Elevated *couch potato* transcripts associated with adult diapause in the mosquito *Culex pipiens*

Qirui Zhang and **David L. Denlinger**[¶]
Departments of Entomology and Evolution, Ecology, and Organismal Biology, Ohio State University, 318 West 12th Avenue, Columbus, OH 43210, USA

Abstract

The *couch potato* (*cpo*) cDNA that we cloned from the northern house mosquito, *Culex pipiens*, encodes the C-terminus containing a highly conserved RNA recognition motif (RRM). Protein structure prediction indicates a canonical RRM structure with a βαββαβ topological structure. Northern blots indicate a single mRNA band over 9.49 kb, and Southern blot analysis suggests that the *cpo* gene contains large introns. Highest expression was noted in first instar larvae and pupae. A comparison of nondiapausing (long daylength) and diapausing (short daylength) adult females showed no difference immediately at adult eclosion, but by day 7 and thereafter, expression of *cpo* was much higher in diapausing adults. When 2-month old diapausing females were transferred from short daylength to diapausing-terminating conditions of long daylength and high temperature, expression of *cpo* declined. Similarly, when a topical application of JH III was used to terminate diapause abundance of the *cpo* transcript declined. Consistent with observations in *Drosophila melanogaster* and several other species levels of *cpo* in *C. pipiens* are influenced by the diapause program, although the direction of change is not the same in all species.

Keywords

couch potato; gene structure; expression pattern; diapause; *Culex pipiens*

1. Introduction

Our interest in *couch potato* (*cpo*) was piqued by the intriguing evidence that populations of *Drosophila melanogaster* having a high incidence of diapause express *cpo* at lower levels than noted in populations with a low diapause incidence (Schmidt et al., 2008). The *Drosophila* results prompted these experiments comparing expression of *cpo* in diapause- and nondiapause-programmed adults of the mosquito *Culex pipiens* to see if the importance of *cpo* in diapause extends beyond *Drosophila*.

Cpo was first identified from *D. melanogaster* during enhancer detection screens to identify genes induced during differentiation of the peripheral nervous system, and it was given the name *cpo* because several insertion alleles caused hypoactive behavior in adults such as slow recovery from ether anesthesia, abnormal geotaxis, phototaxis, and flight behavior (Bellen et al., 1992a; 1992b). *Cpo* belongs to an RNA binding protein family with complex

[¶]Corresponding author. Tel.: +1 614 2926425; Fax: +1 614 2922180 denlinger.1@osu.edu (D.L. Denlinger).

structural characteristics: spanning over 100 kb, encoding at least 3 transcripts with alternative splicing, lacking an AUG start codon, and probably encoding three proteins (Bellen et al., 1992a). Interestingly, in *Drosophila*, besides being present in the peripheral nervous system, *cpo* is also expressed in larval ring glands (Harvie et al., 1998), which suggests that this gene is also likely involved in neuroendocrine activity during larval development.

In many ways, the hypoactivity associated with the mutated *cpo* gene is reminiscent of the dormancy observed during diapause in adults of *D. melanogaster* (Saunders et al., 1989) and other species. The experiments by Schmidt et al. used quantitative trait mapping to identify this gene as the locus controlling the diapause phenotype (Schmidt et al., 2008). An amino acid substitution (Lys/Ile) is suggested to cause diapause phenotype variation in different geographic populations, and manipulation of the *cpo* locus by complementation analysis predictably alters the diapause phenotype, suggesting that the *cpo* locus is related to adult diapause in *D. melanogaster*. Furthermore, *cpo* transcripts were high in both larvae and adult females within a low-diapause incidence strain, while they were low in larvae and adult females within a high-diapause incidence strain (Schmidt et al., 2008). *Drosophila montana*, a species distributed at high latitudes over the northern hemisphere, also enters an adult diapause, and in this species *cpo* mRNA is significant up-regulated during diapause when measured by qPCR (Kankare et al., 2010). Interestingly, a *cpo* homolog in *Sarcophaga crassipalpis* is down-regulated during this fly's pupal diapause, based on microarray analysis (Ragland et al., 2010). Thus, *cpo* shows distinct responses to diapause in several species, although the direction of the response appears to not always be the same.

Our present set of experiments examines expression patterns of this gene in association with adult diapause of the northern house mosquito *Culex pipiens*. The diapause phenotype for the mosquito includes arrested ovarian development, increased lipid accumulation, elevated stress tolerance, and lack of host-seeking behavior (Bowen, 1991; Robich and Denlinger, 2005; Rinehart et al., 2006). We report the cloning of *cpo* cDNA from *C. pipiens*, describe the features of this gene, and investigate expression patterns in different developmental stages in both diapausing and nondiapausing adult females. The *C. pipiens cpo* mRNA is elevated during diapause, and when diapause is terminated naturally or by application of juvenile hormone *cpo* expression declines. The evidence we present in this study demonstrates that expression of *cpo* is affected by the diapause program in *C. pipiens*.

2. Materials and methods

2.1 Mosquitoes

Our colony of *Culex pipiens*, collected in Columbus, Ohio, has been maintained in our laboratory since 2000 (Robich and Denlinger, 2005). Approximately 250 newly hatched larvae were placed in 5×18×28 cm plastic containers at 25°C with a 16L:8D photoperiod, and larvae were fed ground fish food (TetraMin, Melle, Germany). For diapause induction third instar larvae were transferred to 18°C, 8L:16D, and they continued to be held at these conditions during pupal and adult stages. Adults were contained in 30.5×30.5×30.5 cm screened cages and supplied with honey sponges 10–13 days after eclosion. Nondiapause-destined larvae were also reared at 18°C but with a photoperiod of 16L: 8D. Honey sponges were continuously supplied to the nondiapausing adults. Diapause status was confirmed by checking length of the primary follicles in the ovaries and the ratio between the length of the follicle and that of the germarium as described (Sim and Denlinger, 2008).

2.2 Amplification of couch potato cDNA from adult *Culex pipiens*

Total RNA was isolated from whole bodies of 2–3 day-old adult females using Trizol® reagent (Invitrogen, CA). The quality and quantity of the RNA were examined using a NanoDrop 2000c spectrophotometer (Thermo Scientific, Wilmington, DE) and agarose gel electrophoresis. Total RNA (1.5 μg) was reverse transcribed by SuperScript™ II Reverse Transcriptase (Invitrogen, CA) and Oligo dT$_{18}$ according to the manufacturer's instructions. The synthesized first strand cDNA (1 μl) was used for PCR.

Degenerate primers CPP2 and CPP3 (Table 1) were designed using deduced *couch potato* homologous sequences reported in GenBank from other insects including *Drosophila melanogaster*, *Anopheles gambiae* and *Apis mellifera*. Using these two primers PCR was performed under the following conditions: 3 min at 94°C; 32 cycles of 40 sec at 94°C, 1 min at 45°C, 30 sec at 72°C; 7 min at 72°C. The PCR product was analyzed on 1.2% agarose gel and a band of approximately 174 bp was excised from the gel and centrifuged in an Ultrafree®- DA tube (Millipore, MA). The purified PCR fragment was ligated into pCR®2.1-TOPO® vector, transformed into TOP10 competent cells according to the manufacturer's instructions (Invitrogen, CA), and positive colonies were picked and sequenced at the Plant-Microbe Genomic Facility, Ohio State University.

Gene specific primers CPP4 and CPP5 (Table 1) were synthesized for 5'- and 3'-end RACE, respectively. cDNA was prepared from the total RNA mentioned above using a SMART™ RACE cDNA Amplification Kit (Clontech, CA) according to the manual. PCR conditions were as following: 94°C for 3 min; 35 cycles of 94°C for 30 sec; 68°C for 30 sec; 72°C for 5 min; extra extension at 72°C for 7 min. PCR products were separated on an agarose gel, and distinct bands were recovered, cloned, and sequenced as above.

2.3 Sequence alignment and molecular modeling of RNA recognition motif in the CPO C-terminus

The deduced Cupip-CPO amino acid sequence was aligned with CPO protein sequences retrieved from the NCBI database and VectorBase in five other insect species including *Anopheles gambiae* (ID: AGAP013145-PA), *Apis mellifera* (XP_392443.3), *Culex quinquefasciatus* (ID: XP_001862318.1), *Drosophila melanogaster* (ID: NP_732283.4), and *Tribolium castaneum* (ID: XP_968800.2). Only C-terminal regions of these sequences were subjected to Multalin for alignment.

Conserved domain search by a specialized BLAST program indicated typical RRMs (RNA recognition motif) in the deduced Cupip-CPO protein. Then, the sequence containing the RRM (aa87–aa176) was subjected to an online program for protein structure prediction, Protein Homology/analogY Recognition Engine (PHYRE) (Kelly and Sternberg, 2009). The predicted model with the highest accuracy score was adopted and visualized by DeepView/ Swiss-Pdb Viewer version 4.0.1 (Guex and Peitsch, 1997).

2.4 Northern and Southern blots

Based on the cloned *cpo* cDNA sequence, primers CPP8 and CPP9 (Table 1) were designed and synthesized to amplify a cDNA fragment as the template for probe labeling. DIG-High Prime (Roche, IN) was employed and a digoxigenin-dUTP labeled probe was produced for *cpo* Northern and Southern blots.

For Northern blots, total RNAs from day 1 first to fourth larval instars, day 1 pupae, and day 1 and day 7 adults were extracted and checked as above. Five μg of total RNA from each sample were loaded onto a 0.8% agarose/0.41 M formaldehyde gel and transferred to a Hybond-N$^+$ positively charged Nylon Membrane (GE Healthcare, UK) using a

Turboblotter™ Rapid Downward Transfer System (Whatman, NJ). For mRNA size determination, 1.5μg of mRNA purified from either diapausing or nondiapausing adult females was loaded for Northern blots as above.

For Southern blots, genomic DNA was extracted from newly-emerged male adults using GenomicPrep Cells and Tissue DNA Isolation Kit (Amersham Biosciences, UK). The quality and concentration of the isolated DNA were determined by agarose gel electrophoresis and by measuring absorbance with a UV spectrophotometer. Ten μg of genomic DNA was digested at 37°C for 5 h, using restriction endonucleases DraI, EcoRI, HindIII, RsaI, Sau3AI, and SspI. The digested DNA was separated by 0.8% agarose gel electrophoresis. The gel was treated in 0.25 M HCl for DNA fragmentation and in 0.4 M NaOH for neutralization. Then, DNA was transferred to a Hybond-N$^+$ positively charged nylon membrane. Probe hybridization, washing, and signal detection procedures were as described (Zhang and Denlinger, 2010).

2.5 Quantitative PCR

Total RNA was isolated from mosquito samples using Trizol® reagent (Invitrogen, CA), and concentration and absorbance values at both 260 nm and 280 nm were determined by spectrophotometer. 1.5 μg of RNA was treated by amplification-grade DNase I (Invitrogen, CA) and then reverse-transcribed using a SuperScript™ III first-strand synthesis system (Invitrogen, CA) according to the manufacturer's directions. The synthesized cDNAs were stored at −20°C until used.

Rpl19 and *28S rRNA* (GenBank No. FJ266017 and DQ401446) were first used as internal control genes for comparing gene expression levels between diapausing and nondiapausing individuals in *C. pipiens* (Sim and Denlinger, 2009), and we found that *rpl40* (GenBank No. DQ401463) was also a relatively stable gene among four potential control genes that were evaluated in diapause and nondiapause comparisons: Ribosomal protein L32 (*rp49*), Glyceraldehyde-3-phosphate dehydrogenase (*gapdh*), Putative poly A binding protein (*pbp*), and Ribosomal protein L40 (*rpl40*) in both diapausing and nondiapausing *C. pipiens* females (data not shown). Thus three genes (*rpl19*, *28S rRNA*, and *rpl40*) were used as controls for comparing *cpo* expression between diapausing and nondiapausing females and to monitor response to JH treatment. Primer sequences for *rpl19*, *rpl40*, *28S rRNA*, and *cpo* are shown in Table 1.

Target fragments of each gene were cloned and sequenced as above to verify sequence. Plasmids or cDNA samples were serially diluted 7 times using 10-fold or 5-fold dilutions to construct a standard curve for each gene. Based on standard curves, only those primers with efficiencies between 95 and 105%, $R^2 >0.995$, and showing only a single peak in the melt curve were employed for the next qPCR experiments.

Quantitative PCR was performed in 96-well plates with a 20 μl volume for each PCR reaction. The final concentration of each primer was 400 nM. iQ™SYBR® Green Supermix was used, and the PCR reaction was performed on an iQ™ 5 qPCR machine (Bio-Rad, Hercules, CA) under the following conditions: 3 min at 94°C; 40 cycles of 10 sec at 95°C, 30 sec at 60°C, 30 sec at 72°C. After qPCR, products were loaded on a 1% agarose gel to verify the presence of a single product.

Three biological replicates were collected for each sample, and 15 individuals were pooled for each replicate. Two or three technical replicates were performed for each biological replicate on the qPCR plates. The geNorm program was employed to compare expression stability of reference genes, and normalization factors were calculated for the target gene (Vandesompele et al., 2002). Expression levels of *cpo* were presented as the arithmetic

means of three biological replicates with standard deviations. Student's t-test was used to examine statistical differences between samples. P<0.05 and P<0.01 were considered significant and very significant differences respectively.

2.6 Juvenile hormone application to diapausing females

Juvenile hormone III (JH III) (Sigma-Aldrich, St. Louis, MO) was dissolved in acetone, and three doses (5, 50, and 500 ng) were topically applied (1 ul volume) to the thorax of two-week-old diapausing females. Three groups of mosquitoes including non-treated diapausing females at 18°C, acetone-treated diapausing females at 18°C, and nondiapausing females at 18°C were employed as controls. JH-treated and control mosquitoes were sacrificed one week after JH treatment for determination of *cpo* expression levels and to measure primary follicle lengths in the ovaries. Three replicates were employed for each sample, and fifteen individuals were collected for each replicate.

3. Results

3.1 Deduced Culex pipens CPO protein C-terminus has high homology with other CPOs

Using degenerate primers CPP2 and CPP3, we cloned and sequenced a 174-bp cDNA fragment that encodes 58 amino acid residues having 95% identity to the corresponding fragment in *Drosophila melanogaster* (NP_732283) CPO variant A, and 97% identity to a fragment of the same size in *Anopheles gambiae* (AGAP013145). Later, when the *Culex quinquefaciatus* genome was released, a 100% identical fragment was retrieved from the deduced protein sequence in CPIJ012174 reported in VectorBase (www.vectorbase.org). Thus, we conclude that the cloned cDNA fragment is a portion of the *cpo* gene in *C. pipiens*.

RACE was employed to obtain upstream and downstream sequences and a 3430-bp cDNA was obtained (GenBank No. HQ603004). The cloned cDNA contains an ORF encoding 252 amino acid residues (Fig.1). Other CPOs including *D. melanogaster* CPO isoform A (NP_732283), deduced CPOs from *Tribolium castaneum* (XP_968800) and *Apis mellifera* (XP_392443) were retrieved from the NCBI nucleotide database, and those of *Anopheles gambiae* (AGAP013145) and *C. quinquefaciatus* (CPIJ012174) were from VectorBase. Sequence analysis suggests that a canonical ribonucleoprotein consensus 2 (RNP2) sequence (LFVSGL), a ribonucleoprotein consensus 1 (RNP1) sequence (SPVGFVTF), and a typical RNA recognition motif (RRM) dimerization site (LEF) are present (Fig.1). Sequence alignment indicates that this protein has a highly conserved region in the C-terminus beginning with the RRM sequence (see Fig.1). In this region, the identity between *C. pipiens* and *C. quinquefaciatus* is 100% (167/167). It is 96% identical between *C. pipiens* and *A. gambiae* (160/167), 93% between *C. pipiens* and *D. melanogaster* (156/168), 90% between *C. pipiens* and *T. castaneum* (151/167), and 87% between *C. pipiens* and *A. mellifera* (123/141). These results indicate that CPOs share a highly homologous C-terminus, at least in the insect species reported thus far.

3.2 Structural prediction indicates presence of an RNA recognition motif (RRM)

Since RNP2 and RNP1 motif sequences were detected in the CPO sequences (Fig.1), the partial CPO sequence was subjected to an online protein structural prediction program, PHYRE, and a typical RNA recognition motif (RRM) was predicted (Fig.2). Between aa86 and aa174, four anti-parallel β-strands are packed against two α-helices forming a βαββαβ topological structure which is the conserved RRM structure found in many RNA binding proteins. When viewed from the surface of the sheet, the arrangement of these four β-strands from left to right is $\beta_4\beta_1\beta_3\beta_2$. In the β-sheet, a six amino acid residue RNP2 motif sequence (LFVSGL) forms the β_1 strand and an eight amino acid residue RNP1 motif sequence (SPVGFVTF) forms the β_3 strand. These two β strands form the major interface for RNA

binding, in which three conserved amino acid residues, F in RNP2 plus V and F in RNP1, are all available in Cupip-CPO as well as in the other CPOs compared here. Finally, three amino acid residues (LEF) form a conserved RRM dimerization site, suggesting that the RRM motif in CPO binds to other proteins with RRM motifs for complete function.

3.3 Developmental expression and size of cpo mRNA evaluated by Northern blot

Expression levels of *cpo* transcripts in nondiapause-destined *C. pipiens* were examined by total RNA Northern blots at seven developmental times from day 1 of the first larval instar to day 7 of adult life (Fig.3A). Highest expression was noted on day 1 of the first larval instar (L1), and expression continued to decline until reaching an undetectable level on day 1 of the fourth larval instar (L4). Expression was again high in day 1 pupae (P1), dropped by day 1 of adult life, and dropped further by day 7. These differences in *cpo* mRNA levels throughout development suggest that *cpo* functions differently at different developmental stages.

mRNA was purified from total RNA extracted from adult females on day 7 to determine the size of the *cpo* mRNA and to compare expression levels between diapausing and nondiapausing females. The *cpo* mRNA was slightly higher than the 9.49 kb RNA molecular weight marker, and only one band hybridized with the probe in either the diapausing or nondiapausing group (Fig.3B). The *cpo* mRNA level in diapausing adult females appeared slightly greater than that of nondiapausing females when the same amount of mRNA was loaded, thus prompting more detailed experiments using qPCR.

3.4 Southern blot analysis of the cpo gene

Southern blots were used to examine the genomic organization of the *cpo* gene. No cutting sites were present for DraI, HindIII, and SspI in the 567-bp probe cDNA; there was one cutting site for EcoRI and three cutting sites for both RsaI and Sau3AI. There were multiple bands in each lane despite the different numbers of cutting sites for different enzymes in the cDNA probe (Fig.4). The only difference was that the sizes of bands produced by low-selective enzymes such as RsaI (GTAC) and Sau3AI (GATC) were smaller than the others (most were over 2.5 kb). These results indicate that there are either multiple copies of *cpo* in the *C. pipiens* genome or there are large introns between the exons employed as hybridization probes.

3.5 qPCR reveals higher levels of cpo transcript in diapausing mosquitoes

Cpo expression levels were compared between long-day (nondiapause-destined) and short-day (diapause-destined) mosquitoes at various time points from the final larval instar to two months of age in adult females (Fig.5). There were no significant differences of *cpo* mRNA levels between short-day and long-day individuals on day 1 of the final larval instar, in day 2 pupae, or in day 2 adult females ($P>0.05$, student's t-test). But as adult development progressed, a significant difference between the short-day and long-day groups appeared in day 7 adult females ($P<0.05$, student's t-test), and this difference became very significant by day 14 ($P<0.01$, student's t-test) and reached the highest level of significance by day 28 ($P<0.001$, student's t-test), with short-day females having 3.4-fold more *cpo* mRNA than long-day females. Even on day 60, the difference of *cpo* mRNA between short-day and long-day females was still very significant ($P<0.01$, student's t-test). When two month-old short-day adult females were transferred from 18°C, 8L:16D to 25°C, 16L:8D for diapause termination, the *cpo* mRNA level dropped sharply ($P<0.01$, student's t-test); nondiapausing adult females also received the same treatment, but no significant change in *cpo* transcript level was noted. These results suggest that *cpo* mRNA is maintained at a higher level in short-day females than in long-day females after day 7, and diapause termination prompts an obvious decrease in the *cpo* transcript level. In addition, the *cpo* mRNA levels in pupae were

6–8 fold higher than in the final larval instar, a result consistent with the qualitative results observed using Northern blots.

3.6 Juvenile hormone application terminated diapause and reduced cpo expression

Primary follicle length, a marker of diapause status, and *cpo* expression were evaluated 7 days after three different doses of JH III were topically applied in acetone to two week-old diapausing adult females (Fig.6). Acetone-treated controls (Acetone, 1 ul) did not differ from non-treated diapausing controls (D). *Cpo* expression in response to 5 ng JH III (JH5) or 50 ng JH III (JH50) did not differ from the controls, but the highest dose used, 500 ng (JH500), caused a significant decrease in levels of the *cpo* transcript. The lowest *cpo* mRNA level was observed in adult females that never entered diapause (ND). JH III also impacted the length of the primary follicles in a similar manner. While follicle length was not different in the untreated or acetone-treated controls, or in the JH5 and JH50 groups, there was a very significant increase in the length of primary follicles in the JH500 group, and follicle length in the JH500 group was not significantly different from that observed in the nondiapausing group. Thus, 500 ng of topically-applied JH III effectively terminated diapause in *C. pipiens*, as noted in previous experiments (Sim and Denlinger, 2008), and significantly decreased the level of *cpo* mRNA, but 5 or 50 ng JH III were not sufficient to elicit this response.

4. Discussion

Cpo was first identified from *D. melanogaster* as an important factor caused hypoactive activity in adults (Bellen et al., 1992a; 1992b), and recently it was proposed as an essential locus controlling adult reproductive diapause in *D. melanogaster* (Schmidt et al., 2008). In the present study, we cloned the *cpo* gene from *C. pipiens* and compared expression of this gene in diapausing and nondiapausing adult females. We report that *cpo* expression is higher in diapausing adult females than in corresponding nondiapausing individuals maintained at the same temperature. Diapause termination, caused by either environmental signals or a topical application of JH III, decreases expression of the *cpo* transcript.

The gene structure and transcripts of *cpo* in *D. melanogaster* and its homologs in humans are large and complex (Bellen et al., 1992a; Shimamoto et al., 1996). In *D. melanogaster*, the *cpo* locus spans over 100 kb with huge introns of 18 and 42.7 kb (Bellen et al., 1992a), encoding 7 different mRNA isoforms (reported in NCBI nucleotide database). The homolog in humans, RBP-MS (RNA-binding protein gene with multiple splicing), is a single-copy gene spanning over 230 kb and is alternatively spliced to produce at least 12 transcripts (Shimamoto et al., 1996). Interestingly, when searching VectorBase, two *cpo* homologs from *C. quinquefaciatus* (CPIJ012174) and *Anopheles gambiae* (AGAP013145) were retrieved, but the sizes and genomic structures of *cpo* from these two species are quite different. In *A. gambiae*, the *cpo* locus spans 124.18 kb containing 9 exons and produces 3 alternative splicing mRNAs (AGAP013145-RA, AGAP013145-RB, and AGAP013145-RC), while in *C. quinquefaciatus*, the *cpo* locus spans only 7.52 kb containing 4 exons and produces only one transcript (CPIJ012174-RA). We performed Southern blot analysis of the *C. pipienscpo* gene, and multiple bands of large size produced by different enzymes suggest that either *cpo* is a multiple copies gene or it has huge introns as noted in *cpo* from *D. melanogaster*. Since only one copy of the *cpo* gene was verified in humans and reported in the insect species mentioned above, we assume that only one copy of *cpo* containing large introns is present in *C. pipiens*, and likely the *C. quinquefaciatuscpo* locus is not yet completely annotated.

Another unique characteristic of *cpo* in *D. melanogaster* is that there are 422 amino acid residues in front of the first Met, thus suggesting that the start codon for CPO proteins is not

AUG (Met) but CUC (Leu) (Bellen et al., 1992a) in all seven mRNA isoforms. Interestingly, in *A. gambiae*, two of the three CPO proteins have the same characteristic in that they have 168 amino acid residues in front of the first Met, while the third isoform begins with Met as shown in Fig.1. Although we observed only one band in Northern blots, with RACE we cloned 5 different *cpo* cDNAs that encode five proteins with a stretch of amino acid residues having different lengths (from 89–191) in front of the first Met (data not shown), but all contained the consensus sequence reported in Fig.1. In the amino acid regions in front of Met, stretches of Ala and Glu are present in the CPOs from all three species. However, in *C. quinquefaciatus* the deduced CPO protein is shorter than that from either *C. pipiens* or *A. gambiae*, and at the N-terminus the first 8 amino acid residues are completely different from the other sequences (Fig.1), while the remainder of the CPO sequence is 100% identical to *C. pipiens* and 96% identical to *A. gambiae*. One possible explanation for this interesting phenomenon is that *cpo* is evolving quickly, leading to multiple molecular structures in different species. But, considering that the Northern blot results in *C. pipiens* indicate that the size of the *cpo* mRNA is over 9.49 kb, it is also possible that *cpo* molecules are huge and the transcripts are so large that most reports or annotations do not reach the real 5'-ends of the *cpo* transcripts. More experiments are needed to decode the complicated structure of *cpo* genes and transcripts in different insect species.

RRM is an abundant domain in higher vertebrates and is present in ~0.5–1% of human genes. A typical RRM has RNP1 and RNP2 motif sequences; it forms a βαββαβ topological structure, and the β-sheet formed by four anti-parallel β-strands is the major interface for single-stranded RNA binding (Cléry et al., 2008). These essential elements are all present in both the *C. pipiens* CPO protein and in the other insect CPOs compared here. This gene is so well conserved that the sequences in this region are nearly the same (see Fig.1). In addition, all have a RRM dimerization site, which possibly binds to RRM structures of other proteins or even to itself. Interestingly, the sequences from different insect species after the RRM are also quite well conserved, especially the last 21 amino acid residues, which are completely the same. Although no conserved domain was detected in this region, CPO most likely carries out some unknown but identical function in the different species, such as single-strand RNA binding or transportation. More structural information is needed from the N-terminal of CPO to fully understand its biological function.

The couch potato-like behavior of adult *D. melanogaster* is caused by P-element insertions at the *cpo* locus (Bellen et al., 1992a; 1992b; Glasscock and Tanouye, 2005). In many respects this suppression of activity is reminiscent of the adult diapause caused by short daylength and low temperature in the Canton-S strain of *D. melanogaster* (Saunders et al., 1989). Indeed, *cpo* was recently proposed as a major genetic locus determining diapause in this species (Schmidt et al., 2008). Variation in the incidence of diapause among different geographic populations was linked to a single amino acid mutation. The relationship between *cpo* expression and diapause incidence was also investigated by hypomorphic alleles, gene duplications/deletions, and RT-PCR, and results consistently indicated that *cpo* mRNA was lower in high-diapause lines than in low-diapause lines (Schmidt et al., 2008).

Experiments comparing *cpo* expression in different populations of *D. melanogaster* are, of course, fundamentally different than a comparison made between diapausing and nondiapausing individuals within a population, as described here. The comparison that is most similar to our study is the diapause and nondiapause comparison of adults of *D. montana* (Kankare et al., 2010), and in both of these cases *cpo* expression is higher in diapausing individuals. But, this does not appear to be so for pupal diapause in the flesh fly, *Sarcophaga crassipalpis*; a recent microarray analysis suggests the opposite response: low *cpo* expression in diapausing pupae (Ragland et al., 2010). Thus, differences may exist among species and/or developmental stages, but in each case distinct patterns of *cpo*

expression have been linked to diapause, suggesting that this is a gene of interest for diapause in diverse species.

Acknowledgments

References

Bellen HJ, Kooyer S, D'Evelyn D, Pearlman J. The *Drosophila* couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. Genes and Development. 1992a; 6:2125–2136. [PubMed: 1427076]

Bellen HJ, Vaessin H, Bier E, Kolodkin A, D'Evelyn D, Kooyer S, Jan YN. The Drosophila *couch potato* gene: an essential gene required for normal adult behavior. Genetics. 1992b; 131:365–375. [PubMed: 1644278]

Bennett-Lovsey RM, Hebert AD, Sternberg MJE, Kelley LA. Exploring the extremes of sequence/ structure space with ensemble fold recognition in the program Phyre. Proteins: Structure, Function, Bioinformatics. 2008; 70:611–625.

Bowen MF. The sensory physiology of host-seeking behavior in mosquitoes. Annual Review of Entomology. 1991; 36:139–158.

Cléry A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. Current Opinion in Structural Biology. 2008; 18:290–298. [PubMed: 18515081]

Glasscock E, Tanouye MA. *Drosophila* couch potato mutants exhibit complex neurological abnormalities including epilepsy phenotypes. Genetics. 2005; 169:2137–2149. [PubMed: 15687283]

Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 1997; 18:2714–2723. [PubMed: 9504803]

Harvie PD, Filippova M, Bryant PJ. Genes expressed in the ring gland, the major endocrine organ of *Drosophila melanogaster*. Genetics. 1998; 149:217–231. [PubMed: 9584098]

Kankare M, Salminen T, Laiho A, Vesala L, Hoikkala A. Changes in gene expression linked with adult reproductive diapause in a northern malt fly species: a candidate gene microarray study. BMC Ecology. 2010; 10:3. [PubMed: 20122138]

Kelly LA, Sternberg MJE. Protein structure prediction on the web: a case study using the Phyre server. Nature Protocol. 2009; 4:363–371.

Ragland GJ, Denlinger DL, Hahn DA. Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:14909–14914. [PubMed: 20668242]

Rinehart JP, Robich RM, Denlinger DL. Enhanced cold and desiccation tolerance in diapausing adults of *Culex pipiens*, and a role for Hsp70 in response to cold shock but not as a component of the diapause program. Journal of Medical Entomology. 2006; 43:713–22. [PubMed: 16892629]

Robich RM, Denlinger DL. Diapause in the mosquito *Culex pipiens* evokes a metabolic switch from blood feeding to sugar gluttony. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:15912–15917. [PubMed: 16247003]

Saunders DS, Henrich VC, Gilbert LI. Induction of diapause in *Drosophila melanogaster*: photoperiodic regulation and the impact of arrhythmic clock mutations on time measurement. Proceedings of the National Academy of Sciences of the United States of America. 1989; 86:3748–3752. [PubMed: 2498875]

Schmidt PS, Zhu CT, Das J, Batavia M, Yang L, Eanes WF. An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:16207–16211. [PubMed: 18852464]

Shimamoto A, Kitao S, Ichikawa K, Suzuki N, Yamabe Y, Imamura O, Tokutake Y, Satoh M, Matsumoto T, Kuromitsu J, Kataoka H, Sugawara K, Sugawara M, Sugimoto M, Goto M, Furuichi Y. A unique human gene that spans over 230 kb in the human chromosome 8p11–12 and codes multiple family proteins sharing RNA-binding motifs. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93:10913–10917. [PubMed: 8855282]

Sim C, Denlinger DL. Insulin signaling and FOXO regulate the overwintering diapause of the mosquito *Culex pipiens*. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:6777–6781. [PubMed: 18448677]

Sim C, Denlinger DL. Transcription profiling and regulation of fat metabolism genes in diapausing adults of the mosquito *Culex pipiens*. Physiological Genomics. 2009; 39:202–209. [PubMed: 19706691]

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology. 2002; 3 research0034.

Zhang Q, Denlinger DL. Molecular characterization of heat shock protein 90, 70 and 70 cognate cDNAs and their expression patterns during thermal stress and pupal diapause in the corn earworm. Journal of Insect Physiology. 2010; 56:138–150. [PubMed: 19782689]
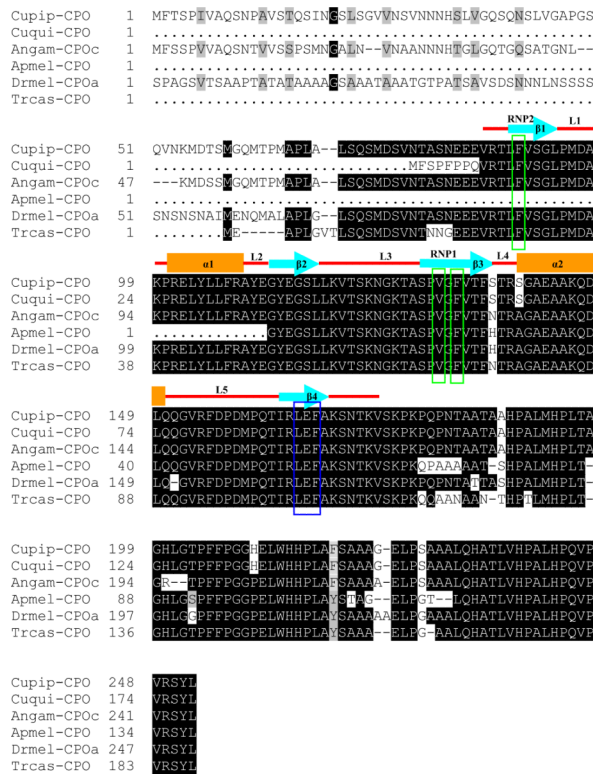
Figure 1.
Sequence alignment of deduced CPO amino acid sequences at the conserved C-terminal and a schematic diagram of the RNA recognition motif presented above the sequences. Sequence of *Culex pipiens* CPO (Cupip) was compared with CPOs from *Culex quinquefasciatus* (Cuqui), *Anopheles gambiae* (Angam), *Apis mellifica* (Apmel), *Drosophila melanogaster* (Drmel), and *Tribolium castaneum* (Trcas). Conserved amino acid residues are indicated as white letters on a black background; similar residues are indicated as white letters on a gray background; non-conserved residues are indicated as black letters on a white background; gaps, indicated as dashes, are inserted to optimize the alignment; dots indicated unavailable amino acid residues. The numbers of amino acid residues are designated arbitrarily for sequence alignment and are not their exact positions in the full sequences. Amino acid residues comprising the RRM are labeled above as follows: blue arrows indicate β-strands and are labeled β1, β2, β3, and β4; orange squares indicate α-helices and are labeled α1 and α2; loops connecting them are indicated in red. β1 and β3 are the conserved RNP2 motif (LFVSGL) and RNP1 motif (SPVGFVTF), respectively. Three conserved residues F, V, and F in these two motifs are boxed in green. Three amino acid residues (LEF) that form the conserved RRM dimerization site are boxed in purple.
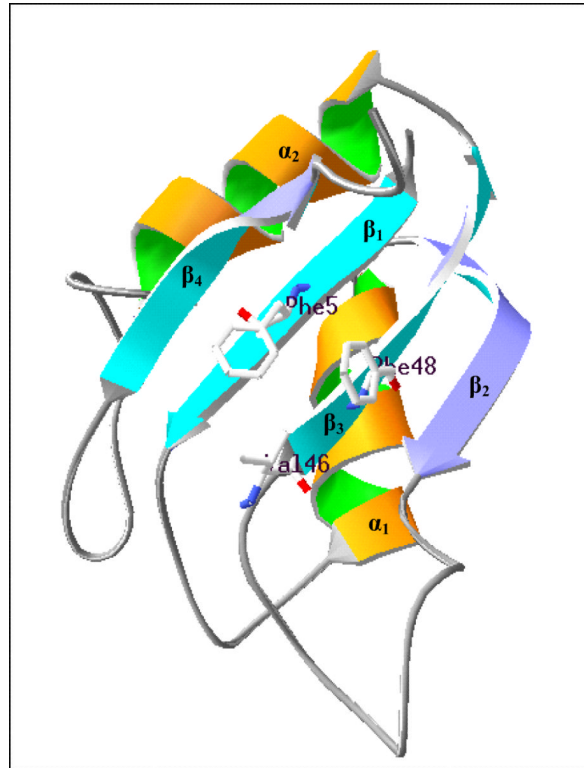
Figure 2.
Three-dimensional structure prediction of the conserved RNA recognition motif (RRM) component in Cupip-CPO (aa86 - aa174 in Fig.1) indicates that it is a typical βαββαβ topological structure. Blue and purple arrows indicate an anti-parallel β-sheet consisting of four β-strands pointing from N-terminal to C-terminal. Orange and green helices are α-helices and the gray threads indicate connecting coil loops. From the front, the order of the β-sheets is $\beta_4\beta_1\beta_3\beta_2$ from left to right. β_1 and β_3 are conserved RNP2 and RNP1 motifs respectively and form the major RNA recognition site. Three conserved residues in β_1 and β_3 are labeled Phe5, Val46 and Phe48. Predicted by Protein Homology/analogY Recognition Engine (PHYRE) and viewed by Deep View / SWISS-Pdb Viewer v4.0.1.
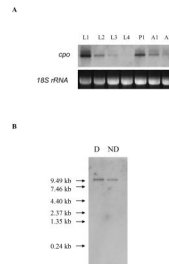
Figure 3.
Northern blot analysis of *cpo* mRNA. A: Developmental expression pattern of *cpo* transcripts from day 1 of the first larval instar (L1), day 1 of the second larval instar (L2), day 1 of the third larval instar (L3), day 1 of the fourth larval instar (L4), day 1 pupae (P1), day 1 adult females (A1), and day 7 adult females (A7). Mosquitoes were reared under nondiapause condition at 25°C with a photoperiod of 16 L: 8 D. 5 µg total RNA from each sample was loaded and 18 S rRNA was employed as a loading control. B: Comparison of *cpo* mRNA in 7 day-old diapausing (D) and nondiapausing (ND) adult females. 1.5 µg mRNA from each sample was loaded and sizes of RNA molecular markers are shown at left.
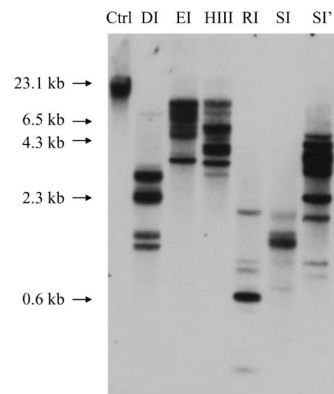
Figure 4.
Southern blot analysis of the *cpo* gene. Ten µg genomic DNA from adult males was digested by 50 units' restriction endonucleases DraI (DI), EcoRI (EI), HindIII (HIII), RsaI (RI), Sau3AI (SI), and SspI (SI') respectively; multiple bands were observed in each lane. Non-digested DNA was used as a control (Ctrl); sizes of DNA molecular weight markers are shown at left.
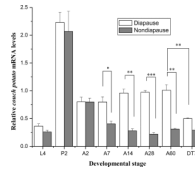
Figure 5.
Expression pattern of *cpo* mRNA by qPCR in short-day (diapause-destined, D) and long-day (nondiapause-destined, ND) mosquitoes in day 1 fourth instar larvae (DL4 and NDL4), day 2 pupae (DP2 and NDP2), day 2 (DA2 and NDA2), 7 (DA7 and NDA7), 14 (DA14 and NDA14), 28 (DA28 and NDA28), and 60 (DA60 and NDA60) adult females. Day 60 diapausing females were transferred to 25°C and long daylength (16L: 8D) for 7 days to break diapause, and nondiapausing females were also transferred as controls (D18–25 and ND18–25). Error bars represent standard deviations of three replicates; 15 individuals were sampled in each replicate. Single asterisk indicates significant difference, with double asterisks and triple asterisks indicating very significant differences between the two tested groups, respectively (*P<0.05, **P<0.01, ***P<0.001, student's t-test).
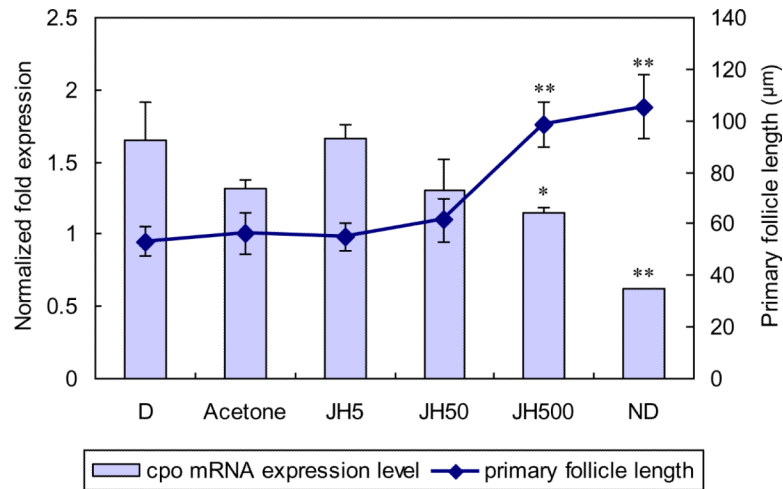
Figure 6.
JHIII application terminated *C. pipens* diapause as indicated by an increase in follicle length and decreased expression level of *cpo* mRNA evaluated by qPCR. Day 14 diapausing adult females were treated with JHIII at three doses: 5 ng (JH5), 50 ng (JH50), and 500 ng (JH500). Day 14 diapausing adult females (D), nondiapausing adult females (ND), and acetone-treated day 14 diapausing adult females (Acetone), were employed as controls. Expression levels of *cpo* and primary ovarian follicle lengths were examined 7 days after treatment. Primary follicle lengths are shown as blocked diamonds and expression levels of *cpo* as columns. Error bars represent standard deviations of three replicates; 15 individuals were sampled in each replicate. Single asterisk and double asterisks indicate significant difference and very significant difference, respectively (*P<0.05, **P<0.01, student's t-test).

Table 1

Primers used for cloning, probe production, and qPCR.

Primer sequence	Purpose	Amplicon size	Efficiency for qPCR(%)
CPP2: (forward) 5 '-GAAGGTTCACTGT/CTGAAA/GGT-3'	For cDNA internal fragment	174 bp	---
CPP3: (reverse) 5'-GTTACTT/CTTGGCA/GAA TTCC-3'			
CPP4: (reverse) 5'-CGTACACCTTGCTGTAGATCCT-3'	For 5'-RACE	---	---
CPP5: (forward) 5'-TTCGTCACGTTCAGCACACGGT-3'	For 3'-RACE	---	---
CPP8: (forward) 5'-TCAGTACACAGTCCATCAACGG-3'	For hybridization probe	567 bp	---
CPP9: (reverse) 5'-TACCTAGGTGTCCAGCAGTAAG-3'			
*qRPL19P1 (forward) 5'-GCTTTGTTTGATCGTGTGTGA-3'	For *rpl19* qPCR	105 bp	101.4
*qRPL19P2 (reverse) 5'-AACATATCTCCCCCAATCCAG-3'			
*q28SP1 (forward) 5'-ACGTGAAACTGCCTAGGGCTC-3'	For *28SrRNA* qPCR	200 bp	97.8
*q28SP2 (reverse) 5'-TGGGACAAGCAACCAGATG-3'			
*qCPOP1 (forward) 5'-TTCGTCACGTTCAGCACACGGT-3'	For *cpo* qPCR	119bp	102.4
*qCPOP2 (reverse) 5'-GTGTTACTCTTGGCGAATTCC-3'			
qRPL40P1 (forward) 5'-ACTGCCGCAAGAAGAAGTGT-3'	For *rpl40* qPCR	104 bp	98.2
qRPL40P2 (reverse) 5'-GGAATCAATCCAGAAGCACAA-3'			

*indicates that the primer sequence was previously reported in Zhang and Denlinger (2011).