

# Cleavage of RNA Hairpins Mediated by a Developmentally Regulated CCCH Zinc Finger Protein

CHUNYANG BAI AND PETER P. TOLIAS\*

Public Health Research Institute, New York, New York 10016

Received 17 June 1996/Returned for modification 1 August 1996/Accepted 30 August 1996

**Control of RNA turnover is a major, but poorly understood, aspect of gene regulation. In multicellular organisms, progress toward dissecting RNA turnover pathways has been made by defining some *cis*-acting sequences that function as either regulatory or cleavage targets (J. G. Belasco and G. Brawerman, *Control of Messenger RNA Stability*, 1993). However, the identification of genes encoding proteins that regulate or cleave target RNAs has been elusive (C. A. Beelman and R. Parker, *Cell* 81:79–183, 1995); this gap in knowledge has made it difficult to identify additional components of RNA turnover pathways. We have utilized a modified expression cloning strategy to identify a developmentally regulated gene from *Drosophila melanogaster* that encodes a RNase that we refer to as Clipper (CLP). Significant sequence matches to open reading frames encoding unknown functions identified from the *Caenorhabditis elegans* and *Saccharomyces cerevisiae* genome sequencing projects suggest that all three proteins are members of a new protein family conserved from lower eukaryotes to invertebrates. We demonstrate that a member of this new protein family specifically cleaves RNA hairpins and that this activity resides in a region containing five copies of a previously uncharacterized CCCH zinc finger motif. CLP's endoribonucleolytic activity is distinct from that associated with RNase A (P. Blackburn and S. Moore, p. 317–433, in P. D. Boyer, ed., *The Enzymes*, vol. XV, part B, 1982) and is unrelated to RNase III processing of rRNAs and tRNAs (J. G. Belasco and G. Brawerman, *Control of Messenger RNA Stability*, 1993, and S. A. Elela, H. Igel, and M. Ares, *Cell* 85:115–124, 1995). Our results suggest that CLP may function directly in RNA metabolism.**

The appropriate expression of specific genes is vital for the proper development of all organisms. When genes are misexpressed, the resulting biochemical defects can give rise to a multitude of cellular and developmental abnormalities. One of the major, but poorly understood, aspects of gene regulation is the control of RNA turnover. In higher eukaryotes, advances in our knowledge of how RNAs are degraded are based on studies defining *cis*-acting sequences that function as either regulatory or cleavage targets (3). A common focus in initiating the turnover of transcripts derived from many developmentally regulated genes appears to involve an endonucleolytic cleavage in the 3' untranslated region of target mRNAs. Though these studies have helped define some of the distinguishing features of specific mRNA turnover targets, little is known about the identity of the proteins that regulate or cleave these transcripts (2, 3). This gap in our understanding of RNA turnover pathways has made it difficult to identify additional components that are required in this process; the identification of these components would help resolve the mechanisms that coordinate transcript degradation during development.

Maternally provided transcripts play a central role in the establishment of egg polarity and embryonic patterning in *Drosophila melanogaster*. The importance of gene products provided by the mother was initially demonstrated by maternal-effect genetic screens (19–21). Subsequent cloning of some genes identified in these genetic screens led to experiments revealing that the establishment of early patterning requires the regulated expression, processing, transport, translation, and turnover of maternal transcripts (reviewed in reference 23). Though it is apparent that during early *drosophila* devel-

opment regulation of maternal transcript turnover is essential, screens for maternal-effect mutations have not identified genes that encode components of RNA turnover pathways. Reasons for this failure could include the involvement of multifunctional genes that are essential in several stages of the *drosophila* life cycle, loci that are haplolethal, mutations that result in weak phenotypes, genes with redundant functions, duplicated genes, possible bypass shunts within a developmental cascade, and genes that are refractory to commonly used mutagenic agents (9, 15, 29). Hence, alternative methods for identification of genes that control the stability of maternal transcripts should be pursued.

Since RNA turnover is a posttranscriptional control mechanism that involves participation of RNases and other RNA-binding proteins, biochemists have traditionally exploited the ability of such proteins to bind nonspecifically to single-stranded DNA (ssDNA) columns for their purification. We have replaced this biochemical approach with a molecular cloning strategy that enabled the identification of ovarian cDNAs expressing such binding activities (24). This screen led to our identification of a developmentally regulated gene from *D. melanogaster* that encodes an RNase that we refer to as Clipper (CLP). CLP specifically cleaves RNA hairpins by using a region that contains five copies of a previously uncharacterized CCCH zinc finger structural motif. This activity is distinct from that associated with other RNases. Our results suggest that CLP defines a new family of conserved proteins and that it may function directly in RNA turnover.

## MATERIALS AND METHODS

**Cloning and purification.** CLP was identified in an expression cloning screen of an ovarian cDNA library for plaques that encode proteins that bind ssDNA (24). The original c6a cDNA encoding CLP was subcloned as an *EcoRI*-*NotI* restriction fragment into the modified pGEX1*NotI* expression vector (24). The N-terminal portion of CLP was subcloned from cDNA c6a as a 646-bp *EcoRI*-

\* Corresponding author. Mailing address: Public Health Research Institute, 455 First Ave., New York, NY 10016. Phone: (212) 578-0815. Fax: (212) 578-0804. Electronic mail address: TOLIAS@PHRI.NYU.EDU.

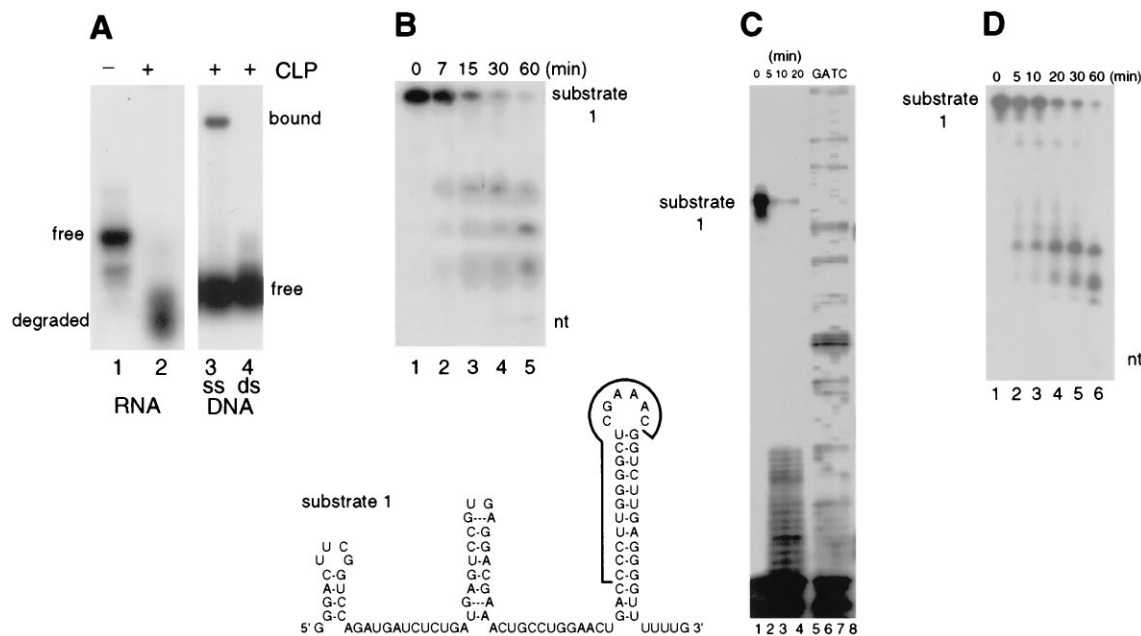


FIG. 1. Characterization of CLP's RNase activity. (A) Gel retardation analysis. Lanes 1 and 2 display the native mobility of reactions performed with an *in vitro*-synthesized  $^{32}\text{P}$ -labeled 65-nt RNA substrate. The GST-CLP fusion protein was omitted from the binding reactions in lane 1 (–), whereas 100 ng of this protein was included in all other reactions (+). Lanes 3 and 4 represent reactions performed with  $^{32}\text{P}$ -labeled ssDNA and dsDNA, respectively, as substrates. The positions of free and degraded RNA and free and bound DNA are indicated. (B) Denaturing gel RNase assay. Lane 1 displays the electrophoretic migration of 96-nt RNA substrate 1 (labeled throughout its length with  $[\alpha\text{-}^{32}\text{P}]\text{GTP}$ ). Lanes 2 to 5 represent reactions that included the GST-CLP fusion protein for the times indicated. (C) Denaturing sequencing gel analysis of CLP-mediated cleavages using a primer extension assay. Unlabeled 96-nt RNA substrate 1 was annealed to a  $^{32}\text{P}$ -end-labeled oligonucleotide primer complementary to 19 nucleotides at the 3' end of the substrate and extended with reverse transcriptase. Lane 1 displays the electrophoretic migration of the fully extended 96-mer near the top of the gel and the unextended 19-mer near the bottom of the gel. Lanes 2 to 4 represent reactions that included the GST-CLP fusion protein for the times indicated prior to addition of primer. Lanes 5 to 8 display a sequence ladder representing the complementary strand of the 96-nt substrate derived from plasmid DNA with the same primer. In the schematic representation of RNA substrate 1, the line indicates the location of CLP-mediated cleavages that were detected by this primer extension experiment. (D) Denaturing gel RNase assay using substrate 1 labeled only at its 5' end (with  $^{32}\text{P}$ ). Lane 1 displays the electrophoretic migration of RNA substrate 1. Lanes 2 to 6 represent reactions that included the GST-CLP fusion protein for the times indicated. nt,  $^{32}\text{P}$ -labeled nucleoside triphosphates.

*PvuII* fragment and placed into pGEX1NotI (24), which was first cleaved with *NotI*, backfilled with Klenow fragment, and recut with *EcoRI* by using standard cloning techniques (17). The remaining C-terminal portion of CLP was subcloned from cDNA c6a as a *PvuII* fragment and placed into the *SmaI* site of pGEX-3X (14). Transformation, growth induction, and affinity purification of the glutathione *S*-transferase (GST)-CLP fusion protein were performed as described by Smith and Johnson (22).

**Activity assays.** The gel retardation assay and the identity and preparation of the labeled RNA and DNAs were performed as described by Bai et al. (1). The denaturing gel RNase assay and the preparation and  $[\alpha\text{-}^{32}\text{P}]\text{GTP}$  labeling of RNA substrates 1, 2, and 3 (throughout their length) were performed as described by Wang et al. (28). In one experiment (see Fig. 1D), substrate 1 was labeled at its 5' end with  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  and polynucleotide kinase. Fifteen-microliter cleavage reaction mixtures containing labeled RNA and 70 ng of GST-CLP fusion protein in CLP binding buffer (10 mM Tris-HCl [pH 7.5], 5 mM  $\text{MgCl}_2$ , 7.5 mM dithiothreitol) were separated on a denaturing 12% polyacrylamide-8 M urea gel. The primer extension assay was performed as previously described (8) with CLP binding buffer. Each 50- $\mu\text{l}$  cleavage reaction mixture contained 400 ng of unlabeled 96-nucleotide (nt) substrate 1 and 100 ng of GST-CLP fusion protein. Extension products were resolved along with products of sequencing reactions (performed with a Sequenase kit [U.S. Biochemical, Cleveland, Ohio]) on a denaturing 5% polyacrylamide-8 M urea gel.

**Expression studies.** Whole-mount *in situ* hybridizations of embryos (27) and ovaries (26) were performed as previously described. The RNase assay was performed with the RPA II kit (Ambion, Austin, Tex.). RNA samples were prepared as described by Brown and Kafatos (6). The two  $^{32}\text{P}$ -labeled *in vitro*-synthesized antisense RNA transcripts were directed against an internal 268-nt region of *clp* (*PstI-XhoI* cDNA fragment) and the terminal 139 nucleotides at the 3' end of a *rp49* cDNA (*PvuII* site). The latter was used as a loading control, but because of its abundance, it was labeled to an approximately 50-fold-lower specific activity.

**Nucleotide sequence accession number.** The sequence of the longest *clp* cDNA is listed under GenBank accession number U26549.

## RESULTS

**Biochemical studies.** Here we report the cloning and biochemical characterization of a novel RNase that may function in RNA turnover. The nucleic acid binding activity encoded by clone c6a, which was represented as a single *Drosophila* ovarian cDNA in a molecular screen for ssDNA-binding activities (24), was examined by using gel retardation analysis. A GST-c6a fusion protein (affinity purified and subsequently visualized as a single band on a Coomassie blue-stained gel) was incubated with  $^{32}\text{P}$ -labeled nucleic acid probes, and the mixture was separated by electrophoresis in a native polyacrylamide gel (Fig. 1A). Lane 1 displays the electrophoretic migration pattern of a 65-nucleotide (nt) RNA transcript (1), selected because it was predicted to contain double-stranded and single-stranded stretches. When this transcript was incubated with the GST-c6a fusion protein under conditions for detecting nonspecific RNA-binding activity (1), we failed to detect bound (retarded) RNA-protein complexes (lane 2). Instead, the mobility increased as if the transcript had been cleaved. The cleaving activity required the presence of  $\text{Mg}^{2+}$  and was never observed with the GST protein alone or with over a dozen unrelated GST fusion proteins. By virtue of its apparent RNase activity, we named this protein CLP. Lanes 3 and 4 show that CLP can nonspecifically bind ssDNA but not double-stranded DNA (dsDNA), the basis of its identification in our expression cloning screen.

Analysis on denaturing gels with a homogeneously labeled

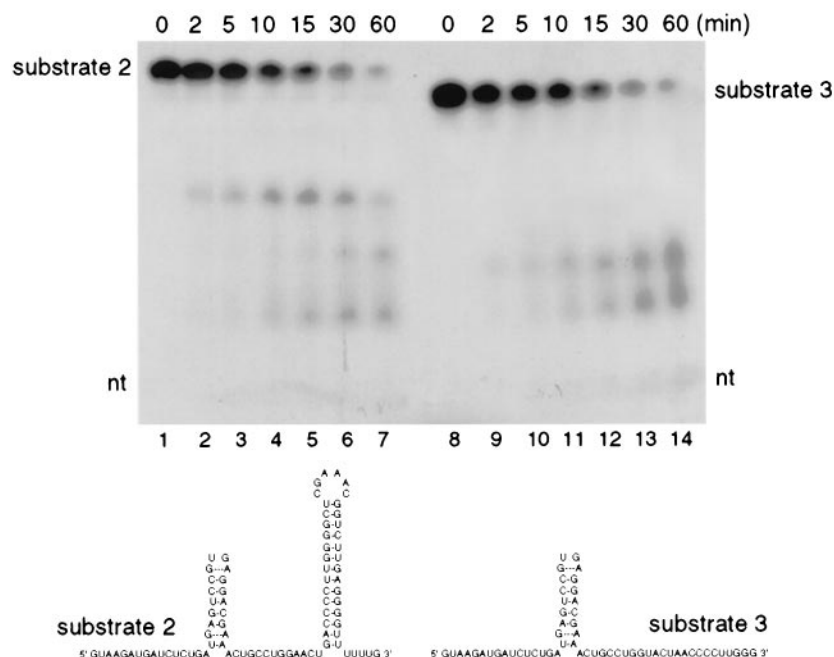


FIG. 2. Denaturing gel RNase assay using other substrates. Lanes 1 and 8 display the electrophoretic migration of RNA substrates 2 and 3, respectively (labeled throughout their length with [ $\alpha$ - $^{32}$ P]GTP). Lanes 2 to 7 represent reactions that included substrate 2 and the GST-CLP fusion protein for the times indicated. Lanes 9 to 14 represent reactions that included substrate 3 and the GST-CLP fusion protein for the times indicated. nt,  $^{32}$ P-labeled nucleoside triphosphates.

96-nt RNA (substrate 1), whose 5' and 3' ends are apparently protected against exonuclease cleavage by hairpins (provided by K. Drlica, Public Health Research Institute, New York, N.Y.), gave the first indication that CLP may have endonucleolytic activity (Fig. 1B). The appearance of three major products suggested a limited number of CLP-mediated cleavages. However, at this point, we could not distinguish whether the cleavages were actually catalyzed by an endo- or an exo-RNase that paused at the hairpins. To further characterize CLP's activity, we treated unlabeled RNA substrate 1 with the GST-CLP fusion protein, annealed a labeled oligonucleotide primer complementary to 19 nucleotides at the 3' end of the substrate, and extended the primer with reverse transcriptase. If CLP possessed a predominant 3' to 5' exo-RNase activity, then extension products should not have been generated by this assay. The reaction products were analyzed by electrophoresis in a denaturing DNA sequencing gel (Fig. 1C); a sequence ladder of the substrate derived from plasmid DNA with the same primer was also included. The structure of RNA substrate 1 and the locations of cleavages are also depicted schematically in Fig. 1. This experiment demonstrated that cleavages occurred throughout the large hairpin at the 3' end of substrate 1. With the exception of three nucleotides at the base of the duplex, digestion occurred throughout the hairpin in both dsRNA and the ss loop region. These results proved that CLP is not a 3'-to-5' exo-RNase.

To distinguish whether CLP is functioning as an endo- versus a 5'-to-3' exo-RNase that is slowed down by RNA secondary structure, we repeated the denaturing gel RNase assay with substrate 1 labeled only at its 5' end (Fig. 1D). If CLP had a predominant 5'-to-3' exo-RNase activity, then specific cleavage products should not be detected and we would expect the accumulation of a single  $^{32}$ P-labeled nucleotide (derived from the 5' end of substrate 1) at the bottom of the gel. The results shown in Fig. 1D displaying three major cleavage products and no accumulation of nucleotides prove that CLP is not a 5'-to-3'

exo-RNase. Hence, on the basis of all of the above-presented data, we conclude that CLP is an endo-RNase.

We also investigated whether removal of some of the hairpins may have affected CLP's substrate preference. To do this, we examined the ability of the GST-CLP fusion protein to cleave RNAs related to substrate 1 that contained only one or two hairpins. Three major cleavage products were detected with a truncated RNA (substrate 2) that lacked the 5'-terminal hairpin (Fig. 2, lanes 1 to 7). Another derivative RNA that lacked both the 5' and 3' hairpins (substrate 3) was cleaved by CLP only at the single eccentrically located hairpin, leaving intact the flanking ssRNA, which migrated as two distinct cleavage products (Fig. 2, lanes 8 to 14). These results further demonstrate the importance of RNA secondary structure with respect to CLP's apparent endoribonucleolytic specificity for RNA hairpins.

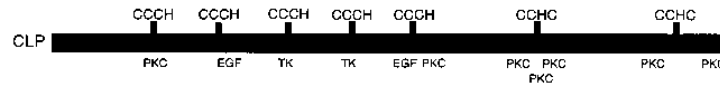
**Sequence analysis.** The original *clp* cDNA was used as a probe to isolate four additional clones. Sequence analysis (Fig. 3A) revealed that they were all identical except for the length of their 5' ends; the sizes of cDNAs varied between 1,180 and 1,246 bp. Computer analysis predicted a 296-amino-acid open reading frame encoding an amphipathic, basic protein with an estimated pI of 9.2 and a molecular mass of 33.5 kDa. CLP appears to contain a total of seven zinc finger motifs (conserved nucleic acid-binding structural domains that form because of the specific arrangement of cysteine and histidine residues chelated around  $Zn^{2+}$  [4]). The first five fingers present at the N-terminal portion of CLP are all members of the CCCH class, whereas the C terminus contains two CCHC motifs (Fig. 3). The CCHC motif, also known as the zinc knuckle, binds  $Zn^{2+}$  (25) and is often present in proteins that bind exclusively to ssDNA or ssRNA. The CCCH motif is a rare and uncharacterized zinc finger. It was first noticed in a group of proteins encoded by genes that are coinduced in cultured mammalian cell lines by growth factors, tumor promoters, or serum (7). One of these proteins, TIS11/nup45/

**A**

```

GCACTATTTCGCACTGAAACAACAACACCGCCAAAAAAGAAGGAAACTGCATTGTTTACCAGATTTTCGGAGTTTTTCCTCATCTATCCGAGCATCC 100
GCGAATCCGTCAGAG ATG GAC ATC CTT TTG GCC AAC GTA AGT GGG CTG CAG TTC AAG GCG GAG CCG GAC CTC ATC GAG 178
                M D I L L A N V S G L Q F K A E R D L I E 21
CAG GTG GGT GCC ATC CCG CTG CCG TTC TAC GGC ATG GAC AAG TCC ATT GCG GCG STC TGC AAC TTC ATC ACC AGG 253
Q V G A I F L F F Y G M D K S I A A V C N F I T R 46
AAC GGA CAG GAG TGC GAC AAA GGG AGC GCC TGT CCC TTC CGA CAC ATT GCG GGG GAC CCG ACG ATC GTG TGC AAG 328
N G Q E C D K G S A C P F R H I R G D R T I V C K 71
CAC TGG CTG AGA GGA CTG TGC AAG AAG GGC GAC CAG TGC GAG TTC CTG CAC GAG TAC GAC ATG ACC AAG ATG CCC 403
H W L R G L C K K G D Q C E F L H E Y D M T K M P 96
GAG TGC TAC TTC TAC TCG AGG TTC AAC GCC TGC CAC AAC AAA GAG TGT CCC TTT TTG CAC ATC GAC CCG CAG AGC 478
E C Y F Y S R F N A C H N K E C P F L H I D P Q S 121
AAG GTG AAG GAT TGT CCG TGG TAC AAG AGA GGC TTC TGC CGC CAC GGT CCC CAC TGC CCG CAC CAG CAT CTG CGC 553
K V K D C P W Y K R G F C R H G F H C R H Q H L R 146
CGT CTC CTG TGC ATG GAC TAC CTA GCC GGC TTC TGC CCC GAG GCG CCG TCC TGC AAG CAC ATG CAT CCG CAC TTC 628
R V L C M D Y L A G F C P E A P S C K H M H P F 171
GAG CTG CCC CCG CTG GCG GAA CTG GGT AAG GAT CAG CTG CAC AAG AAA CTG CCC ACG TGC CAC TAT TGC GGC GAG 703
E L P P L A E L G K D Q L H K K L P T C H Y C G E 196
CTG GGC CAC AAG GCC AAC TCG TGC AAG CAG TAT GTG GGC AGC CTG GAG CAT CGC AAC AAT ATC AAC GCA ATG GAT 778
L G H K A N S C K Q Y V G S L E H R N N I N A M D 221
CAC TCC GGC GGA CAC TCG GGC GGC TAC TCT GGA CAC TCC GGC CAC ATC GAG GGT GCC GAT GAC ATG CAA TCC AAT 853
H S G G H S G G Y S G H S G H I E G A D D M Q S N 246
CAC CAC AGT CAG CCG CAT GGT CCC GGC TTC GTC AAG GTG CCC ACG CCC CTG GAG GAG ATC ACC TGC TAC AAG TGC 928
H H S Q P H G P G F V K V P T P L E E I T C Y K C 271
GGG AAC AAG GGC CAC TAC GCG AAC AAG TGT CCC AAG GGT CAC CTC GCC TTC CTC TCT AAC CAA CAT AGT CAT AAG 1003
G N K G H Y A N K C P K G H L A P L S N Q H S H K 296
TAGGACGCTCTGCTGTATGATTAGACTGTAAGAGTGCAGAAATAAAGGTCGAAAGCGGCTGTCATAGTCCCGCTGTGTAGTCCAGAGGTCGTAGGTTCC 1103
CAACTCCCGCAGCGCAACGTTCAATCATTTCTGTGTTTAAAGTGTATGTTAAGAAACTCATTTCGCATTTTTGAAAAAGTCCCTAGAACTGCCAGTAGAA 1203
CGTATTGTGTTATCAGCGAATAAACTTCCATAAGATACAAGT 1246
    
```

**B**



**C**

```

41 C N P I T R N G Q E C D K G S A C P F R H
70 C K H W L R G L C K K G D O C E F L K
98 C Y F Y S R Y N A C H N K E C P F L H
126 C P W Y K R G F C R H G P H C R H O H
150 C M D Y L A G F C P E A P S C K H M H
191 C H Y C G E L G H K A N S C
268 C Y K C G N K G H Y A N K C
C x (4-12) G x C x x G x x C x F x H C x x C G x x G H x x x x C
    
```

FIG. 3. (A) cDNA and deduced protein sequences of CLP. In the protein sequence, the five CCCH motifs (underlined) are followed by two CCHC motifs (double underlined). Eleven putative serine, threonine, and TK phosphorylation consensus sequences are indicated by italics. The cDNA sequence contains a polyadenylation consensus sequence (underlined nucleotides) and is flanked by a poly(dA) tail (not shown). (B) Linear schematic representing the full-length CLP protein and the relative positions of five CCCH and two CCHC zinc fingers, as well as putative phosphorylation consensus recognition sequences for protein kinase C (PKC), epidermal growth factor (EGF), and TK. (C) Alignment of seven putative zinc finger repeats in CLP. The number to the left of each finger indicates its relative position in the CLP protein sequence. The known consensus sequences for both CCCH and CCHC zinc finger motifs are also displayed.

TTP, was shown to be a nuclear protein that is capable of binding  $Zn^{2+}$ . A member of this class of zinc finger proteins was later identified in *D. melanogaster* and shown to be encoded by *unkempt*, a gene that is essential for postembryonic development (12). Though these reports have all speculated on a putative nucleic acid binding activity for members of the CCCH family, such an activity has never been demonstrated.

When the predicted amino acid sequence of CLP was compared with those of all proteins listed in the GenBank database, two significant matches were identified. The first match was to the product of open reading frame F11A10.3 from *Caenorhabditis elegans*, identified by the Nematode Genome Sequencing Project (accession number Z68297) and displaying

more than 52% identity (69% similarity) over the N-terminal 204 amino acids of CLP (Fig. 4), with additional identities extending to CLP's C terminus (data not shown). A second match was obtained with the 190-amino-acid product of open reading frame P8283.17 from *Saccharomyces cerevisiae*, identified by the Yeast XVI Chromosome Sequencing Project (accession number U32445) and displaying greater than 32% amino acid matches (47% conservative substitutions) over this same region of CLP. All three proteins contain five tandemly repeated copies of the CCCH zinc finger motif; the second finger is almost identical in all three species. In addition, the first of CLP's two CCHC zinc knuckle motifs is also conserved in the worm protein. Though the functions of the putative

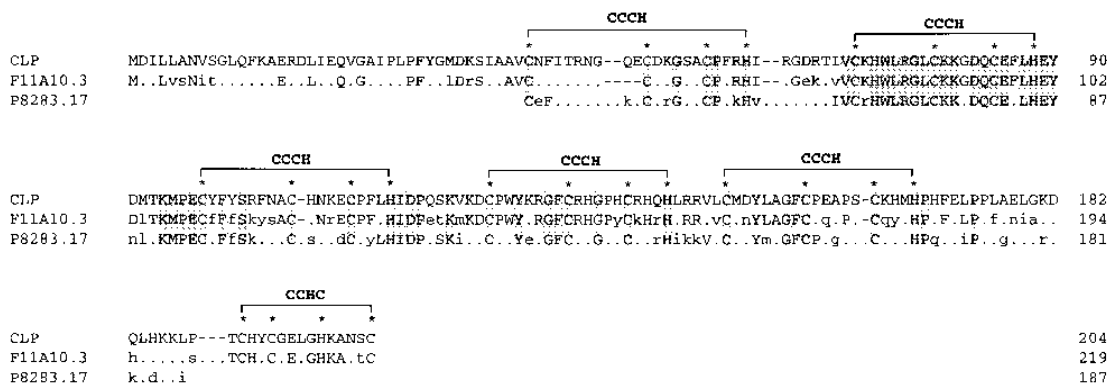


FIG. 4. Protein alignment of CLP and two related hypothetical proteins identified in the nematode (F11A10.3) and yeast (P8283.17) genome sequencing projects. Alignments were performed with sequence data provided by GenBank (sequences were retrieved with the program BLAST). Uppercase letters denote the CLP sequence and identical residues present in either of the other proteins. Residues identical in all three proteins are shaded. Lowercase letters indicate conservative changes. Dots denote nonconservative changes. Dashes indicate gaps in protein sequences introduced to achieve optimum alignment. The CCCH and CCHC zinc finger consensus sequences are indicated with brackets and stars. GenBank accession numbers for the yeast and nematode proteins are Z68297 (F11A10.3 from *C. elegans*) and U32445 (P8283.17 from *S. cerevisiae*).

worm and yeast proteins are unknown, their significant sequence conservation with CLP suggests that all three are members of a new family of proteins conserved from lower eukaryotes to invertebrates.

To determine whether the endoribonucleolytic activity of CLP might be associated with the CCCH domain, we fused distinct segments of CLP that contained either the five N-terminal CCCH or the two C-terminal CCHC fingers to GST, affinity purified the fusion proteins, and tested them for RNase activity using the denaturing gel assay. The results presented in Fig. 5 demonstrate that the RNase activity associated with CLP resides in the region containing the five N-terminal CCCH zinc fingers.

A striking feature of the predicted CLP amino acid sequence is the presence of an unusually high number of putative serine, threonine, and tyrosine kinase (TK) consensus phosphorylation recognition sites (16), most of which are contained within or near the zinc fingers (Fig. 3A and B). This raises the possibility that CLP may be regulated by protein kinase C-, TK-, and epidermal growth factor-mediated signal transduction pathways.

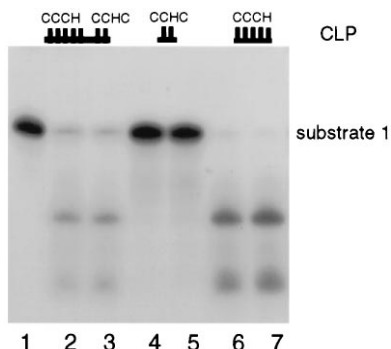


FIG. 5. Mapping CLP's RNase activity to the five N-terminal CCCH zinc fingers by using the denaturing gel assay. Labeled RNA and 70 ng of the fusion proteins indicated below were incubated for 30 min. Lane 1 displays the electrophoretic migration (on a denaturing 9% polyacrylamide-8 M urea gel) of the <sup>32</sup>P-labeled 96-nt transcript. Lanes 2 and 3 represent duplicate reactions that included the GST-CLP fusion protein. Lanes 4 and 5 represent duplicate reactions that included a GST fusion protein containing the two C-terminal CCHC zinc fingers. Lanes 6 and 7 represent duplicate reactions that included a GST fusion protein containing the five N-terminal CCCH zinc fingers.

**Developmental expression.** We also examined the expression of *clp* mRNA throughout fly development by Northern (RNA) blot analysis. A single 1,300-nt transcript, barely detected, was confined to early embryos and adult female ovaries (data not shown). These observations prompted us to focus on the distribution of the *clp* transcript during oogenesis and embryogenesis (Fig. 6) by whole-mount in situ hybridization using a digoxigenin-labeled *clp* antisense RNA probe. Expression of *clp* mRNA was first detected in both the nurse cells and the oocyte of stage 3 egg chambers (Fig. 6A). Throughout oogenesis, *clp* transcripts were absent from somatically derived follicular epithelial cells surrounding the follicle (Fig. 6A to C). By stage 8, the oocyte no longer displayed *clp* mRNA. Subsequent expression was restricted to nurse cells (Fig. 6B, which shows stage 10) until stage 12 (Fig. 6C), when the nurse cells rapidly transferred all of the *clp* mRNA into the oocyte. During early embryonic syncytial nuclear divisions, maternally provided *clp* transcripts were homogeneously distributed (Fig. 6D). When the synchronous division cycles ended at blastoderm cellularization, maternally derived *clp* transcripts were barely visible (Fig. 6E). They were no longer detected by 5 h postfertilization. Zygotic expression of *clp* mRNA during the rest of embryonic development was not detected.

Since *clp* does not encode an abundant transcript, we decided to examine its developmental expression with the more sensitive and quantitative RNase protection assay. RNA samples prepared from particular developmental stages were annealed with two <sup>32</sup>P-labeled, in vitro-synthesized, antisense RNA transcripts directed against a 268-nt region of *clp* and a 139-nt region of *rp49* (the latter was used as a loading control). Samples were then treated with RNase A and RNase T<sub>1</sub> to degrade ssRNA, and uncleaved duplexes were separated by electrophoresis in a denaturing gel. This approach revealed *clp* to be differentially expressed throughout drosophila development (Fig. 7). *clp* mRNA was most abundant in adults (Fig. 7, lanes 10 and 11), and the majority of the female expression was ovarian (lanes 12 and 13). Maternal transcripts were then provided to the embryo (Fig. 7, lane 1), and transcript accumulation gradually decreased (lanes 2 to 5) until a zygotically derived burst was observed in third-instar larvae (lane 6). This increased transcript accumulation gradually decayed to barely detectable levels in late pupae (Fig. 7, lanes 7 to 9).

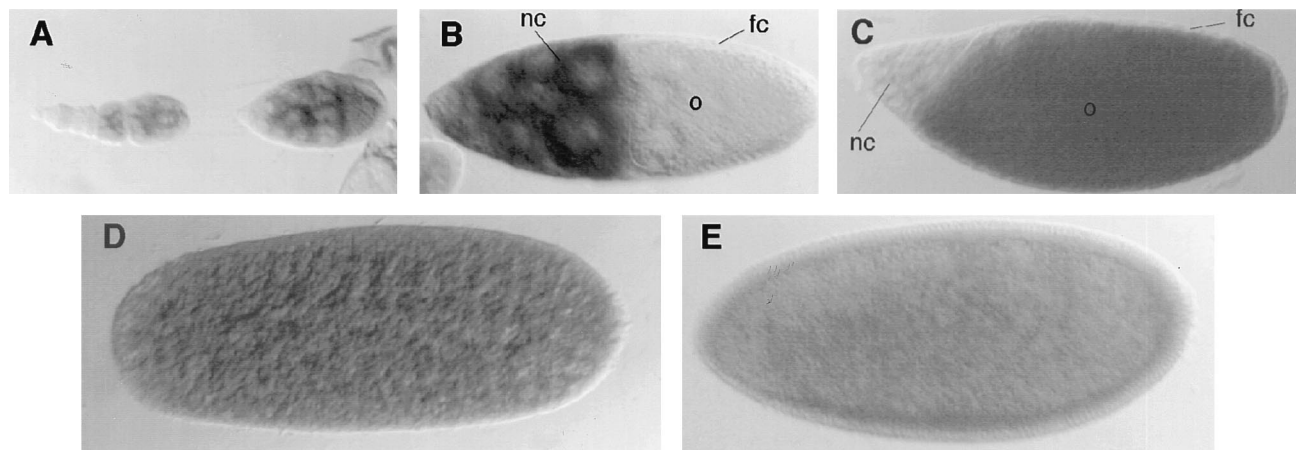


FIG. 6. Expression of *clp* transcripts during oogenesis and embryogenesis. Whole-mount in situ hybridization was performed with a digoxigenin-labeled *clp* antisense RNA probe (prepared with the Genius 4 kit; Boehringer Mannheim, Indianapolis, Ind.). All photographs were taken with a 20 $\times$  objective and 10 $\times$  ocular lens, and expression is reflected by the dark staining. Egg chambers (A to C) are oriented with the anterior end on the left. (A) Expression in an early ovariole commencing at stage 3 (10) was restricted to the germ line-derived nurse cells and the oocyte but was absent from the somatically derived follicle cells. (B) Expression in stage 10b egg chambers was restricted to nurse cells (nc), o, oocyte; fc, follicle cells. (C) *clp* transcripts were completely transferred to the oocyte by stage 12. Throughout oogenesis, expression was not detected in follicle cells. (D) Early (syncytial blastoderm) embryos maintained maternally derived *clp* transcripts. (E) Maternal transcripts were barely detected at the onset of cellularization and were subsequently lost.

## DISCUSSION

We have demonstrated that *clp* is a developmentally regulated gene that encodes a bipartite zinc finger protein. The significant sequence conservation among CLP and apparent homologs in *C. elegans* and *S. cerevisiae* defines a new family of proteins conserved in *S. cerevisiae*, *D. melanogaster*, and *C. elegans*. CCCH zinc fingers of the type found in CLP have been found in only a handful of proteins; we show that five copies of a region in CLP containing the CCCH motif confer a RNase activity with specificity for RNA hairpins. The structure and activity of the CCCH zinc fingers are distinct from those of motifs conserved in bacterial and yeast endo-RNases, such as

RNase III, that are known to process rRNA and tRNA precursors by recognizing specific dsRNA sequences and cleaving at a single position within each target site (3, 8). Moreover, the activity of CLP was not affected by RNasin (Promega, Madison, Wis.) (data not shown), a known inhibitor of some RNases (such as RNase A, B, and C) that are active against a broad range of RNA substrates (5, 11). Hence, CLP possesses a novel activity that distinguishes it from other known eukaryotic endo-RNases.

In multicellular organisms, it is vital to rapidly degrade specific mRNAs that encode critical regulatory proteins that function only briefly during development. A central focus of many developmentally regulated mRNA turnover pathways is the initiation of degradation by an endonucleolytic cleavage, usually in the 3' untranslated region of target transcripts that typically exhibit extensive secondary structure (3). This cleavage is believed to render target mRNAs vulnerable to subsequent degradation by exonucleases. CLP's unique preference for RNA hairpins and its specific and thorough destruction of these structures fulfills a criterion for initiation of mRNA turnover by an endo-RNase. If CLP is truly initiating mRNA turnover, what prevents it from launching the destruction of all RNAs in cells that express *clp* transcripts? One possibility is that *clp* transcripts may be regulated at the level of translation. Other possibilities include controlling the transport, activity, or specificity of the CLP protein, possibly by phosphorylation at some of its 11 putative serine, threonine, and TK consensus recognition sites. In this respect, it is interesting that the protein kinase C, TK, and epidermal growth factor signal transduction pathways have all been implicated in controlling the turnover of specific eukaryotic mRNAs (3). Another scenario is that the two C-terminal CCHC zinc fingers may mediate recognition of specific target sequences, thus exposing only RNAs that contain such sites to the N-terminal ribonucleolytic activity for hairpins.

With respect to drosophila development, the abundant expression of *clp* transcripts in the maternal germ line suggests that CLP may function directly in the control of maternal mRNA turnover. While this process is not well understood, it clearly plays an important role in the regulation of transcripts

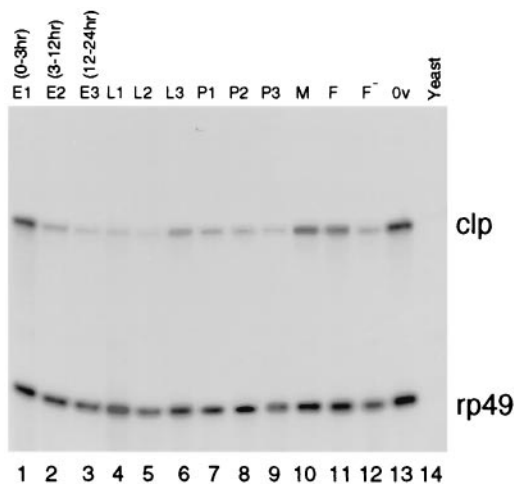


FIG. 7. Developmental RNase protection assay. The assay was performed with 20  $\mu$ g of total RNA isolated at the indicated developmental stages (defined below). The mobilities of the labeled *clp* and *rp49* antisense RNA probes that annealed to particular samples and resisted degradation are indicated. RNA samples were prepared from 0- to 3-h embryos (E1), 3- to 12-h embryos (E2), 12- to 24-h embryos (E3), first-instar larvae (L1), second-instar larvae (L2), third-instar larvae (L3), early pupae (P1), middle pupae (P2), late pupae (P3), adult males (M), adult females (F), ovariectomized females (F<sup>-</sup>), and ovaries (Ov). Yeast RNA (20  $\mu$ g) was included (lane 14) as a negative control to ensure that all of the input labeled antisense probes were completely degraded.

that control the initial divisions, growth, and establishment of polarity in the early syncytial embryo. Our results also support the genetic studies of Myers et al. (13), who demonstrated that turnover of drosophila maternal transcripts depends on expression of maternally contributed gene products. Hence, CLP may be one of these maternally contributed gene products, and its activity, specificity, and/or localization may be directly regulated by phosphorylation through signal transduction cascades employing protein kinases. A second wave of *clp* transcript accumulation during late larval development suggests that CLP-mediated mRNA turnover may help prepare for pupal metamorphosis.

Using *in situ* hybridization to larval polytene chromosomes (18), we have cytologically mapped *clp* to interval 21D1,2 on the second chromosome (data not shown), where lethal saturation mutagenesis screens performed by others have revealed nine essential loci in this region (16a). Future plans will focus on identification of *clp* mutants and characterization of RNA turnover pathways in *clp* loss-of-function alleles.

#### ACKNOWLEDGMENTS

We thank Z. Li and C. R. Chen for technical assistance and K. Drlica and L. Qiu for helpful suggestions and for providing DNA templates. We also thank K. Drlica, M. Konsolaki, J. Bijlsma, D. Dubnau, and L. Mindich for critical reading of the manuscript.

This work was supported by National Science Foundation grant IBN-9418722 awarded to P.P.T.

#### REFERENCES

- Bai, C., Z. Li, and P. P. Tolias. 1994. Developmental characterization of a *Drosophila* RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen. *Mol. Cell. Biol.* **14**:5123–5129.
- Beelman, C. A., and R. Parker. 1995. Degradation of mRNA in eukaryotes. *Cell* **81**:179–183.
- Belasco, J. G., and G. Brawerman. 1993. Control of messenger RNA stability. Academic Press, San Diego, Calif.
- Berg, J. 1990. Zinc fingers and other metal-binding domains. *J. Biol. Chem.* **265**:6513–6516.
- Blackburn, P., and S. Moore. 1982. Pancreatic ribonuclease, p. 317–433. *In* P. D. Boyer (ed.), *The enzymes*, vol. XV, part B. Academic Press, New York.
- Brown, N. H., and F. C. Kafatos. 1988. Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.* **203**:425–437.
- Dubois, R. N., M. W. McLane, K. Ryder, L. F. Lau, and D. Nathans. 1990. A growth factor-inducible nuclear protein with a novel cysteine-histidine repetitive sequence. *J. Biol. Chem.* **265**:19185–19191.
- Elela, S. A., H. Igel, and M. Ares. 1996. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell* **85**:115–124.
- Jurgens, G., E. Wieschaus, C. Nusslein-Volhard, and H. Kluding. 1984. Mutations affecting the pattern of larval cuticle in *Drosophila melanogaster*. II. Zygotic loci on the third chromosome. *Roux's Arch. Dev. Biol.* **193**:283–295.
- King, R. C. 1970. Ovarian development in *Drosophila melanogaster*. Academic Press, New York.
- Kobe, B., and J. Deisenhofer. 1995. A structural basis of the interactions between leucine rich repeats and protein ligands. *Nature (London)* **374**:183–186.
- Mohler, J., N. Weiss, S. Murli, S. Mohammadi, K. Vani, G. Vasilakis, C. H. Song, A. Epstein, T. Kuang, J. English, and D. Cherdak. 1992. The embryonically active gene, *unkempt*, of *Drosophila* encodes a Cys<sub>2</sub>His finger protein. *Genetics* **131**:377–388.
- Myers, F. A., H. Francis-Lang, and S. F. Newbury. 1995. Degradation of maternal string mRNA is controlled by proteins on maternally contributed transcripts. *Mech. Dev.* **51**:217–226.
- Nagai, K., and C. Thoegerson. 1984. Generation of  $\beta$ -globin by sequence-specific proteolysis of hybrid protein produced in *Escherichia coli*. *Nature (London)* **309**:810–812.
- Nusslein-Volhard, C., E. Wieschaus, and H. Kluding. 1984. Mutations affecting the pattern of larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Roux's Arch. Dev. Biol.* **193**:267–282.
- Pearson, R. B., and B. E. Kemp. 1991. Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol.* **200**:62–81.
- 16a. Roberts, D. B., H. W. Brock, N. C. Rudden, and S. Evans-Roberts. 1985. A genetic and cytogenetic analysis of the region surrounding the LSP-1 gene of *Drosophila melanogaster*. *Genetics* **109**:145–156.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Saunders, R. D. C., D. M. Glover, M. Ashburner, I. Siden-Kiamos, C. Louis, M. Monastirioti, C. Savakis, and F. C. Kafatos. 1989. PCR amplification of DNA microdissected from a single polytene chromosome band: a comparison with conventional microcloning. *Nucleic Acids Res.* **17**:9027–9037.
- Schüpbach, T., and E. Wieschaus. 1986. Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo. *Roux's Arch. Dev. Biol.* **195**:302–317.
- Schüpbach, T., and E. Wieschaus. 1989. Female sterile mutations on the second chromosome of *Drosophila melanogaster*. I. Maternal effect mutations. *Genetics* **121**:101–117.
- Schüpbach, T., and E. Wieschaus. 1991. Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**:1119–1136.
- Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31–40.
- St. Johnston, D., and C. Nusslein-Volhard. 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**:201–219.
- Stroubakis, N. D., Z. Li, and P. P. Tolias. 1994. RNA- and single-stranded DNA-binding (SSB) proteins expressed during *Drosophila melanogaster* oogenesis: a homolog of bacterial and eukaryotic mitochondrial SSBs. *Gene* **143**:171–177.
- Summers, M. F. 1991. Zinc finger motif for single-stranded nucleic acids? Investigations by nuclear magnetic resonance. *J. Cell. Biochem.* **45**:41–48.
- Suter, B., and R. Steward. 1991. Requirement for phosphorylation and localization of the bicaudal-D protein in *Drosophila* oocyte differentiation. *Cell* **67**:917–926.
- Tautz, D., and C. Pfeifle. 1989. A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**:81–85.
- Wang, J.-Y., L. Qiu, E.-D. Wu, and K. Drlica. 1996. RNases involved in ribozyme degradation in *Escherichia coli*. *J. Bacteriol.* **178**:1640–1645.
- Wieschaus, E., C. Nusslein-Volhard, and G. Jurgens. 1984. Mutations affecting the pattern of larval cuticle in *Drosophila melanogaster*. III. Zygotic loci on the X chromosome and on the fourth chromosome. *Roux's Arch. Dev. Biol.* **193**:296–307.