1988b). In trx embryos, its expression is almost completely abolished in PS11 in the ectoderm and is significantly reduced in the rest of the domain (Fig. 5F). Promoter P3 of Abd-B produces β class transcripts which are expressed in the mesoderm and ectoderm of PS14/15 (Kuziora and McGinnis, 1988b). We did not find any significant changes in its expression level in trx embryos at any stage. We were also unable to detect any consistent changes in the expression of Scr and Antp RNAs in trx\textsuperscript{B11} embryos until embryonic stages 16-17.

DISCUSSION

Structural complexity of trx

Genetic and cytological studies have revealed a requirement for trx in the proper expression of multiple homeotic genes. The complex and diverse functions of trx are to some extent reflected in the gene structure described here. The trx transcription unit gives rise to five mRNAs by alternative splicing (Fig. 1). All trx mRNAs share a common start site (Fig. 2). Sequencing of exons 1 (Mazo et al., 1990) and 2 (this work) did not reveal any ORF that might be continued in exons 3 or 4. Differences in the coding regions of the zygotic trx mRNAs are, therefore, generated by the alternative usage of exon 3. The first AUG codon is located in the beginning of exon 3, and may be used for translational initiation in both the E1 and L mRNAs (Fig. 1). Translational initiation of mRNAs ME, E2 and E3 may occur at the first AUG located 21 bp from the beginning of exon 4. Both AUG codons are located within sequences optimal for translational initiation (Cavener and Ray, 1991). Sequence analysis of the coding region of the trx gene from D. viridis showed that these two AUG codons are conserved between the two species (S. T., unpublished data). Since there
are no sequence similarities between the N-terminal peptide sequence encoded by exon 3 with any other known protein, it is difficult to assess its significance. The corresponding portion of the human ALL-1/HRX protein contains ‘AT-hook’ motifs related to the DNA-binding domains of the HMG proteins (Tkachuk et al., 1992). Interestingly, alternative splicing also occurs in the N-terminal coding portion of the human homolog of trx (Domer et al., 1993).

We detected two sites of polyadenylation and showed that all zygotic trx mRNAs have long 3’ noncoding regions. In the 3’ portion of the longer trx noncoding region, we found two potentially important sequences. The first one consists of two repeats of ATTATA. This motif is found in mRNAs with a short half-life and is believed to play a direct role in regulating rapid mRNA turnover (Gillis and Malter, 1991). Indeed, two early mRNAs, ME and E1, are detected by both in situ hybridization and northern blot analysis during only the first few hours of embryogenesis. In addition, the major late zygotic RNA (L) shows a sharp peak of expression, suggesting that it turns over quite rapidly (Fig. 3A). Interestingly, three of these ATTATA motifs are found in the mouse and human homologs of trx, suggesting the functional importance of this sequence (Ma et al., 1993). Another element is identical in sequence to the cytoplasmic polyadenylation element (CPE) implicated in the developmental control of polyadenylation in Xenopus (Simon et al., 1992).

Expression patterns of zygotic trx RNAs suggest distinct functions in the early embryo

Major zygotic trx transcripts are expressed at two distinct periods of embryonic development. Two mRNAs, ME and E1, are expressed in a unique restricted pattern beginning from syncytial blastoderm. The pattern of early expression described here is consistent with previous genetic data. In experiments with a ts allele trx<sup>1</sup>, Ingham and Whittle (1980) demonstrated that for normal adult segment identities trx function is required after egg deposition, but within the first four hours of embryogenesis. The finding of patterned trx expression at 2-4 hours supports the existence of a distinct early zygotic function. The early trx pattern can also explain a paradox described previously. Ingham has demonstrated that, although animals that are homozygous for a trx null allele die before or upon hatching, the absence of maternal and/or zygotic trx only slightly affects the segmental differentiation of the late embryonic epidermis (Ingham, 1983). Therefore, trx protein has a vital function(s) in other tissues. The presence of a strong ventral stripe of expression corresponding to the presumptive mesoderm, contributed mainly by the more abundant species ME, suggests that trx function may be required in mesoderm-derived tissues.

Another interesting aspect of early trx expression is that its pair-rule-like domains imitate the early pattern of Ubx RNA. Ubx transcripts are first apparent at cellular blastoderm as a stripe corresponding to PS6 and shortly thereafter as a four-stripe pattern (Shermoen and O'Farrell, 1991) which strikingly resembles that of trx trx expression in the presumptive mesoderm and ectoderm overlaps with domains of expression of Antp, Ubx and abd-A. ABD-B1 is the only Abd-B protein isoform expressed in PS10-13 in the mesoderm and in PS13 in the ectoderm, and it is encoded by the α class mRNA transcribed from P4, one of the four Abd-B promoters (Casanova et al., 1986; Celniker et al., 1989, 1990). Thus, we suggest that at early stages of embryo development trx may be required for proper expression of the Antp, Ubx, abd-A and Abd-B (P4 promoter) genes in the mesoderm and ectoderm.

Early zygotic trx function is required for regulation of BX-C, but not ANT-C genes

Our analyses show that trx function is required to maintain transcription levels of Ubx, abd-A and Abd-B in PS6-13 beginning from embryonic stage 10 (4-5 hour embryos). Expression of all three genes is affected in both mesodermal and ectodermal tissues where early trx RNAs are expressed. Consistent with the absence of trx RNA in PS14-15, we did not detect any alterations in the expression of the β class Abd-B RNA, which is transcribed in this region from the P3 promoter. (Kuziora and McGinnis, 1988b; Celniker et al., 1990). Thus proteins derived from early zygotic trx RNAs are required to specifically mediate transcription from the P4 promoter of Abd-B. In both trx mutant alleles tested, there is no effect on the level of expression of early Scr and Antp transcripts. This is expected for Scr because its primary domain of expression lies outside of the region of expression of early trx RNAs (LeMotte et al., 1989). Although two Antp promoters produce rather strong levels of expression in the posterior half of the embryo, especially at stages 10-11 (Bermingham et al., 1990), we did not find any effect of trx mutation at these stages. This is consistent with that in late trx embryos; expression of Antp RNA is only slightly reduced in the abdominal parasegments (Breen and Harte, 1993, and Fig. 5J). Thus, there is no (or minimal) requirement for products of early trx transcripts for normal levels of both Antp promoters.

Our results demonstrate that early zygotic trx function is not required to initiate homeotic gene expression. However, absence of zygotic trx function affects BX-C transcription level at early stages of embryo development long before the major late zygotic trx RNA is expressed at its maximal level. The transcription pattern and time of expression of the early trx RNAs ME and E1 in ectoderm and presumptive mesoderm correlates well with the zygotic trx requirement for transcription of BX-C genes. Thus we suggest that products of the ME and E1 RNAs are required early to regulate BX-C transcription. trx function is not required, however, for normal expression of genes of the ANT-C and the β class Abd-B transcripts during early embryonic stages. It is not clear at present whether products of the minor RNA species E2 (Fig. 1) have any functional significance.

Regulation of the ANT-C requires a distinct protein domain not required for the BX-C, and possibly a distinct protein product encoded by late RNA L

The finding of a distinct BX-C-related early embryonic function of trx raises the question of which trx transcripts encode products required for proper expression of ANT-C genes in late embryos? Reduced transcript levels from these genes are detected only at embryonic stages 16-17 at the earliest, i.e. in 13-16 hour embryos. Results of northern experiments and in situ hybridization suggest that at these stages the major trx transcript is RNA L, and that it is expressed at its maximal level in 11-15 hour embryos in a broad region (Fig. 3A). We therefore conclude that trx function is required for proper expression of the ANT-C genes only at late stages of
embryogenesis, and that this function is most likely supplied by protein products from RNA L.

The analyses of homeotic gene transcription in the homoy-

genous lethal trxE3 allele show that this mutation does not lead to

misexpression of either the Ubx protein (Mazo et al., 1990)

or BX-C transcripts (this work) at any embryonic stage. In

contrast, we show here that expression of Antp and Scr are

changed. The trxE3 mutation is an in-frame deletion in the trx-
coding region. It deletes 271 amino acids including one pre-

sumptive Zn-finger domain (Mazo et al., 1990). We did not

find any changes in distribution of the trx RNAs in trxE3

embryos (not shown), i.e. this allele does not contain any

second site mutation affecting trx transcription. Since the

trxE3 mutation does not alter BX-C transcription, it is clear that

cannot contain the BX-C, the deleted portion of the Protein is not

required to maintain BX-C transcripts levels, at least during

embryo development. Although the protein domain deleted in

trxE3 is functionally important, this is not reflected in structural

differences between the early (ME, E1) and late (L) RNA iso-

forms (Fig. 1). Moreover, although the E1 and L RNAs are

presumably involved in regulation of different sets of genes,

coding capacity is identical (Fig. 1). The difference in

function at different times suggested by the temporally distinct

regulation of the BX-C and the ANT-C, combined with the dif-

ferential effect of the removal of the region deleted in trxE3

suggests that there may be important differences in the

mechanism of trx action at early versus late embryonic stages.

Such differences may be mediated by different post-transla-

tional modifications or protein processing which in turn may

be affected by the differential inclusion of the N-terminal

protein domains at early times. We are currently examining

these ideas by characterization of antibodies raised against
different trx protein domains.

An important question is whether late trx function is

required for proper expression of genes other than those of

the ANT-C. Our recent analysis showed that, in embryonic and

larval tissues, trx is also required to maintain transcription

levels of the region-specific homeotic gene fork head (fkh) (our

unpublished observations). At cellularization fkh is expressed

in anterior and terminal domains of the embryo (Weigel et al.,

1989) that do not overlap with the early trx domain of

expression. Similarly to our results with the ANT-C, we have

found that in several trx alleles, including trxE3 and trxB1,

expression of fkh RNA is decreased only at late embryonic

stages. Therefore, in addition to ANT-C genes, expression of

other trx target genes probably requires products of the late

zygotic RNA.

In conclusion, we suggest that multiple genetically defined

functions of the trx gene are performed by distinct products of

alternatively spliced trx transcripts. Surprisingly, the require-

ments for trx function are different for genes in the two

homeotic complexes. It is likely that the early embryonic BX-

C-related function is evolutionarily conserved, because we find

a similar posteriorly restricted transcription pattern of the trx

gene in the distantly related species D. virilis (Yu. S., unpub-

lished observations).

A. M. wish to thank Igor Dawid for his generous and stimulat-
ing support during his stay in NIH, where the analysis of trithorax was

initiated. We are extremely grateful to Jim Jaynes for helpful dis-

cussions and comments on the manuscript. We also thank Judy Kassis

for critical reading of the manuscript; Jim Jaynes and Brian Mozer for
cDNA clones.

REFERENCES

Akam, M. E. and Martinez-Arias, A. (1985). The distribution of

Ultrabithorax transcripts in Drosophila embryos. EMBO J. 4,

1689-1700.

Akam, M. E. (1987). The molecular basis for metameric pattern in


Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G.,


Biologte. New York: John Willey and Sons.


of the Drosophila homeotic gene Deformed. EMBO J. 9, 4287-4297.

Bermingham, J. R. Jr., Martinez-Arias, A., Petitt, M. G. and Scott, M. P.

(1990). Different patterns of transcription from the two Antennapedia


Drosophila visceral mesoderm from autoregulation and exclusion.

Nature 333, 576-578.


trithorax gene, a positive regulator of homeotic gene expression in


genes in the bithorax and Antennapedia complexes and exerts different

tissue-specific, parasegment-specific and promotor-specific effects on each.

Development 117, 119-134.


and trithorax with cis regulatory regions of Ultrabithorax during the

development of Drosophila melanogaster. EMBO J. 9, 4267-4275.


characterization of a parasegment specific regulatory element of the


sites. Nucl. Acids Res. 19, 3185-3192.

Celinka, S. E., Keelam, D. J. and Lewis, E. B. (1989). The molecular genetics

of the bithorax complex of Drosophila: characterization of the products of

the Abdominal-B domain. Genes Dev. 3, 1424-36.

Celinka, S. E., Sharma, S., Keelam, D. J. and Lewis, E. B. (1990). The

molecular genetics of the bithorax complex of Drosophila: cis-regulation in

the Abdominal-B domain. EMBO J. 9, 4277-86.


sites. Nucl. Acids Res. 19, 3185-3192.

Celinka, S. E., Keelam, D. J. and Lewis, E. B. (1990). The molecular genetics

of the bithorax complex of Drosophila: cis-regulation in

the Abdominal-B domain. EMBO J. 9, 4277-86.


sites. Nucl. Acids Res. 19, 3185-3192.


sites. Nucl. Acids Res. 19, 3185-3192.


sites. Nucl. Acids Res. 19, 3185-3192.


(Accepted 22 March 1994)