Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2112–2116, March 1990
Genetics

The trithorax gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains

(homeotic gene/Ultrabithorax gene/zinc finger)

ALEXANDER M. MAZO, DER-HWA HUANG, BRIAN A. MOZER*, AND IGOR B. DAWID

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892

Contributed by Igor B. Dawid, December 29, 1989

ABSTRACT **The trithorax (*trx*) gene functions in segment determination in *Drosophila* through interaction with genes of the bithorax complex and Antennapedia complex. Genetic evidence suggests that *trx* may be considered a positive regulator of homeotic genes. Sequencing of cDNAs corresponding to the entire *trx* transcription unit revealed the existence of an unusually long open reading frame encoding 3759 amino acids. The main features of the predicted *trx* protein are several cysteine-rich regions which can be folded into zinc finger-like domains. Cysteine-rich portions expressed from *trx* cDNAs in *Escherichia coli* are capable of zinc binding *in vitro*, suggesting a possible function for the *trx* product as a metal-dependent DNA-binding protein. Analysis of *trx* mutant embryos with antibody to the Ultrabithorax (*Ubx*) gene product showed decreased staining in parasegment 6 of the ventral nerve cord of late embryos. However, expression of *Ubx* was not affected in embryos carrying the lethal mutation trx^{E3}, in which one of the putative zinc finger-like domains of the *trx* protein is deleted. This differential effect of the *E3* mutation suggests that *trx* exhibits other function(s) besides those involved in the regulation of *Ubx* expression in the ventral nerve cord of the embryo.**

In the *Drosophila* embryo, homeotic genes including members of the bithorax complex (BX-C) and Antennapedia complex (ANT-C) directly control the unique pathways of differentiation followed by particular segments (1, 2). One of the central questions in understanding homeotic gene function concerns the mechanism by which segment-specific expression patterns are achieved. Genes of the gap and pair-rule classes are believed to define domains in which homeotic genes are activated during embryogenesis (3–6), and cross-regulatory interactions between BX-C and ANT-C genes are important in pattern maintenance by preventing ectopic expression of individual genes (7–9). In addition, regulatory genes have been identified by mutations which cause homeotic transformations in multiple segments. One such regulatory gene is trithorax (*trx*), originally named Regulator-of-bithorax (10); mutations in *trx* cause transformations of the first and third thoracic segments towards the second thoracic segment and transformations in the abdomen to a more anterior pattern (11). Since the *trx* effect is suppressed by extra copies of the BX-C and enhanced by a reduced copy number, it has been suggested that *trx* behaves genetically as an activator of the BX-C (12). However, appropriate expression of BX-C genes can be initiated in the embryo in the absence of *trx* function, and therefore it appears that *trx* is required mainly for the maintenance of BX-C activity (13). In addition to its interactions with the BX-C, *trx* also affects the function of the ANT-C (14, 15). A requirement for *trx* activity has been demonstrated in the imaginal cells that give rise to different tissues of the adult (14).

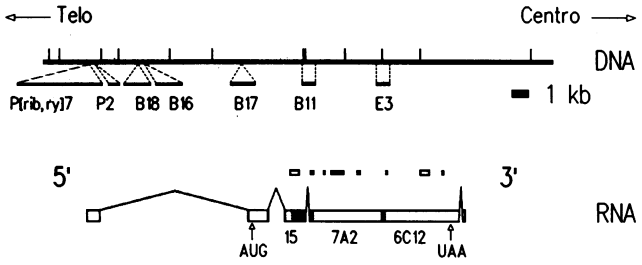

FIG. 1. Structure of the *trx* locus. (*Upper*) Genomic DNA from the *trx* locus with mapped mutations indicated: *P{rib, ry}7*, *P2*, *B18*, *B16*, and *B17* are insertions; *B11* and *E3* are deletions. The *B11* deletion leads to a frameshift and thus terminates the ORF, while *E3* is an in-frame deletion (see Fig. 2). Telo, telomeric; Centro, centromeric. (*Lower*) Composite structure of *trx* mRNA, based on cDNAs (15, 7A2, and 6C12) isolated previously (16) and sequenced in this work. Gaps between the cDNA clones were bridged by synthesis of DNA from poly(A)$^+$ RNA with oligonucleotide primers, followed by PCR amplification (refs. 21 and 22; stippled regions). The PCR products and the corresponding genomic region were sequenced to provide complete overlaps. The ORF is indicated by its AUG start and UAA stop. Introns 2–4 and the boundaries of intron 1 were also sequenced. Filled and open boxes above the RNA map indicate the positions of cysteine-rich domains, which have possible zinc finger structure, and runs of acidic amino acids, respectively (see text and Figs. 2 and 3).

The *trx* gene has been cloned by isolating DNA from the chromosomal region 88B, to which five insertional *trx* mutations map (16). The gene comprises a transcription unit of about 25 kilobases (kb) from which two major RNAs of similar size are derived. These RNAs are distributed uniformly in the early embryo, consistent with the requirement for *trx* function in many if not all segments (16). In the present paper we extend the molecular and genetic analysis of the *trx* gene. A set of overlapping cDNAs corresponding to the major transcripts has been obtained and sequenced,† revealing the existence of zinc finger-like structures in the predicted *trx* protein. Fusion proteins derived from the open reading frame (ORF) of the *trx* gene bind zinc *in vitro*, providing an additional basis for the hypothesis that the *trx* product is a metal-dependent nucleic acid-binding protein. We also provide evidence that *trx* functions as a positive regulator of Ultrabithorax (*Ubx*) expression in the embryo and has additional functions required for survival.

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: ANT-C, Antennapedia complex; BX-C, bithorax complex; ORF, open reading frame; PCR, polymerase chain reaction; PS, parasegment.

*Present address: Division of Biology 156–29, California Institute of Technology, Pasadena, CA 91125.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M31617).

MATERIALS AND METHODS

cDNA Sequencing. Nucleotide sequence was determined for both strands by the Sequenase protocol of United States Biochemical, after subcloning of *Eco*RI fragments in Bluescript plasmid, by the exonuclease III/mung bean or S1 nuclease procedure of Statagene. In cases of ambiguities both strands were also sequenced with oligonucleotide primers.

To localize mutations, both wild-type and mutant genomic DNAs were used as templates in the polymerase chain reaction (PCR) to synthesize fragments which cover the entire *trx* gene except the large intron. Fragments from mutants covering the deletions were subcloned in Bluescript plasmid and sequenced.

***lacZ* Fusion Constructs.** Parts of the *trx* cDNA (see legend to Fig. 4) were obtained by cleavage with restriction enzymes or by PCR and cloned in frame at the 3′ end of the *lacZ* gene of pUR 290, 291, or 292 vector (17). Western blotting and ^{65}Zn binding assay were performed as described (18).

Staining of Embryos with Antibody to the *Ubx* Protein. Twelve- to 15-hr-old embryos were collected from each cross and fixed as described (19). After blocking with 5% horse serum in PBS/Triton (phosphate-buffered saline containing 0.15% Triton X-100) for 3 hr, embryos were incubated overnight at room temperature with diluted culture supernatant (1:1) containing FP.3.38, a monoclonal antibody against *Ubx* protein (20). Embryos were washed several times for 2 hr each with PBS/Triton and then incubated for 2 hr with biotin-conjugated anti-mouse antibody (Vector Laboratories) diluted 100-fold in PBS/Triton containing 5% horse serum. Before and after a further 2-hr incubation with the avidin/biotin-peroxidase complex (Vector Laboratories), embryos were washed extensively with PBS/Triton. Diaminobenzi-

```
MGRSKFPGKPSKSINRKRISVLQLEDDAANPAEPQQPAPESQQPSGSGSGSSAAREKGNNCDNDEDDNAPGGASISGNTASSSAGSGNSGNGSSSGSSTG     100
SGSSGSGSTNGGSVNGGTHHKSAANLDKEAVTKDQNGDGDKTRGNVSSAPSGKLSAAASGKALSKSSRTFSASTSVTSSGRSSGSSPDGNSGASSDGASS     200
GISCGKSTAKSTEAASSGKLAKTTGAGTCSSAKSSKASSLEQLVKQQPLVSGACLKALFVATPATSTGLACALVSPGGSSQGGTFPISAALLRARKNSNKK     300
FKKLNLARGEVMLPSTSKLKQLNSPVVDNPSPSPPILSGSTPSVEGGIGVGGVVSPGEDAALKRVLTEMPNEVARDPSPSSCTAAANGAASGKGSASNGP     400
PAMASSGDGSSPKSGADTGPSTSSTTAKQKKTVTFRNVLETSDDKSVVKRFYNPDIRIPIVSIMKKDSLNRPLNYSRGGECIVRPSILSKILNKNSNIDK     500
LNSLKFRSAGASSSSSNQESRSSSNVFGLSRAFGAPMDEDDEGGVTFRRNDSPEDQNNAEDDEMDDDDDDEEAEEDDQNEDDNDEAASEKSAETEKSAGA     600
DERDPDEKQLVMDSHFVLPKRSTRSSRIIKPNKRLLEEGAISTKKPLSLGDSKGKNVFGTSSSSAGSTASTFSASTNLKLGKETFFNFGTLKPNSSAAGI     700
FVLRQPRLQFQADNQQATFAAPKACPTSPSAIPKPANSLATSSFGSLASTNSSIVTPTPSACSICSAVVSSKEVTQARKYGVVACDVCRKFFSKMTKKSI     800
SANSSTANTSSGSQQYLQCKGNEGSPCSIHSAKSQLKNFKKFYKDRCTACWLKKCMISFQLPAAHRSRLSAILPPGMRGEAAAREEKSAELLSPTGSLRF     900
TSTASSSSPSVVASTSVKWKSSGDSTSALTSIKPNPLAENNVTFGSTPLLRPAILENPLFLKISNAADQKLAAAEAISPSLTKKNSKQEKEKVKESEQSE    1000
KLLSPTQAGTKKSGAAEAQVEEVQPQKEEAPQTSTTTQPSASNGASHGVPQAELAGETNATGDTLKRQRIDLPGPRVKHVCRSASIVLGQPLATFGEDQQ    1100
PEDAADMQQEIAAPVPSAIMEPSPEKPTHIVTDENDNCASCKTSPVGDESKPSKSSGSAQAEVKKATALGKEGTASAAGGSSAKVTTRNAAVASNLIVAA    1200
SKKQRNGDIATSSSVTQSSNQTQGRKTKEHRQQRTLISIDFVENYDPAEVCQTGFGLIVTETVAQRALCFLCGSTGLDPLIFCACCCEPYHQYCVQDEYN    1300
LKHGSFEDTTLMGSLLETTVNASTGPSSSLNQLTQRLNVLCPRCTVCYTCNMSSGSKVKCQKCQKNYHSTCLGTSKRLLGADRPLICVNCLKCKSCSTTK    1400
VSKFVGNLPMCTGCFKLRKKGNFCPICQRCYDDNDFDLKMMECGDCGQVVHSKCEGLSDEQYNLLSTLPESIEFICKKCARRNESSKIKAEEVRQAVMEE    1500
FKASLYSVLKLLSKSRQACALLKLSPRKNVRCTCGASSNQGKLQPKALQFSSGSDNGLGSDGESQNSDDVYEFKDQQQQQQQRNANMNKPRVKPLPCSCQ    1600
QHISHSQSFSLVDIKQKIAGNSYVSLEEFNYDMSQVIQQSNCDELDIAYKELLSEQFPVFQNETKACTDALEEDMFESCSGGNYEDLQDAGGVSASVYNE    1700
HSTSQAESRSGVLDIPLEEVDDFGSCGIKMRLDTRMCLFCRKSGEGLSGEEARLLYCGHDCVVHTNCAMVSAEVFEEIDGSLQNVHSAVARGRMIKCTVC    1800
GNRGATVGCNVRSCGEHYHYPCARSIDCAFLTDKSMYCPAHAKNGNALKANGSPSVTYESNFEVSRPVYVELDRKRKKLIEPARVQFHIGSLEVRQLGAI    1900
VPRFSDSYEAVVPINFLCSRLYVSSKEPVKIVEYTVRTTIQNSSSTLTALDVGRNYTVDHTNPNSKEVQLGMAQIARVHTSLARSEFLENGGTDVSGEFP    2000
NPNSCVPPDQNTEEEPQQQADLLPPVLTSPLKFLGLSTHGGLLLWLLLGVVVRLKQGGELKDAIFEDLPHELLDGISMLDIFLYDDKTDLFAISEQSKDG    2100
TQAMTSNQAQNQNQQAGGANSVSICDEDTRNSNTSLGNGVPASNPVEDAMLSAARNSSQVQMLKTLAVPKLDGNSAMATAIKRRKLSKNLAEGVFLTLSS    2200
QQRNKKEMATVAGVSRRQSISETSVEGVATTSGSVRSKSFTVSAAKRYFEKSEGREEAAKMRIMQMDGVDDSITEFRIISGDGNLSTAQFSGQVKCDRCQ    2300
CTYRNYDAFQRHLPSCSPTMSSNETESDVSGQGMTNNATQISAESLNELQKQLLANAGGLNCLQSATSFPQVQSLGSLGQFGLQGLQQLQLQPQSLGDGF    2400
FLSQPDPATQADTDDLQIYANSLQSLAANLGGGFTLAQPTVTAPAQPQLIAVSTNPDGTQQFIQIPQTMQATTTPTATYQTLQATNTDKKIMLPLTAAGK    2500
PLKTVATKAAQQAAVKQRQLKSGHQVKPIQAKLQPHPQQHQQQQQTQVQQPITVMGQNLLQPQLLFQSSTQTQAPQIILPQAQPQNIISFVTGDGSQGQP    2600
LQYISIPTAGEYKPQPQPTATPTFLTTAPGAGATYLQTDASGNLVLTTTPSNSGLQMLTAQSLQAAPQVIGTLIQPQTIQLGGGADGNQPGSNQQPLILG    2700
GTGGGSSGLEFATTSPQVILATQPMYYGLETIVQNTVMSSQQFVSTAMPGMLSQNASFSATTTQVFQASKIEPIVDLPAGYVVLNNTGDASSAGTFLNAA    2800
SVLQQQTQDDTTTQILQNANFQFQSVPTSSGASTSMDYTSPVMVTAKIPPVTQIKRTNAQAKAAGISGVGKVPPQPQVVNKVLPTSIVTQQSQVQVKNSN    2900
LKQSQVKGKAASGTGTTCGAPPSIASKPLQKKTNMIRPIHKLEVKPKVMKPTPKVQNQNHSLLQQQQQQQPQLQQQIPAVVVNQVPKVTISQQRIPAQTQ    3000
QQQLQQAQMIHIPQQQQPLQQQQVQVQPSMPIITLAEAPVVQSQFVMEPQALEQQELANRVQHFSTSSSSSSSNCSLPTNVVNPMQQQAPSTTSSSTTRP    3100
TNRVLPMQQRQEPAPLSNECPVVSSPTPPKPVEQPIIHQMTSASVSKCYAQKSTLPSPVYEAELKVSSVLESIVPDVTMDAILEEQPVTQSIYTEGLYEK    3200
NSPGESKTEQLLLQQQQREQLNQQLVNNGYLLDKHTFQVEPMDTDVYREEDLEEEEEDEDDDFSLKMRTSACNDHEMSDSEEPAVKDKISKILDNLTNDDC    3300
ADSIATATTMEVDASAGYQQMVEDVLATTAAQSAPTEEFEGALETAAVEAAATYINEMADAHVLDLKQLQNGVELELRRRKEEQRTVSQEQEQSKAAIVP    3400
TAAAPEPPQPIQEPKMTGPHLLYEIQSEDGFTYKSSSITEIWEKVFEAVQVARRAHGLTPLPEGPLADMGGIQMIGLKTNALKYLIEQLPGVEKCSKYT    3500
PKYHKRNGNVSTAANGAHGGNLGGSSASAALSVSGGDSHGLLDYGSDQDELEENAYDCARCEPYSNRSEYDMFSVLASRHRKQPIQVFVQPSDNELVPRR    3600
GTGSNLPMAMKYRTLKETYKDYVGVFRSHIHGRGLYCTKDIEAGEMVIEYAGELIRSTLTDKRERYYDSRGIGCYMFKIDDNLVVDATMRGNAARFINHC    3700
CEPNCYSKVVDILGHKHIIIFAVRRIVQGEELTYDYKFPFEDEKIPCSCGSKRCRKYLN
```

FIG. 2. Amino acid sequence of the predicted *trx* protein. The nucleic acid sequence of 12,445 nucleotides has been submitted to GenBank†; the single long ORF predicts a protein of 3759 amino acids, which is shown in one-letter code. Two introns (at positions 362 and 754) are indicated by triangles. Cysteine and some histidine residues found in putative finger-like structures are indicated by thick underlining, as are the aspartic and glutamic residues in two highly acidic regions. The beginning of the deletion in trx^{B11} is indicated by a filled circle after residue 658; this deletion of 833 nucleotides (2817–3649) leads to a frameshift and thus terminates the ORF. trx^{E3} is a deletion of 832 nucleotides (7336–8167) together with an insertion of 19 nucleotides of unknown origin, and is therefore in frame; the deleted residues are underlined lightly.

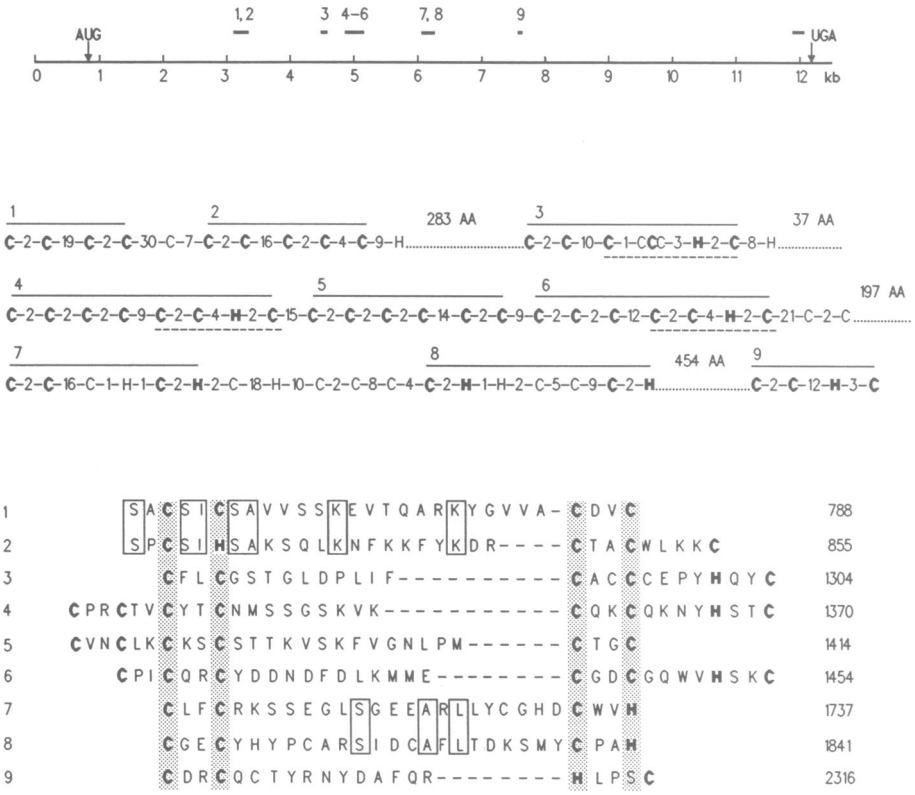

FIG. 3. Putative finger domains of the predicted *trx* protein. (*Top*) cDNA map indicating the positions of cysteine-rich regions (indicated by bars above the map). (*Middle*) Overall pattern of cysteine (C) and histidine (H) residues in the cysteine-rich regions. Possible zinc finger structures are indicated by horizontal lines above the sequence, and numbered. The central region (domains 4–6) may be folded in different ways, yielding up to 5 finger-like structures. Domain 9 is deleted by the trx^{E3} mutation. AA, amino acids. (*Bottom*) Alignment of the proposed finger structures. Domain numbers are given on the left; numbers on the right identify the last residue of each structure. Amino acids that are identical in domains 1 and 2 or 7 and 8 are boxed.

dine was used as the substrate for peroxidase as recommended by the vendor (Amersham). Embryos with suitable staining were mounted in Permount and examined under a Zeiss Axioplan microscope equipped with Nomarski optics.

RESULTS

Sequence of the *trx* cDNA. Fig. 1 *Upper* shows a map of the *trx* region, indicating the five insertion mutations mapped previously (16) and two additional deletion mutations (see also below). Three large cDNAs (15, 7A2, and 6C12; ref. 16) were selected for sequence analysis. Overlaps between these cDNA clones were generated by the PCR technique, and the entire transcript is represented in Fig. 1 *Lower*. Comparison of cDNA sequences with genomic DNA showed the presence of only four introns. The combined sequence revealed a single ORF more than 11 kb long, predicting a protein of 3759 amino acids (Fig. 2). A main feature of the predicted *trx* protein is its cysteine-rich regions, mostly in the central part of the protein; these regions are indicated by the solid blocks above the cDNA map in Fig. 1 and by heavy underlines in Fig. 2 and will be discussed further below. In addition there are two acidic domains consisting of runs of aspartic and glutamic residues, which are heavily underlined in Fig. 2 and indicated by open boxes in Fig. 1. Additional regions with a net negative charge occur at amino acids 1700–2300 and 3100–3450.

The cysteine-rich regions of the *trx* protein can be arranged into sequence motifs that have some similarity with the zinc finger domain, as outlined in Fig. 3. The map at the top shows the location of the cysteine-rich domains and is numbered for reference to the lower parts of the figure. The middle section illustrates the spacing of the cysteine motifs in the central part of the protein, with putative finger domains overlined and numbered; clustering of groups of domains is apparent. The lower part of Fig. 3 gives an alignment of nine regions which emphasizes their arrangement as putative zinc finger motifs. The alignment in the central region (motifs 3–6) is somewhat arbitrary since the multiple cysteine residues permit several arrangements of putative finger structures. The main motif is C2C(H) (C, cysteine; H, histidine; 2 indicates the number of other amino acids), which occurs 27 times in the *trx* protein. Additional similarities are restricted to adjoining structures (1 and 2; 7 and 8) which share some residues beyond the cysteines and have similar loop lengths. The arrangement of cysteines in the *trx* protein shows some similarity to the C_x type of nucleic acid-binding fingers found in hormone receptors, the *Drosophila* transposon *copia*, and certain viral proteins (23–26). This similarity, however, is largely restricted to a general cysteine (and occasionally histidine) motif; with respect to the presence of other conserved amino acids, loop length, and the number of fingers in the protein the *trx* protein is quite distinct from previously reported examples.

Expression Studies: Cysteine-Rich Regions Bind Zinc *in Vitro*. Portions of the *trx* cDNA were fused to a *lacZ* expression vector and expressed in *Escherichia coli*. The various fusions cover almost the entire ORF in overlapping segments (Fig. 4*A*). All of these constructs led to the production of proteins of the predicted molecular size (see Fig. 4*B* for examples), confirming the existence of a continuous ORF extending through more than 11 kb of DNA sequence.

The possibility that the cysteine-rich regions of the *trx* product fold into zinc finger-like structures predicts at a minimum that these regions should be capable of binding zinc. That this is indeed the case is illustrated in Fig. 4*C*. In

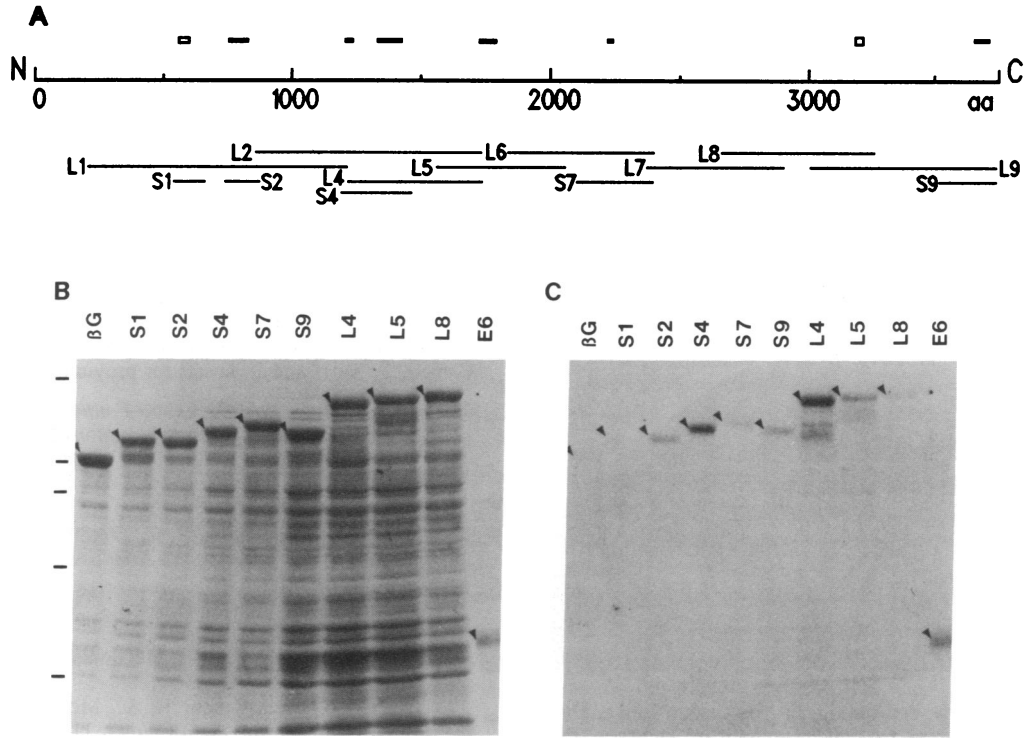

FIG. 4. Zinc binding by *trx* fusion proteins. (*A*) Fusion constructs containing the following parts of *trx* cDNA are indicated below the map: S1 [nucleotides (nt) 2657–3043; amino acids (aa) 605–733], S2 (nt 3118–3490; aa 759–883), S4 (nt 4594–5391; aa 1251–1516), S7 (nt 7319–8107; aa 2159–2421), S9 (nt 11,470–12,115; aa 3542–3757), L1 (nt 1670–4672; aa 276–1276), L2 (nt 3426–6058; aa 861–1738), L4 (nt 4672–6058; aa 1276–1738), L5 (nt 5629–7078; aa 1595–2078), L6 (nt 6421–8107; aa 1859–2421), L7 (nt 8036–9703; aa 2398–2953), L8 (nt 8954–10,668; aa 2704–3275), and L9 (nt 9990–12,115; aa 3049–3757). Indication of cysteine-rich regions and acidic domains is as in Figs. 1 and 3. The ORF starts at nt 842 and ends at nt 12,121. (*B*) Amido black staining of bacterial extracts expressing different fusion proteins (arrowheads). E6, partially purified fusion protein including sequence encoded by the human papillomavirus *E6* gene (18); βG, β-galactosidase. (*C*) Autoradiogram of a blot of the gel shown in *B* after renaturation and labeling with $^{65}ZnCl_2$.

general, the fusion proteins derived from different portions of the *trx* ORF bind zinc in proportion to the number of finger-like motifs they contain. Protein S4, representing the multiple cysteine-motif region, binds zinc most intensely. Protein S7, which contains one putative finger motif and is of special interest since it is deleted by the mutation trx^{E3} (Fig. 1, see also below), shows only weak zinc binding which is of uncertain significance. As an example of a known zinc-binding protein, Fig. 4 *B* and *C* includes a protein derived by fusion of the ORF of the human papillomavirus *E6* gene to the bacterial *trpE* gene (18).

Expression of *Ubx* in the Ventral Nerve Cord Is Affected by Mutations in *trx*. *trx* is known to interact with genes of the BX-C and may be a positive regulator of *Ubx* function (12). To study the nature of this interaction in early development we stained *trx* mutant embryos with antibody to the *Ubx* product. In the nervous system of wild-type embryos the expression domain of *Ubx* extends from parasegment 5 (PS5) to PS12; within this domain PS6 is stained most strongly (refs. 27 and 28 and Fig. 5*a*). In contrast, *Ubx* expression in PS6 is substantially reduced in *trx* mutant embryos (Fig. 5*b*). In addition, a subset of cells in all parasegments that normally expresses *Ubx* appears to be negative in mutant embryos, consistent with an activator function of *trx* (12).

We wished to determine whether the effect on *Ubx* expression requires the entire *trx* product. For this purpose we studied *Ubx* staining in embryos mutant for several *trx* alleles (see legend to Fig. 5). Particularly interesting in this context are two lethal *trx* alleles, trx^{B11} and trx^{E3}, both resulting from short deletions (Fig. 1). The deletion in trx^{B11} leads to a frameshift which terminates the *trx* protein after the first 658 residues; trx^{B11} affects *Ubx* expression in the same way as a total deletion of the *trx* locus [i.e., *Df(3)red*]. In contrast, trx^{E3} deletes 271 amino acids in frame within the coding region (Figs. 1 and 2). While trx^{E3} is a lethal allele, it does not affect *Ubx* staining in the embryonic nervous system; thus, the deleted segment of the protein is dispensable for regulating *Ubx* expression up to this stage. Since trx^{E3} is a lethal mutation and was originally isolated as a suppressor of Polycomb (*Pc*) (30), it is clear that *trx* must have additional functions beyond those involved in *Ubx* expression in the embryonic nervous system.

DISCUSSION

The *trx* gene is required in multiple segments and tissues of the fly and at different stages of development (14). The product of *trx* interacts with genes of both the BX-C and the ANT-C and behaves as an activator of the BX-C (12, 14, 15). Two structural features of the predicted *trx* protein suggest that this effect may be mediated by the function of the *trx* product as a transcription factor. First, the *trx* protein contains several cysteine-rich domains that resemble, in a distant way, zinc finger structures of the C_x (hormone receptor) type (Fig. 3). The protein segments containing these domains are capable of binding zinc *in vitro* (Fig. 4), consistent with such an interpretation. If, in fact, the *trx* product proves to be a metal-dependent nucleic acid-binding protein, then it is clear that it represents a new type with only general similarity to known classes of such proteins. Second, two runs of acidic residues in the predicted *trx* protein resemble acidic stretches in the protein products of the *myc* family and the *Drosophila* engrailed gene (31). These motifs may be functionally similar to activating domains of certain transcription factors (32, 33).

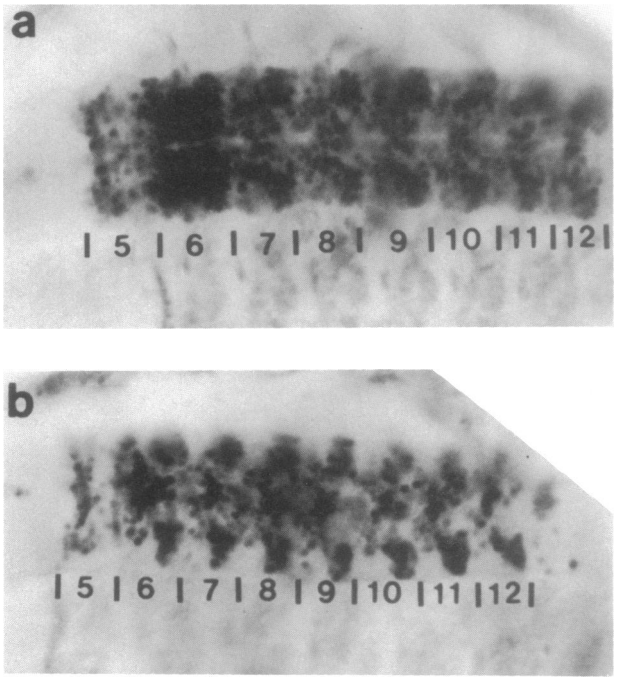

FIG. 5. Staining of the embryonic nervous system with antibody to *Ubx* protein in wild-type (*a*) and *trx* mutant (*b*) embryos. Parasegments are numbered in the photographs. Wild-type embryos show much stronger staining in PS6 than in the other parasegments; in the mutant, staining is almost uniform throughout the major *Ubx*-positive domain. The embryo shown in this figure was homozygous for trx^{3} (29); the same staining pattern was seen with homozygotes for the large deletion $Df(3)red^{P93}$ and for mutations trx^{E2}, trx^{B11}, and trx^{E8} (30). In contrast, the lethal mutation trx^{E3} (30) showed a wild-type staining pattern (see also text).

We observed a distinct effect of *trx* mutations on the expression of *Ubx* in the ventral nerve cord, leading to a substantial decrease in PS6 and lesser reductions in some cells in other parasegments. This finding is consistent with a role for *trx* as a specific regulator of *Ubx*, at least at this stage of development. The role of *trx* could be mediated through one or several regulatory elements of the very long control regions of *Ubx* (34). A comparison between the effect of *trx* mutations and certain mutations in the *Ubx* control regions reveals only limited similarities and suggests the possibility of multiple actions of *trx* on the BX-C. A decrease of *Ubx* expression in PS6 and more posterior parasegments, and an increase in PS5, is associated with mutations in the *bxd/pbx* region located upstream of the *Ubx* promoter (27, 28), while a decrease of expression in PS5 is caused by the mutation abx^{2}, which maps in the 3′ region of the *Ubx* transcription unit (28). The pattern of *Ubx* expression in *trx* mutants has features reminiscent of both *bxd/pbx* and *abx* mutants. Further, *trx* mutants show abdominal transformations (11), implying that this locus also interacts with the regulatory regions of *abd-A* or *Abd-B*, the other two genes in the BX-C.

trx may be considered a member of a class of regulatory genes that are characterized by their interactions. Two unlinked loci, *ash-1* and *ash-2*, are related to *trx* as shown by their similar phenotypes, reciprocal enhancement, and sensitivity to the allelic state of the maternal-effect gene *fsh* (35, 36). Further, *trx* shares with these and a series of additional loci the property of acting as suppressors of *Pc* (30). *Pc* is the prototype of a class of genes with properties of repressors of BX-C activity (37, 38). While both the *Pc* and *trx* groups undoubtedly include genes with different molecular characteristics and biological functions, the existence of regulatory networks within and between these classes appears likely. *Pc* was shown to encode a nuclear protein which is localized at specific chromosomal sites in the *Drosophila* salivary gland chromosomes, including those of the BX-C and ANT-C, implying a plausible model for its apparent repressor activity (39). The sequence of the predicted *trx* protein, as reported here, suggests that *trx* as well may have a nuclear site of action and may be an important member of such a regulatory network.

We thank S. Haynes and J. Kennison for advice and help throughout; S. Haynes, J. Kennison, and J. Kassis for comments on the manuscript; G. Michaels for help with computer searches; N. Patel and C. Goodman for supplying the anti-*Ubx* hybridoma line of White and Wilcox (20); M. Barbarosa for supplying partially purified TrpE–E6 fusion protein and helpful discussions; and P. Beachy, S. Carroll, T. Karr, and J. Kassis for providing staining protocols.

1. Garcia-Belido, A. (1975) *Ciba Found. Symp.* **29,** 161–182.
2. Lewis, E. B. (1978) *Nature (London)* **276,** 565–570.
3. White, R. A. H. & Lehmann, R. (1986) *Cell* **47,** 311–321.
4. Ingham, P. W., Ish-Horowicz, D. & Howard, K. R. (1986) *EMBO J.* **5,** 1659–1665.
5. Duncan, I. M. (1986) *Cell* **47,** 297–309.
6. Ingham, P. W. & Martinez-Arias, A. (1986) *Nature (London)* **324,** 592–597.
7. Hafen, E., Levine, M. & Gehring, W. J. (1984) *Nature (London)* **307,** 287–289.
8. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. (1985) *Science* **229,** 1236–1242.
9. Carroll, S. B., Laymon, R. A., McCutcheon, M. A., Riley, P. D. & Scott, M. (1986) *Cell* **47,** 113–122.
10. Lewis, E. B. (1968) *Proceedings of the 12th International Congress on Genetics* (Science Council of Japan, Tokyo), Vol. 2, pp. 96–97.
11. Ingham, P. W. & Whittle, R. (1980) *Mol. Gen. Genet.* **179,** 607–614.
12. Ingham, P. W. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 201–208.
13. Ingham, P. W. (1983) *Nature (London)* **306,** 591–593.
14. Ingham, P. W. (1985) *J. Embryol. Exp. Morphol.* **89,** 349–365.
15. Sato, T. & Denell, R. E. (1987) *Genetics* **116,** 389–398.
16. Mozer, A. B. & Dawid, I. B. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 3738–3742.
17. Rüther, V. & Müller-Hill, B. (1983) *EMBO J.* **2,** 1791–1794.
18. Barbarosa, M. S., Lowy, D. R. & Schiller, J. T. (1989) *J. Virol.* **63,** 1404–1407.
19. Carroll, S. B. & Scott, M. P. (1985) *Cell* **43,** 47–57.
20. White, R. A. H. & Wilcox, M. (1984) *Cell* **39,** 163–171.
21. Frolman, M. A., Dush, M. K. & Martin, G. L. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
22. Gyllensten, V. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.
23. Klug, A. & Rhodes, D. (1987) *Trends Biochem. Sci.* **12,** 464–469.
24. Evans, R. M. & Hollenberg, S. M. (1988) *Cell* **52,** 1–3.
25. Thomas, S. M., Lamb, R. A. & Paterson, R. G. (1988) *Cell* **54,** 891–902.
26. Berg, J. M. (1986) *Science* **232,** 485–487.
27. Beachy, P. A., Helfand, S. L. & Hogness, D. S. (1985) *Nature (London)* **313,** 545–551.
28. White, R. A. H. & Wilcox, M. (1985) *Nature (London)* **318,** 563–567.
29. Ingham, P. W. (1981) *Wilhelm Roux Arch. Dev. Biol.* **190,** 365–369.
30. Kennison, J. A. & Tamkun, J. W. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8136–8140.
31. Kassis, J. A., Poole, S. J., Wright, D. K. & O'Farrell, P. M. (1986) *EMBO J.* **5,** 3583–3589.
32. Ptashne, M. (1988) *Nature (London)* **335,** 683–689.
33. Sigler, P. B. (1988) *Nature (London)* **333,** 210–212.
34. Peifer, M., Karch, F. & Bender, W. (1987) *Genes Dev.* **1,** 891–898.
35. Digan, M. E., Haynes, S. R., Mozer, B. A., Dawid, I. B., Forquignon, F. & Gans, M. (1986) *Dev. Biol.* **114,** 161–169.
36. Shearn, A. (1989) *Genetics* **121,** 517–525.
37. Lewis, E. B. (1978) *Nature (London)* **276,** 565–570.
38. Jürgens, G. (1985) *Nature (London)* **316,** 153–155.
39. Zink, B. & Paro, R. (1989) *Nature (London)* **337,** 468–471.