Structural Analysis and Identification of Novel Isoforms of the Circadian Clock Gene period in the Silk Moth Bombyx mori

Authors: Takeda, Yukimasa, Chuman, Yoshiro, Shirasu, Naoto, Sato, Seiji, Matsushima, Ayami, et al.

Source: Zoological Science, 21(9) : 903-915

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.21.903

ZOOLOGICAL SCIENCE **21**: 903–915 (2004)

Structural Analysis and Identification of Novel Isoforms of the Circadian Clock Gene *period* in the Silk Moth *Bombyx mori*

Yukimasa Takeda[1†], Yoshiro Chuman[1†], Naoto Shirasu[1], Seiji Sato[1],
Ayami Matsushima[1], Atsushi Kaneki[1], Yoshiya Tominaga[2],
Yasuyuki Shimohigashi[1*] and Miki Shimohigashi[2*]

[1] *Laboratory of Structure-Function Biochemistry, Department of Chemistry, Faculty and Graduate School of Sciences, Kyushu University, Fukuoka 812-8581, Japan*
[2] *Division of Biology, Faculty of Science, Fukuoka University, Fukuoka 814-0180, Japan*

ABSTRACT—The molecular basis of the circadian clock is an autoregulatory feedback loop in which the PAS domain-containing protein PERIOD periodically inhibits its own transcription. In the present study on PERIOD of the silk moth *Bombyx mori*, we have cloned two distinct *period* mRNA homologues with different PAS domain sequences either with or without the pentapeptide GTQEK. A *period* cDNA fragment first amplified by PCR exhibited a 15 bp-deleted nucleotide sequence in the PAS domain, compared with the database sequence. A possible alternative splicing mechanism was examined by PCR analyses, and a 15 bp-inserted clone was also amplified. The entire sequences of these *period* α and *period* β isoforms were then determined by the 3' and 5' RACE methods. Isoform *period* α consists of a 3,324 bp oligonucleotide encoding 1,108 amino acid residues, whereas isoform *period* β comprises 3,309 bp corresponding to 1,103 amino acids. Isoforms *period* α and *period* β were found to be exactly identical except for the 15 bp deletion/insertion site. Such a pair of isoforms with a deletion/insertion sequence, namely two splice variants, has previously been reported only for the PERIOD proteins of the two honeybees, *Apis mellifera* and *A. cerana*. The occurrence of an alternative splicing mechanism in the *B. mori period* gene was hypothesized based on the genome structure recently clarified. *Bombyx mori* PERIOD α and β proteins are the isomers that reveal firstly the different PAS domain sequences.

Key words: circadian rhythm, cDNA cloning, alternative splicing, isoforms

---

INTRODUCTION

Circadian clocks are endogenous timing mechanisms that generate daily rhythms in diverse organisms from cyanobacteria to humans. Although rhythms controlled by circadian clocks are self-sustained and persist robustly with a period close to 24 hours under conditions of constant darkness and temperature, they are entrained by environmental time cues (zeitgebers) such as light, temperature, and food intake (Soriano, 1981; Rusak *et al*., 1993). The *period* gene has been demonstrated to play a crucially important role in circadian rhythms of both eclosion behavior and locomotor activity in the fruit fly *Drosophila melanogaster* (Konopka and Benzer 1971), the giant silk moth *Antheraea pernyi* (Sauman *et al*., 1996), and the silk moth *Bombyx mori* (Itoh *et al*., 1995). The major role played by *period* genes has been clarified by analysis of various types of mutants.

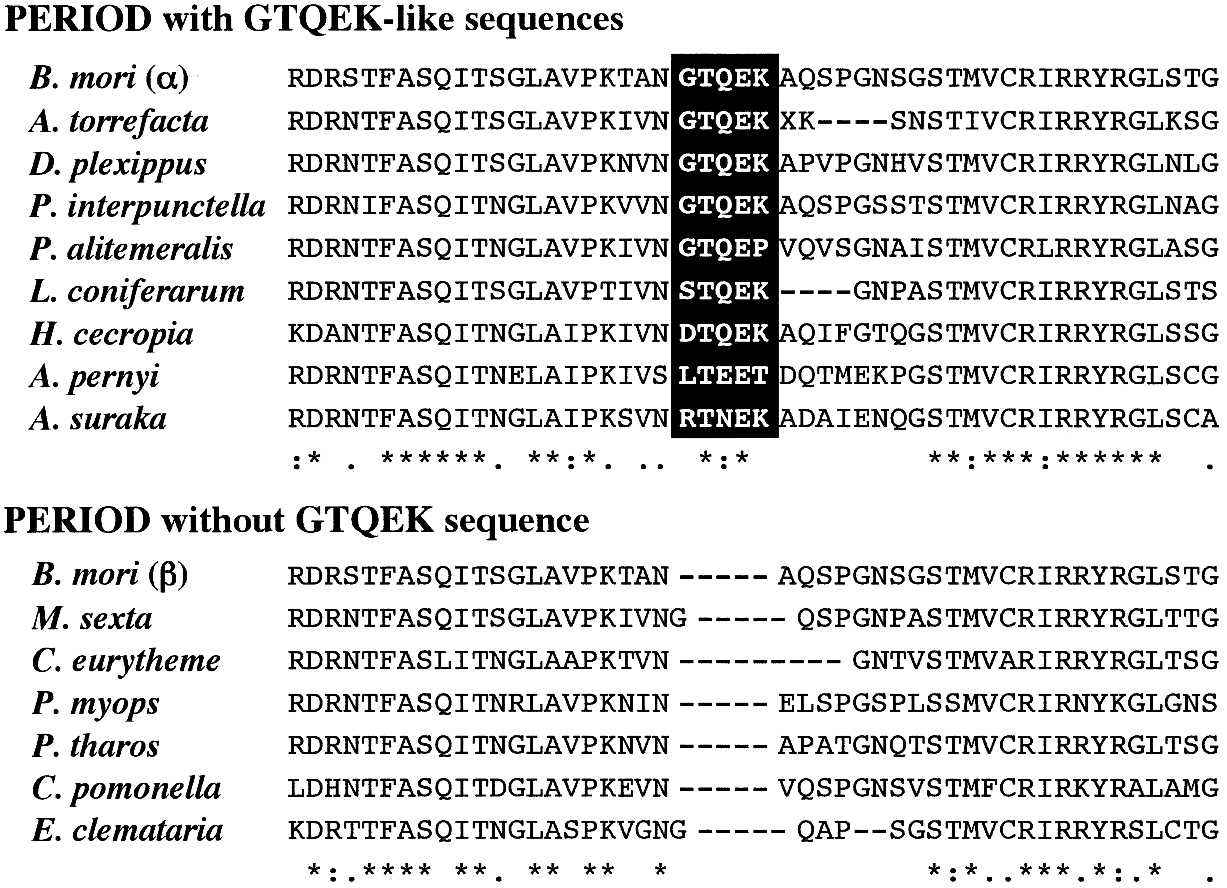
Circadian oscillations of a wide variety of organisms from insects to vertebrates are controlled through autoregulatory feedback loops involving PERIOD gene expression (Dunlap, 1999). In *Drosophila*, the lateral neurons (LNs), candidates for the fly's circadian pacemaker, in addition to *period* also express a series of other clock genes, such as *timeless*, *clock*, and *cycle* (Shirasu *et al*., 2003). Each corresponding protein except for TIMELESS produced from *timeless*, namely, PERIOD, CLOCK, and CYCLE, respectively, contains a so-called PAS domain that enables the molecule to construct rigid heterodimers. The PERIOD/TIMELESS dimer is translocated from the cytoplasm to the nucleus to function as a negative regulator of its own positive transcription factor CLOCK/CYCLE (Saez and Young, 1996; Lee *et al*., 1999).

The importance of the PAS domain for dimerization has

---

* Corresponding author: Tel. +81-92-642-2584;
FAX. +81-92-642-2584.
E-mail: shimoscc@mbox.nc.kyushu-u.ac.jp

[†] These authors contributed equally to this work.



PERIOD with GTQEK-like sequences

```
B. mori (α)       RDRSTFASQITSGLAVPKTAN GTQEK AQSPGNSGSTMVCRIRRYRGLSTG
A. torrefacta     RDRNTFASQITSGLAVPKIVN GTQEK XK----SNSTIVCRIRRYRGLKSG
D. plexippus      RDRNTFASQITSGLAVPKNVN GTQEK APVPGNHVSTMVCRIRRYRGLNLG
P. interpunctella RDRNIFASQITNGLAVPKVVN GTQEK AQSPGSSTSTMVCRIRRYRGLNAG
P. alitemeralis   RDRNTFASQITNGLAVPKIVN GTQEP VQVSGNAISTMVCRLRRYRGLASG
L. coniferarum    RDRNTFASQITSGLAVPTIVN STQEK ----GNPASTMVCRIRRYRGLSTS
H. cecropia       KDANTFASQITNGLAIPKIVN DTQEK AQIFGTQGSTMVCRIRRYRGLSSG
A. pernyi         RDRNTFASQITNELAIPKIVS LTEET DQTMEKPGSTMVCRIRRYRGLSCG
A. suraka         RDRNTFASQITNGLAIPKSVN RTNEK ADAIENQGSTMVCRIRRYRGLSCA
                  :* . ******. **:*. ..  *:*          **:***:******  .
```

PERIOD without GTQEK sequence

```
B. mori (β)       RDRSTFASQITSGLAVPKTAN ----- AQSPGNSGSTMVCRIRRYRGLSTG
M. sexta          RDRNTFASQITSGLAVPKIVNG ----- QSPGNPASTMVCRIRRYRGLTTG
C. eurytheme      RDRNTFASLITNGLAAPKTVN --------- GNTVSTMVARIRRYRGLTSG
P. myops          RDRNTFASQITNRLAVPKNIN ----- ELSPGSPLSSMVCRIRNYKGLGNS
P. tharos         RDRNTFASQITNGLAVPKNVN ----- APATGNQTSTMVCRIRRYRGLTSG
C. pomonella      LDHNTFASQITDGLAVPKEVN ----- VQSPGNSVSTMFCRIRKYRALAMG
E. clemataria     KDRTTFASQITNGLASPKVGNG ----- QAP--SGSTMVCRIRRYRSLCTG
                   *:.**** **. ** **  *               *:*..***.*:.*  .
```

Fig. 1. Comparison of the primary structures of Lepidoptera PERIOD proteins aligned around the GTQEK-like sequence. The amino acid sequences in a black box represent GTQEK-like sequences. '*' Residues in that column are identical in all sequences in the alignment. '**:**' Conserved substitutions have been observed. '**.**' Semi-conserved substitutions have been observed.

been shown by mutations at the PER^L site and PAS B repeat regions, both of which substantially decreased dimerization efficiency (Huang *et al.*, 1993; Gekakis *et al.*, 1995). The PAS domain is an important signaling module not only for protein-protein interaction but also for monitoring several environmental changes such as light, oxygen, and redox potential (Getzoff, 2002). The PAS domain itself would be exposed in the isolated PERIOD protein, but should be covered by the binding partner protein when PERIOD is in a dimer configuration.

In addition to dimerization with TIMELESS, the PERIOD protein may also interact with CLOCK or CYCLE to suppress the transcriptional function of the CLOCK/CYCLE dimer. PERIOD may also exist as a homodimer, though no physiological evidence has yet been reported on its function in that configuration (Huang *et al.*, 1993). It is thus likely that PERIOD possesses a molecular contrivance to select a counterpart protein. In our effort to identify the clock genes and their proteins in the silk moth *Bombyx mori*, we initiated the cDNA cloning of the clock protein PERIOD. When the region corresponding to the PAS domain was amplified by PCR, a 15 bp-deleted nucleotide sequence was identified with respect to the sequence reported in the GenBank database (accession number AF063429). This 15 bp nucleotide sequence corresponds to the amino acid sequence GTQEK. PAS domain sequences are highly conserved among the various insect PERIOD homologues (Regier *et al.*, 1998). When we aligned the sequence of PAS domains very carefully, however, the domains were classified into two groups: namely, PAS domains with GTQEK or a related pentapeptide and PAS domains lacking such a fragment (Fig. 1). All these results strongly suggested the occurrence of an alternative splicing mechanism in the *B. mori period* gene. Thus, in the present study, we carried out a full sequence analysis of possible alternative splicing mRNA products to confirm the presence of PERIOD isoforms that possess PAS domain with either a deletion or an insertion of a GTQEK pentapeptide. Here we report the cDNA cloning of two splice variants of *B. mori* PERIOD protein and their circadian profile of accumulation.

MATERIALS AND METHODS

Animals

Silk moths (female) were obtained from Kyoya (Kyoto), and maintained in a light/dark LD 17:7 photoperiod at 25°C. To collect samples for cDNA cloning, *B. mori* brains and optic lobes were excised at ZT17 and ZT9 (ZT: Zeitgeber time denotes the time entrained by environmental time cues, whereby lights on is ZT0 and lights off is ZT17). The samples were frozen in liquid nitrogen and stored at –80°C until use.

cDNA cloning with degenerate PCR

mRNA was purified using a QuickPrep® *Micro* mRNA Purification Kit (Amersham Biosciences; Piscataway, NJ, USA) according to the manufacturer's instructions. For the extraction of mRNA, frozen brains and optic lobes described above were homogenized in a buffer solution containing guanidinium thiocyanate. The extract was diluted by 10 mM Tris-HCl (pH 7.5) containing 1 mM EDTA, then the solution was incubated with oligo deoxythimidylic acid dT$_{17}$-attached cellulose to bind the poly A of mRNA molecules. After the cellulose was washed, extracted mRNA was reverse transcribed into cDNA by dT$_{17}$-adapter primer (5'-GGCCACGCGTCGACTAGTAC-dT$_{17}$-3') using AMV reverse transcriptase (Promega, Madison, WI, USA) as previously described (Frohman *et al.*, 1988). After phenol extraction followed by ethanol precipitation, the resulting single strand cDNA was used as a template for subsequent PCR. The PCR mixture (100 µl) included 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl$_2$, 0.2 µM of each primer, 0.2 mM dNTP mix, and 2.5 units of DNA polymerase.

PCR was carried out by using a sense primer PAS-AF (5'-GGVMGDTCVTTYATHGAYTTTYGTKCA-3') and an antisense primer PAS-CR (5'-TCRTGRTGNGGHGRDATHYCDCCVAGYAT-3'). These primers were designed with reference to highly conserved amino acid sequences of insect PERIODs aligned for the so-called C2 region (Colot *et al.*, 1988). In these primers, the letters D, H, K, M, N, R, V, Y denote the nucleotides (not C), (not G), (G or T), (A or C), (any), (A or G), (not T), and (C or T), respectively. The reaction was performed on a Geneamp PCR system 2400 (Perkin Elmer, MA, USA) using PLATINUM *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA) with the following conditions: 3 min at 94°C followed by 35 cycles denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec, extension at 72°C for 90 sec, and final extension for 10 min at 72°C. Single-primer control PCRs were also carried out in parallel. Two primer-specific PCR products were electrophoresed on a 1% agarose gel, and recovered by phenol extraction followed by ethanol precipitation. The gel-purified PCR products were subcloned into pCR 2.1-TOPO cloning vector (Invitrogen).

3' RACE for identification of the 3' end of *Bombyx mori period* mRNA

To amplify the 3' end of *B. mori period* cDNA, we carried out 3' RACE as described by Frohman *et al.* (1988) by using the primers designed from the identified fragment sequence of *B. mori period* cDNA. Reverse-transcribed cDNAs obtained above were amplified with a series of sense primers and an antisense adapter primer (5'-GGCCACGCGTCGACTAGTAC-3'). Sense primers utilized are bmpersense1 (5'-ATCGCTGACGGCTACACTC-3'), bmpersense2 (5'-TTCGTTCATCCAAGGGACCG-3'), and bmpersense4 (5'-GGTTATGATATATGAAGAGAACGCTCC-3') (Fig. 2). Nested PCR was performed by the touchdown PCR method (Don *et al.*, 1991) with a combination of bmpersense1 primer and adapter primer in the first PCR followed by the subsequent second PCR with bmpersense2 primer and adapter primer using the first PCR product as a template. The first PCR was carried out by *Pfu Turbo*® hotstart DNA polymerase (Stratagene, La Jolla, CA, USA) with the following conditions: 3 min at 94°C followed by 3 cycles of 94°C for 60 sec, 62°C for 60 sec, and 72°C for 10 min, 4 cycles of 94°C for 60 sec, 61°C for 60 sec, and 72°C for 4 min, 8 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 4 min, 10 cycles of 94°C for 30 sec, 59°C for 30 sec, and 72°C for 4 min, and then 15 cycles of 94°C for 30 sec, 58°C for 30 sec, and finally 72°C for 4 min. The second PCR was done by PLATINUM® *Taq* DNA polymerase (Invitrogen) under the same conditions described for the first PCR. After electrophoresis, gel-purified PCR products were subcloned into pDrive cloning vector (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

5' RACE for identification of the 5' end of *Bombyx mori period* mRNA

For amplification of the 5' end of *B. mori period* cDNA, we performed 5' RACE using rapid amplification of cDNA ends (version 2.0: Invitrogen) mostly according to the manufacturer's protocol. First-strand cDNA was synthesized from mRNA isolated by using a gene specific primer bmperantisense1 (5'-GAGGTAGATGACATTTCCCTCTTC-3') (Fig. 1) and ThermoScript™ RNase H$^-$ Reverse Transcriptase (Invitrogen) at 60°C. The product was digested by RNase H to remove the original mRNA template. The remaining dNTP, primer, and contained proteins were separated from synthesized cDNA by using a GLASSMAX® DNA Isolation Spin Cartridge System (GIBCO BRL®; Rockville, MD, USA). A homopolymeric tail poly(C) was attached to the 5' end of purified cDNA by using terminal deoxynucleotidyl transferase and a substrate dCTP in a PCR-compatible buffer.

PCR amplification was accomplished using PLATINUM *Taq* DNA polymerase (Invitrogen) and gene specific antisense primers such as bmperantisense2 (5'-TGCTCAGACCTCTGTAACGTC-3'), bmperantisense3 (5'-CCGACACACCATAGTCGATCCAGA-3') (Fig. 1) and deoxyinosine-containing anchor primers provided in the kit (GIBCO BRL®). PCR products with bmperantisense1 and Abridged Anchor Primer (AAP) (5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3') was further amplified with bmperantisense2 and Abridged Universal Anchor Primer (AUAP) (5'-GGCCACGCGTCGACTAGTAC-3'). These nested PCR amplifications were carried out under the following conditions: 3 min at 94°C followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 90 sec. PCR products were then electrophoresed on an agarose gel and the resulting pure product was subcloned into pDrive cloning vector.

PCR amplifications for identifying *Bombyx mori period* isoforms

To identify spliced and non-spliced clones, two PCR strategies were systematically performed, focusing on the expected sequence differences between the isoforms. The two predicted gene isoforms were denoted as *period* α including a 15 bp oligonucleotide sequence that corresponds to the pentapeptide GTQEK, and *period* β which lacks this sequence (Fig. 2). In the first PCR, a set of synthetic oligonucleotides of bmperGTQEK primer and bmperantisense1 primer were used to amplify *period* α specifically. In the second PCR, a set of bmpersense3 (5'-GACCGTAGCACGTTCGCGTCACAG-3') and bmperantisense3 primers was used to amplify *period* α and *period* β simultaneously. These PCRs were carried out by PLATINUM® *Taq* DNA polymerase as follows: 3 min at 94°C followed by 35 cycles (29 cycles for the second PCR) of denaturation at 94°C for 30 sec, annealing at 56°C for 30 sec, and extension at 72°C for 30 sec. These PCR products were used for electrophoresis on a 2% agarose gel, and the gel-purified PCR products were subcloned into pBluescript II SK$^+$ cloning vector (Stratagene) for subsequent sequence analyses.

Quantitative analyses by RNase protection assay and RT-PCR for *Bombyx mori period* mRNA

Bombyx mori brains and optic lobes (10 each) were collected every 4 h for a full day. Total RNAs were extracted by centrifugation using a silica column. RNase protection analysis was performed using 5 µg of total RNA from each time-point. A cDNA of ribosomal protein 49 (abbreviated here as: RP49 for the protein, and *rp49* for the gene) was utilized as an internal standard for mRNA transcription. [^{32}P]UTP-labeled *rp49* complementary RNA antisense probe was prepared by the *in vitro* transcription driven from the T3 RNA polymerase promoter. The *period* cDNA fragment obtained by PCR amplification using the primer bmpersense2 and antisense primer

(5'-AAGCATTACCGAATCACGC-3') was digested by restriction enzyme *Eco*RI and then the smaller fragment was subcloned into pBluescript II SK$^+$ cloning vector (Stratagene). The plasmid was linearized by restriction enzyme *Xho*I and [^{32}P]UTP-labeled *period* complementary RNA antisense probe was prepared by the *in vitro* transcription driven from T3 RNA polymerase promoter. The *period*

```
                                                             attcgtgtttagagcgtttaaacctcaag  29
acataagttttgtgtagaggcgatgctgataaaatatttttatgctattaatgttcatatagaaagtaatttaggattttaccgttcaac  119
aatgtgggtgtctctaagtaacttttgcaaataattaactacaataacagtgccatcaataaataaacatagtttagaaacgaacaaata  209
aatttacacttttgaaacatacttctctacaactgtatggacttttcacactattaagtttggattaacgttcattttaaagtcttcata  299
atggataaccttgaaaactctgagaacactgacaaaatatcagattcggcatactcaaacagctgcagttacagtcaatcaagaagaagt  389
M  D  N  L  E  N  S  E  N  T  D  K  I  S  D  S  A  Y  S  N  S  C  S  Y  S  Q  S  R  R  S   (30)
cacagctcaaaatctactcattccgggagttattctagcggcagtagcggatacggaggtcgacatgttactggcactaacagtaataag  479
H  S  S  K  S  T  H  S  G  S  Y  S  S  G  S  S  G  Y  G  G  R  H  V  T  G  T  N  S  N  K   (60)
ttggctcagtcaccaaataaagataaagagttaaagaaaaagaaacaatcacaaacagagctaacgactaccaaagaacaaaatgaagga  569
L  A  Q  S  P  N  K  D  K  E  L  K  K  K  K  Q  S  Q  T  E  L  T  T  T  K  E  Q  N  E  G   (90)
gaaagcggacctcaagaatctgtgaacctggaaactctagaggaggagaagcatgatgttacgagcgcaccctctccagcgcatgtacag  659
E  S  G  P  Q  E  S  V  N  L  E  T  L  E  E  E  K  H  D  V  T  S  A  P  S  P  A  H  V  Q   (120)
gtcgagaagggacctgagaatatggatatagcttcaacagacactgaaggcgacaaggtagaggttttatcgtgtgacctgcaaccgcag  749
V  E  K  G  P  E  N  M  D  I  A  S  T  D  T  E  G  D  K  V  E  V  L  S  C  D  L  Q  P  Q   (150)
aactctactgtcgatgtagttacgagtcgttcgaatctaatctgttacactgatggattttcatgcgttatatcgatgcacgatggaatc  839
N  S  T  V  D  V  V  T  S  R  S  N  L  I  C  Y  T  D  G  F  S  C  V  I  S  M  H  D  G  I   (180)
                bmpersense1 primer                                                bmpersense2
gtaatgtacacgacatcatcgctgacggctacactcggtttcccgaaggacatgtggatcgggaggtcatttatcgatttcgttcatcca  929
V  M  Y  T  T  S  S  L  T  A  T  L  G  F  P  K  D  M  W  I  G  R  S  F  I  D  F  V  H  P   (210)

primer    bmpersense3 primer                                  bmperGTQEK primer    bmperantiGTQEK
agggaccgtagcacgttcgcgtcacagataaccagtggtttggccgtcccgaaaactgcaaacgGTACGCAAGAAAAAGctcaatcgccg  1019
R  D  R  S  T  F  A  S  Q  I  T  S  G  L  A  V  P  K  T  A  N  g  t  q  e  k  A  Q  S  P   (240)
          bmperantisense3                bmperantisense2
ggtaattctggatcgactatggtgtgtcggataagacgttacagaggtctgagcacggggttcggagttaaggaaagagtggtcacgttc  1109
G  N  S  G  S  T  M  V  C  R  I  R  R  Y  R  G  L  S  T  G  F  G  V  K  E  R  V  V  T  F   (270)
                                                    bmperantisense1
atgccgttccttttgaaattcacgttcaagaacataagcgatgaagagggaaatgtcatctacctcgtcatacaggccacaccattttt  1199
M  P  F  L  L  K  F  T  F  K  N  I  S  D  E  E  G  N  V  I  Y  L  V  I  Q  A  T  P  F  F   (300)
tcagcgttcaaaacgtcgtttgaaatattaccgaaagttaatcctttcgtgatgcgtcactccgcgaatggtaatctcgagtacttagac  1289
S  A  F  K  T  S  F  E  I  L  P  K  V  N  P  F  V  M  R  H  S  A  N  G  N  L  E  Y  L  D   (330)
ccagaatctgtaccttatctcggatatttacctcaagatgtccaagacaaggacgcattgcaactctatcatccggaagatttggaatac  1379
P  E  S  V  P  Y  L  G  Y  L  P  Q  D  V  Q  D  K  D  A  L  Q  L  Y  H  P  E  D  L  E  Y   (360)
ctccagcaagtatacgaggtaatcgttaaggatggtggcatgccgcgttccaaaacttacagaatgatgacacaaaacggtgattacata  1469
L  Q  Q  V  Y  E  V  I  V  K  D  G  G  M  P  R  S  K  T  Y  R  M  M  T  Q  N  G  D  Y  I   (390)
aaaatcgaaaccgaatggtcttcgtttattaatccttggtcgaagaagctggaattcgtcatcgggaaacattacatcattgaaggcccc  1559
K  I  E  T  E  W  S  S  F  I  N  P  W  S  K  K  L  E  F  V  I  G  K  H  Y  I  I  E  G  P   (420)
gagaatccagatgtattccagtcccaagatccggagaaacatacaaaattatgcgacgaacaaataaaaaagtccatggtatttcgcgag  1649
E  N  P  D  V  F  Q  S  Q  D  P  E  K  H  T  K  L  C  D  E  Q  I  K  K  S  M  V  F  R  E   (450)
aatatcgttaaacttatgaatgaggctctaacgaaaccagcggaagttgccaaacagcaaatgagcaagcgttgccaggacttagcgtcg  1739
N  I  V  K  L  M  N  E  A  L  T  K  P  A  E  V  A  K  Q  Q  M  S  K  R  C  Q  D  L  A  S   (480)
ttcatggagagcctcatggaggaaccaccgaaaaatgacgaggaattacgcctagaaatacaggatccggaccacagttattatgagcgt  1829
F  M  E  S  L  M  E  E  P  P  K  N  D  E  E  L  R  L  E  I  Q  D  P  D  H  S  Y  Y  E  R   (510)
gattcggtaatgctcggcgggatatctccccatcacgactacaatgacagcaaatcaagcacagaaacaccgctgagttataatcaacta  1919
D  S  V  M  L  G  G  I  S  P  H  H  D  Y  N  D  S  K  S  S  T  E  T  P  L  S  Y  N  Q  L   (540)
aactataacgagacattacaaagatatttcgacagtcatgaacccttttcgagcgagaacttcaatcccttgattgtggcaaaggagcca  2009
N  Y  N  E  T  L  Q  R  Y  F  D  S  H  E  P  F  S  S  E  N  F  N  P  L  I  V  A  K  E  P   (570)
aataatttaaaatcgatactaagtacttgtttatcaccaatggctcagaactctggcgattcgaccgagttgaatagttcgtccgattct  2099
N  N  L  K  S  I  L  S  T  C  L  S  P  M  A  Q  N  S  G  D  S  T  E  L  N  S  S  S  D  S   (600)
agtagtatggccctgggaggagtaagtcctgtgggtatattaggcgactaccaccacgttcggctcaccgagttcctactaaccaaacac  2189
S  S  M  A  L  G  G  V  S  P  V  G  I  L  G  D  Y  H  H  V  R  L  T  E  F  L  L  T  K  H   (630)
aatgatgaaatggaaaaggaactgataaatatgcatagagagagccggtccaatagtaaaggagaaagagataagacatctaacgaaacg  2279
N  D  E  M  E  K  E  L  I  N  M  H  R  E  S  R  S  N  S  K  G  E  R  D  K  T  S  N  E  T   (660)
agacagaagaaaaaggaacatctagcgaggtgtaaagcttcgtttcatccaaccgcgaccagcaccacacctgtggacaaagaagtgtat  2369
R  Q  K  K  K  E  H  L  A  R  C  K  A  S  F  H  P  T  A  T  S  T  T  P  V  D  K  E  V  Y   (690)
aaaaaacctcacggtgtcaaacgtgcttcgaaacatatcgaaacggaaactgtatcgcacaaataccattgtccatcgcccagagcgtct  2459
K  K  P  H  G  V  K  R  A  S  K  H  I  E  T  E  T  V  S  H  K  Y  H  C  P  S  P  R  A  S   (720)
agacctaggcaaaccaccagtgcggcgcctgtacaaacttcgtccactttgactacttctgttgcctcgactacctggccacccagcacc  2549
R  P  R  Q  T  T  S  A  A  P  V  Q  T  S  S  T  L  T  T  S  V  A  S  T  T  W  P  P  S  T   (750)
aatgctgctggaaatatgaacacattcatactaggagtcggaatggctccacagatgtcccttataagtccggttccgccaatggcagga  2639
N  A  A  G  N  M  N  T  F  I  L  G  V  G  M  A  P  Q  M  S  L  I  S  P  V  P  P  M  A  G   (780)
atgcttccgttgtattacactccgatggccacaatggctcccgtgccttccacatcggaggccgcgaaccatcaaaatcttcataataat  2729
M  L  P  L  Y  Y  T  P  M  A  T  M  A  P  V  P  S  T  S  E  A  A  N  H  Q  N  L  H  N  N   (810)
ccacagcagtatgctccacctccaatgcaatgtgtattgtacggtcagcctatatacggtcagccgatgtacagttcgccgttcgtgtac  2819
P  Q  Q  Y  A  P  P  P  M  Q  C  V  L  Y  G  Q  P  I  Y  G  Q  P  M  Y  S  S  P  F  V  Y   (840)
tcaccaatgaacccgcacactaactatcctatgcaacagaccacgccgcaaccgaatgctcagttcacacccacaaacacaatgaatccc  2909
S  P  M  N  P  H  T  N  Y  P  M  Q  Q  T  T  P  Q  P  N  A  Q  F  T  P  T  N  T  M  N  P   (870)
ttgtgtttggccaatagcaactatgaagagagaacagaagcagggcatgccacaacgccatctaaagacattaagaaaacgagcaggttg  2999
L  C  L  A  N  S  N  Y  E  E  R  T  E  A  G  H  A  T  T  P  S  K  D  I  K  K  T  S  R  L   (900)
ggtaacaccgaggaggttgtagacaaaactgacggagaatctagctattcctcgttttactcgtcgtttttcaagacggaatctggcagc  3089
G  N  T  E  E  V  V  D  K  T  D  G  E  S  S  Y  S  S  F  Y  S  S  F  F  K  T  E  S  G  S   (930)
```

Figure 2-1

```
gcggaagagagtggcgaaagcaagaaaagggacacaagaaatcacgtcaaattttgggataaaaactatcgtggcgacgatgggatggcg  3179
A  E  E  S  G  E  S  K  K  R  D  T  R  N  H  V  K  F  W  D  K  N  Y  R  G  D  D  G  M  A   (960)
atgcctggttcaacggaggtccgccaaactttttccgatcaccatattccagcgaaggtacccagacgaaacattgaaccaccatggatg  3269
M  P  G  S  T  E  V  R  Q  T  F  S  D  H  H  I  P  A  K  V  P  R  R  N  I  E  P  P  W  M   (990)
gagcaggtatgcctgacgtcagagctcgtctacaattaccagataaggacaaaaacaatggaggaagtactttcatctgacaagcagaag  3359
E  Q  V  C  L  T  S  E  L  V  Y  N  Y  Q  I  R  T  K  T  M  E  E  V  L  S  S  D  K  Q  K  (1020)
attagtaaccttgagcagccgtcactagtgaacgagcaacttggtcagctgtacctggaccttcagttggaaggtgtcgccgcgcgactc  3449
I  S  N  L  E  Q  P  S  L  V  N  E  Q  L  G  Q  L  Y  L  D  L  Q  L  E  G  V  A  A  R  L  (1050)
accctcgaggagggtatcaccagttccagcagttccggtgaagagacaacaacctcgtccaagaggaaacgggaatacagcaaattggtt  3539
T  L  E  E  G  I  T  S  S  S  S  S  G  E  E  T  T  T  S  S  K  R  K  R  E  Y  S  K  L  V  (1080)
 bmpersense4 primer
atgatatatgaagagaacgctcccttcccacctccccttctgaagagatacaagagataccgacagccaatcagtccgagtcctagtcg  3629
M  I  Y  E  E  N  A  P  F  P  P  P  P  S  E  E  I  Q  E  I  P  T  A  N  Q  S  E  S  *     (1108)
aacagtatttctaccacttcgtgttgtcgaggtacaaaagaattacgacattactttttgcaattattatatattgcaatgattatact  3719
gggtgttaggaaatcgcttccgaaaacgaagacagacgatagtacaaatgacaccttagatgataggtcggaaaaaaatagatcatgatt  3809
taaaaaaaaaaaaaaaaa                                                                        3827
```

Fig. 2. Nucleotide and deduced amino acid sequence of *Bombyx mori period* cDNA. Residue numbers in nucleotide and amino acid sequences are shown at the right side of each line. The end of translation (stop codon) is marked by an asterisk (*). The alternatively spliced nucleotides and deduced amino acid residues in PAS domain are boxed and denoted with capital and small letters conversely. Locations of DNA primers used for the RACE methods and for identification of isoforms are indicated by the arrows.

riboprobe protected nucleotides 1526–1844 of the *period* cDNA. The sample of total RNAs was resuspended in a hybridization buffer (40 mM piperazine-1,4'-bis(2-ethanesulfonic acid) (PIPES) (pH 6.4), 400 mM NaCl, 1 mM EDTA in the final solution) containing a set of probes (50,000 cpm each of *rp49* and *period*). Hybridization was carried out at 45°C overnight after heat denaturation at 85°C for 3 min. The reaction mixture was treated with RNases A and T1 at 13°C for 30 min, and then the digest was treated with phenol/chloroform followed by centrifugation. The supernatant was treated with EtOH to precipitate RNA. Protected double-stranded RNAs were separated on a gel of 6% acrylamide with 8 M urea and visualized by autoradiography. Quantification of the protected fragments was performed with a Lane and Spot Analyzer (ATTO, Tokyo) by calibrating the amount of *period* relative to that of *rp49*. Two independent experiments were performed with similar results.

A semi-quantitative RT-PCR was carried out by amplifying the *actin* gene as an internal standard. Primers used were a gene-specific sense primer bmactin1 (5'-ATGTGCAAGGCCGGTTTCGC-3') and an antisense primer bmactin2 (5'-CGACGCAGCTCATT-GTAG-3'), which correspond to nucleotides 49–68 and 332–341 respectively (Nicole *et al.*, 1987). The *actin* gene was amplified under the following conditions: 3 min at 94°C, followed by 28 cycles of denaturation at 94°C for 30 sec, annealing at 58°C for 30 sec, and extension at 72°C for 30 sec. PCR for the *period* gene was performed with bmpersense3 and bmperantisense3 as follows: 3 min at 94°C followed by 29 cycles of denaturation at 94°C for 30 sec, annealing at 56°C for 30 sec, and extension at 72°C for 30 sec. These PCR products were analyzed by electrophoresis on a 2% agarose gel.

Sequence analyses

Sequence analysis was carried out on a Thermo Sequenase Cy5.5 dye terminator cycle sequencing kit (Amersham Biosciences) to determine the nucleotide sequence of plasmid DNA containing *period*. All the subcloned PCR products described above were analyzed from both 5' and 3' directions. For sequence determination, at least three individual clones were analyzed.

RESULTS

Isolation of a *Bombyx mori period* cDNA fragment

The first PCR amplification with a pair of the primers PAS-AF and PAS-CR for *B. mori* cDNA afforded an oligonucleotide fragment with 969 bp (Fig. 2). The deduced sequence with 323 amino acids was judged to be the region corresponding to PAS domain from its similarity to those of other insect PERIOD PAS domains. When this amino acid sequence was compared with the sequence reported by Regier *et al.* (1998), however, we found that the sequence of *Bombyx mori* PERIOD reported in the database contained an additional five amino acid residues in this PAS domain. The insertion/deletion of GTQEK pentapeptide was found between the Asn(=N) and Ala(=A) residues in a proline-truncated site PKTANAQSP.

As shown in Fig. 1, there is a group of PERIOD proteins with GTQEK or related peptides in the PAS domain, while there is another group of PERIOD proteins lacking GTQEK at the same site. This strongly suggested the presence of an alternative splicing mechanism for the insertion/deletion of the pentapeptide. When the database oligonucleotide sequence was examined carefully, it was indeed found that the *B. mori period* mRNA consisted of consensus splicing sequences at this particular site (see Discussion). To confirm the existence of PERIOD with a proline-truncated site of PKTAN*GTQEK*AQSP, we further carried out PCR experiments to identify a clone containing the oligonucleotide fragment corresponding to the pentapeptide inserted.

Identification of novel *period* isoforms that differed in their PAS domain open reading frame

In order to amplify a clone consisting of oligonucleotide GTACGCAAGAAAAAG that encodes GTQEK, we carried out PCR by using a set of primers of bmperGTQEK and bmperantisense1 (Fig. 2). The primer bmperGTQEK that contains an oligonucleotide fragment GTACGCAAGAA-AAAG (=GTQEK) is expected to hybridize to a clone of cDNA isoform that is being explored. As a result, a single cDNA of approximately 193 bp was amplified (Fig. 3A), and the sequence analysis of this PCR product revealed indeed the presence of 15 bp oligonucleotide corresponding to the pentapeptide GTQEK. The result clearly indicated that a clone having GTQEK was definitely present. We designated the clone having the oligonucleotide corresponding to GTQEK as *period* α. Thus, PERIOD α appears to be a dis-

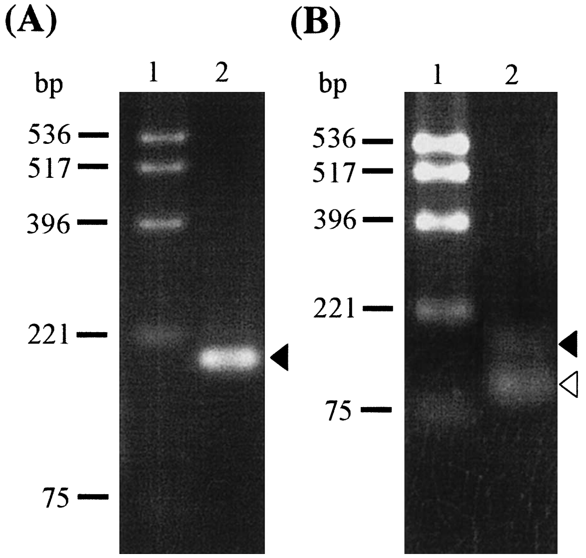


Fig. 3. Electrophoresis profiles of PCR amplification products for detection of the alternative splicing site. (A) Lane 1: marker; and Lane 2: Amplification product of *period α* specific band corresponding to 193 bp with bmperGTQEK primer and bmperantisense1 primer. (B) Lane 1: marker; Lane 2: Amplification of the gel bands corresponding to 102 bp *period α* and 117 bp *period β* oligonucleotide fragments, respectively, with bmpersense3 and bmperantisense3 at ZT 17. Black arrowheads indicate *period α*, while white arrowhead shows *period β* .

tinct protein isoform in the silk moth *B. mori*.

When PCR was carried out with bmpersense3 and bmperantisense3 primers (Fig. 2), two distinct bands of cDNA clones were detected in the gel electrophoresis, *i.e.* one with 102 bp, and the other with 117 bp, 15 bp larger (Fig. 3B). The results of their sequence analysis revealed that the shorter cDNA is a fragment of cDNA with a 15 bp-deletion in the PAS domain, while the longer cDNA is a fragment of *period α*. We designate hereafter the cDNA clone with a 15 bp-deletion as *period β*. It should be noted that both *period α* and *period β* are simultaneously identified from the same source.

5' RACE to determine the 5' end sequence of *Bombyx mori* PERIOD cDNA isoforms

To determine the 5' end structure of the *B. mori period* mRNA isoforms, the 5' RACE method was carried out for both respective isoforms. First-strand cDNAs having a 5' structure with a 15 bp splice site was synthesized by the reverse-transcription reaction using mRNAs and a gene-specific antisense primer bmperantisense1. The first PCR was performed with an AAP primer and a gene-specific antisense primer bmperantisense2, and using this PCR product as a template, subsequent nested PCR was carried out with a AUAP primer and a gene-specific bmperantisense3 primer. Because all these three antisense primers were designed downstream of the 15 bp splice site, the resulting PCR products should contain both isoforms of *period α* and *period β*. A band with the largest molecular size in the electrophoresis products from the second PCR was isolated and the cDNA obtained was subcloned into pDrive cloning vector. From the sequence analysis of several clones, as expected, two different oligonucleotide fragments were determined. These contained 5' UTR, and were 5' terminal fragments originating from the *period α* and *period β* mRNA isoforms.

In order to confirm the amplification of a unique *period α* fragment, another nested PCR was carried out with AUAP and a *period α*-specific primer bmperantiGTQEK. The primer bmperantiGTQEK has an antisense sequence against an oligonucleotide corresponding to the entire GTQEK, and was expected to hybridize to the clone of the *period α* cDNA. Eventually, the sequence of *period α* 5' region was ascertained, by repeatedly analyzing a number of clones. The overlapped oligonucleotide sequences of *period α* and *period β* isoforms were found to be exactly identical. In addition to the 15 bp corresponding to GTQEK, a sequence difference between *period α* and *period β* was seen for the 5' end in the 5' UTR. The isomer *period α* possesses about a 65 bp longer UTR at the 5' terminal (Fig. 2).

3' RACE to determine the 3' end sequence of *Bombyx mori* PERIOD cDNA isoforms

To obtain individual *period α* and *period β* cDNA clones including 3' UTR, 3' RACE by the touchdown PCR procedure was first performed using a gene-specific sense primer bmpersense1 and adapter primer. Subsequent nested PCR was carried out with another sense primer bmpersense2 and adapter primer. These sense primers bmpersense1 and bmpersense2 were designed for the region upstream of the alternative splice site, and thus amplified PCR products should comprise both oligonucleotide fragments from *period α* and *period β* (Fig. 2). PCR products isolated from a gel band of approximately 2,800 bp in electrophoresis were subcloned into the cloning vector for sequence analysis. When several clones were analyzed randomly, exactly the same nucleotide sequence was clarified from a set of clones and it was judged to be a complete 3' terminal fragment of *period β* . This fragment lacked the 15 bp that corresponds to GTQEK. On the other hand, another set of clones was also sequenced and identified to be a 3' terminal fragment of *period α*. These included a 15 bp oligonucleotide sequence. Collectively, the sequences analyzed by the 3' RACE method for fragments of *period α period β* were found to be exactly identical except for the alternative splicing site of 15 bp.

The entire sequences of these *period α* and *period β* isoforms were thus determined by the 3' and 5' RACE methods. In summary, isoform *period α* consists of a 3,324 bp oligonucleotide encoding 1,108 amino acid residues, whereas isoform *period β* comprises 3,309 bp corresponding to 1,103 amino acid residues. Isoforms *period α* and *period β* were found to be indistinguishable with the exception of the 15 bp deletion/insertion site. Such a pair of isoforms with a deletion/insertion sequence, namely two splice variants, has previously been reported only for the PERIOD proteins of the two honeybees, *Apis mellifera* and *Apis cerana japonica* (Toma *et al.*, 2000; Shimizu *et al.*, 2001).

In these 3' RACE reactions, a band of smaller molecular size was often observed in the gel electrophoresis. If we were to carry out nested PCR using sense primer bmpersense4 designed downstream from the alternative splice site, amplified PCR products should comprise only a single oligonucleotide fragment from both *period* α and *period* β. Our careful search for the cDNA sequence found a fragment (5'-GTCCTAGTCGACGCGTGGCC-3') near the stop codon, to which the antisense nucleotide sequence of adapter primer was very much similar. In particular, 5' terminal sequences of 11 nucleotide residues were almost completely identical. When an adapter primer was utilized as an antisense primer, one major amplified product was obtained with a small amount of sub-product having about 200 bp less than the major product. When we utilized a dT_{17}-adapter primer, no such sub-product was detected. These results clearly show that mis-hybridization occurred with the solo adapter primer, while the addition of dT_{17} prevented this unnecessary hybridization.

RNase protection assay for exploration of circadian oscillation of *period* mRNA

The fruit fly *Drosophila melanogaster* and the silk moth *Antheraea pernyi* both exhibit a prominent circadian oscillation of *period* mRNA, which functions for example for circadian rhythmicity in eclosion behavior and locomotor activity. In order to explore such a temporal pattern of mRNA abundance of the *period* gene in *B. mori* brains and optic lobes, the RNase protection assay was performed. Since the assay to differentiate gene isoforms appeared to complicate the outcome, we set a probe at a site away from the deletion/insertion site. The radio-labeled *B. mori rp49* and *period* cRNA antisense probes were hybridized with the respective mRNA obtained at 4-h intervals throughout the 24-h cycle and then the protected fragments were electrophoresed. The autoradiograph intensities of the gel bands were assessed by calibrating the amount of *period* relative to that of *rp49* (Fig. 4). RP49 served as a control for the amount of RNA loaded in each lane. It was found that the *B. mori period* mRNA also fluctuated in a circadian manner with the highest mRNA level in the early night (around ZT 17) and the lowest level in the day (around ZT 13), showing a cycle similar to those seen in the *period* mRNA expression of other insects. The magnitude of the daily mRNA oscillation was about 3.5–fold.

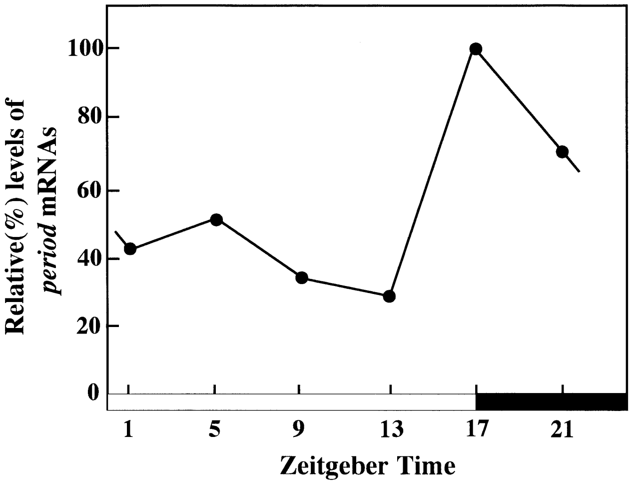


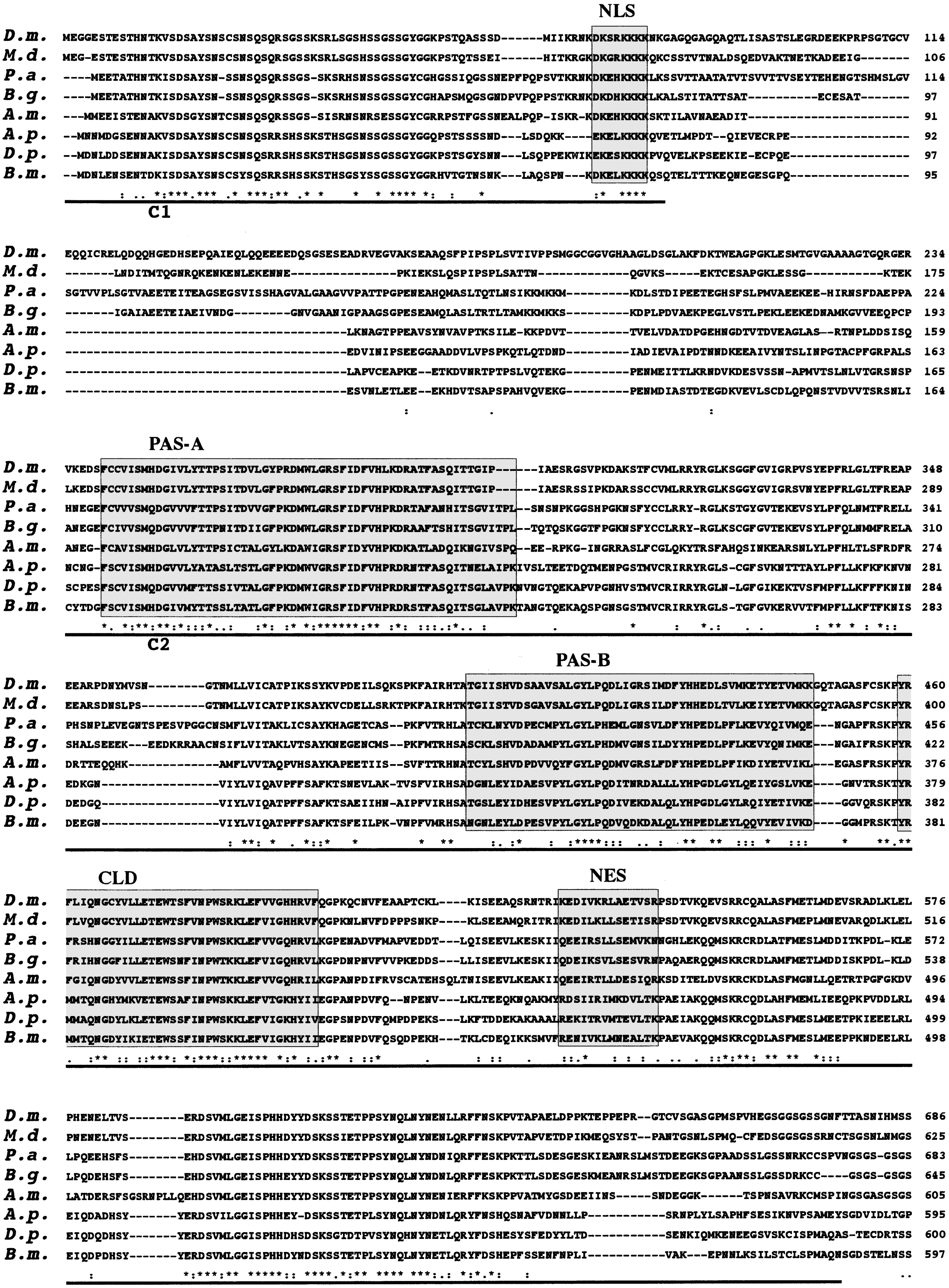
Fig. 4. Cycling of *period* mRNA levels in *Bombyx mori*. RNA protection assay for *period* mRNA expression was carried out twice independently. The white-black bar represents the timing of the 17L:7D light-dark cycle experienced prior to eclosion. Relative level (%) of mRNA refers to the ratio of *period* : *rp49* mRNAs that were converted to percentage of maximal level.

DISCUSSION

Sequence similarity and identity of PERIOD proteins

Whether or not the structurally elucidated *Bombyx mori* proteins are PERIOD was examined by comparing their whole structures, aligning the sequences with other reported PERIOD proteins. When the full-length sequences were compared, *B. mori* PERIOD showed considerably high homologies with other insect PERIOD proteins. Calculated sequence similarity and identity are 48% and 29% for *Drosophila melanogaster* PERIOD, 56% and 42% for *Antheraea pernyi* PERIOD, and 73% and 58% for *Danaus plexippus* PERIOD. PERIOD proteins are distinguished by having a series of characteristic conserved (C) domains numbered C1–C7, and C1–C3 are particularly highly homologous among insect PERIOD proteins (Colot *et al.*, 1988; Lin *et al.*, 2002). As shown in Fig. 5, these C1–C3 characteristic regions are highly conserved in the *B. mori* PERIOD proteins clarified in the present study. It should be noted that *B. mori* PERIOD proteins are significantly highly homologous with *Danaus plexippus* PERIOD in their entire structures.

PERIOD proteins also have several other signals and domains that are functionally important in exhibiting potential cycling activities. When the primary structure of *B. mori* PERIOD proteins was depicted by a rod model as shown in Fig. 6, a domain construction was found to be almost the same as other insect PERIOD proteins. These include the nuclear localization signal (NLS) (Vosshall *et al.*, 1994; Chang *et al.*, 2003; Chang and Reppert, 2003), the PAS domain (Crews *et al.*, 1988), the cytoplasmic localization domain (CLD) (Saez and Young, 1996), and the nuclear export signal (NES) (Fukuda *et al.*, 1997; Vielhaber *et al.*, 2001). It was found that all of these sequences are highly conserved in *B. mori* PERIOD proteins with 70–100% sequence similarities. In particular, the most important characteristic domain structure PAS showed considerably high homologies: *i.e.*, calculated sequence similarity and identity for PAS-A and PAS-B domains, (88% similarity and 69% identity for PAS-A) and (67% similarity and 38% identity for


NLS
C1
PAS-A
C2
PAS-B
CLD
NES
D.m.
M.d.
P.a.
B.g.
A.m.
A.p.
D.p.
B.m.
114
106
114
97
91
92
97
95
234
175
224
193
159
163
165
164
348
289
341
310
274
281
284
283
460
400
456
422
376
379
382
381
576
516
572
538
496
494
499
498
686
625
683
645
605
595
600
597

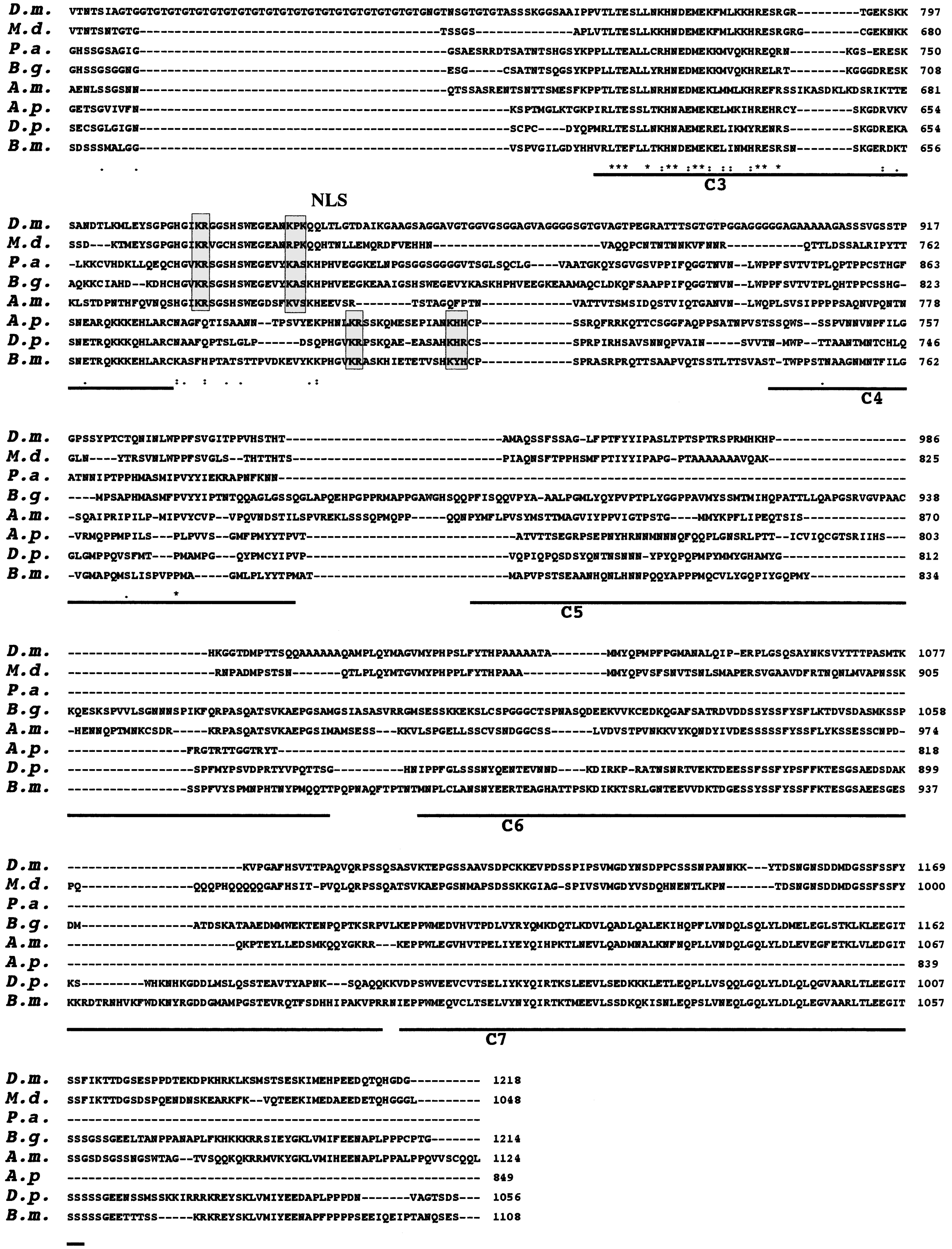


Fig. 5. PERIOD protein sequences alignment comparison among eight insect species. PERIOD protein sequences are aligned by inserting gap (–) to achieve maximum homology. Functional motif structures such as NLS, PAS-A, PAS-B, NES, and CLD are boxed. The defined C1–C7 regions are underlined with labels below. Seven different insect PERIOD protein homologues are aligned with *Bombyx mori* (*B.m.*) PERIOD protein, and those proteins are from *Drosophila melanogaster* (*D.m.*), *Musca domestica* (*M.d.*), *Periplaneta americana* (*P.a.*), *Blattella germanica* (*B.g.*), *Apis mellifera* (*A.m.*), *Antheraea pernyi* (*A.p.*), and *Danaus plexippus* (*D.p.*). '*****'Residues in that column are identical in all sequences in the alignment. '**:**'Conserved substitutions have been observed. '**.**'Semi-conserved substitutions have been observed.

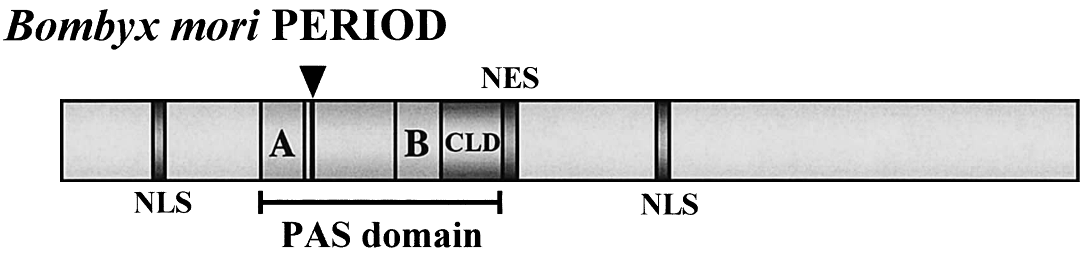


Fig. 6. Schematic look of the primary structure deduced from the nucleotide sequences of *Bombyx mori period* cDNA. Functional motifs and domains are NLS (residue numbers of amino acid sequences: PERIOD α, 68–75, 697, 698, and 711–713 and PERIOD β, 68–75, 692, 693, and 706–708), PAS-A (α, 170–228 and β, 170–228), PAS-B (α, 323–371 and β, 318–366), cytoplasmic localization domain (CLD) (α, 380–417 and β, 375–412), and NES (α, 449–462 and β, 444–457). The arrow head shows the splicing site with 5 amino acid difference for PERIOD α and PERIOD β isoforms.

PAS-B) for *Drosophila melanogaster* PERIOD, (95% and 83% for PAS-A) and (88% and 71% for PAS-B) for *Antheraea pernyi* PERIOD, and (96% and 86% for PAS-A) and (85% and 75% for PAS-B) for *Danaus plexippus* PERIOD. The results clearly indicate that the proteins elucidated in this study are genuine PERIOD proteins.

In the present study, we were able to clone two different *B. mori period* cDNAs. The structures of their open reading frames were completely elucidated and these isoforms showed totally identical sequences except for a deletion/insertion difference of 15 bp. Thus, our initial purpose to find a clone having a fragment corresponding to the pentapeptide GTQEK was successfully achieved by cloning an isoform containing this 15 bp. The resulting PERIOD α and β proteins are the first *period* transcript pairs with different PAS domain sequences present in a single insect species. A similar set of PERIOD isoforms has been reported only for the honeybees, *Apis mellifera* and *Apis cerana*, but in that case for a deletion/insertion of eight amino acid residues near the so-called *per*^{S} mutation site (Shimizu *et al.*, 2001).

Most recently, genome structural analyses have been reported for the silkworm *Bombyx mori* (Mita *et al.*, 2004) and the genome sequence has been opened as an Internet web site (http://sgp.dna.affrc.go.jp). Even though sequence data are said to include only up to about 80% of protein encoding genes, we attempted a BLAST search of genomic DNA sequences using our cDNA sequence in order to reveal the genomic *period*. We hit a series of fragments, but not a single gene, that cover the entire *period* cDNA sequence elucidated in this study. We found three cDNA regions that are absent from the genome DNA sequences, *i.e.* two regions in the open reading frame with about 42 bp and 90 bp corresponding to 14 and 30 amino acid residues, respectively, and one region in 5' UTR with about 48 bp. It is thus evident for *period* that although the genome sequence is incomplete at the moment, the cDNA structure we have clarified is almost completely compatible with the genome sequence and that of 3,800 bp-overlapped regions, only 9 mismatched nucleotides were found.

Alternative splicing mechanism to afford α and β PERIOD isoforms

In the present study, we identified novel *period* isoforms that possess a deletion/insertion of 15 bp in the PAS domain. The cDNA structure of *period a* clearly exhibited the existence of splice consensus sites (Norton *et al.*, 1994; Maniatis and Tasic, 2002), *i.e.* GU (GT in cDNA) at the 5' side and AG at the 3' side of the specific 15 bp sequence **GT**ACGCAAGAAAA**AG**. This strongly suggests that the isoforms with the 15 bp sequence difference were produced by either ordinary processing (α) or by alternative splicing (β). Alternative splicing would be brought about from an intron longer than 15 bp or from DNA itself. Since the genome sequence indicates that there is an intron adjacent to this splicing site, the alternative splicing mechanism illustrated in Fig. 7 is speculated as shown. It is clear that GT and AG splicing consensus sites truncate the intron. At this moment, the structural trigger to cause this mechanism is not known.

As shown in Fig. 1, PERIOD proteins having GTQEK-like sequences have been found in a series of species of Lepidoptera. In particular, PERIOD proteins of *Apatelodes*, *Danaus*, and *Plodia*, consist of exactly same pentapeptide GTQEK. However, no isoform lacking GTQEK has so far been reported for any of the PERIOD proteins having a GTQEK(-like) sequence, and it is not clear at this moment whether their isogene is simply missing. When we carefully examined the oligonucleotide sequences corresponding to these GTQEK(-like) sequences, however, it was found that the splice consensus sites existed in all genes (Fig. 8). Some Lepidopteran *period* genes clearly show a distinct 15 bp GT–AG fragment, while some other genes exhibit more than 15 bp of GT–AG fragments. Of course, this so-called consensus frame sequence for splicing is not always essential for alternative splicing mechanisms. Moreover, PERIOD proteins devoid of GTQEK(-like) sequences have also been isolated from several Lepidoptera insects (Fig. 1). There is a possibility that these PERIOD proteins are the products of alternative splicing and thus there are isoforms containing GTQEK(-like) sequences. No reports have been announced for such splice variants, however.

When the secondary-structure of the PAS domain of *B. mori* PERIOD isoform α was analyzed by the Chou-Fasman

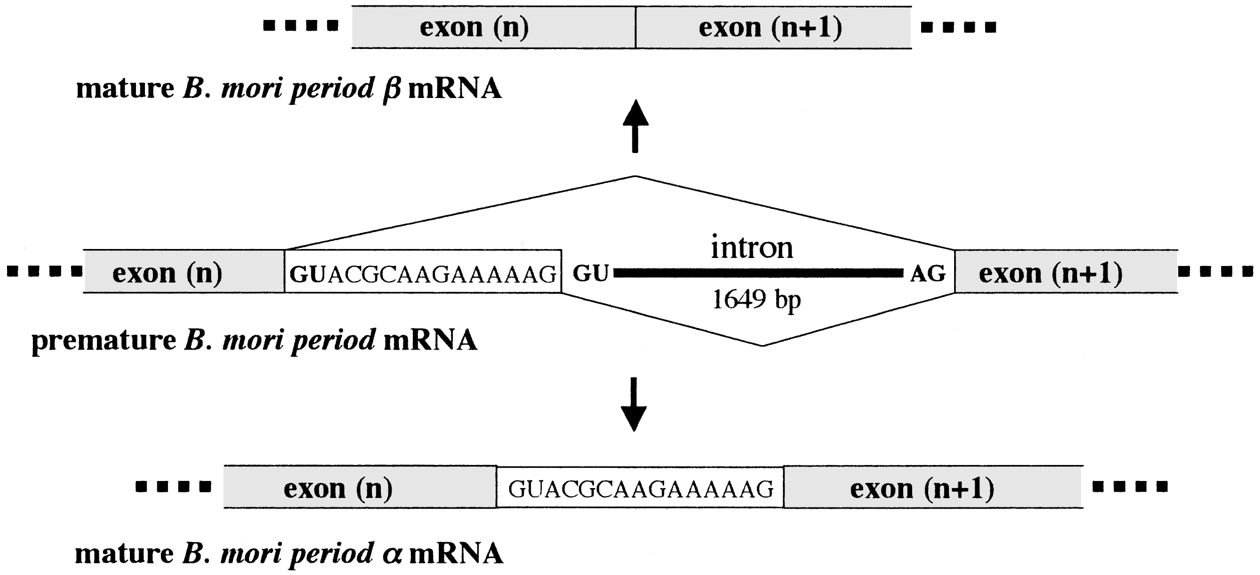


Fig. 7. The putative mechanism of alternative splicing to produce *period α* and *period β* mRNA, respectively, from premature *Bombyx mori period* mRNA. The letters in the parentheses, n and n + 1, denote the number of exon in the *period* gene.

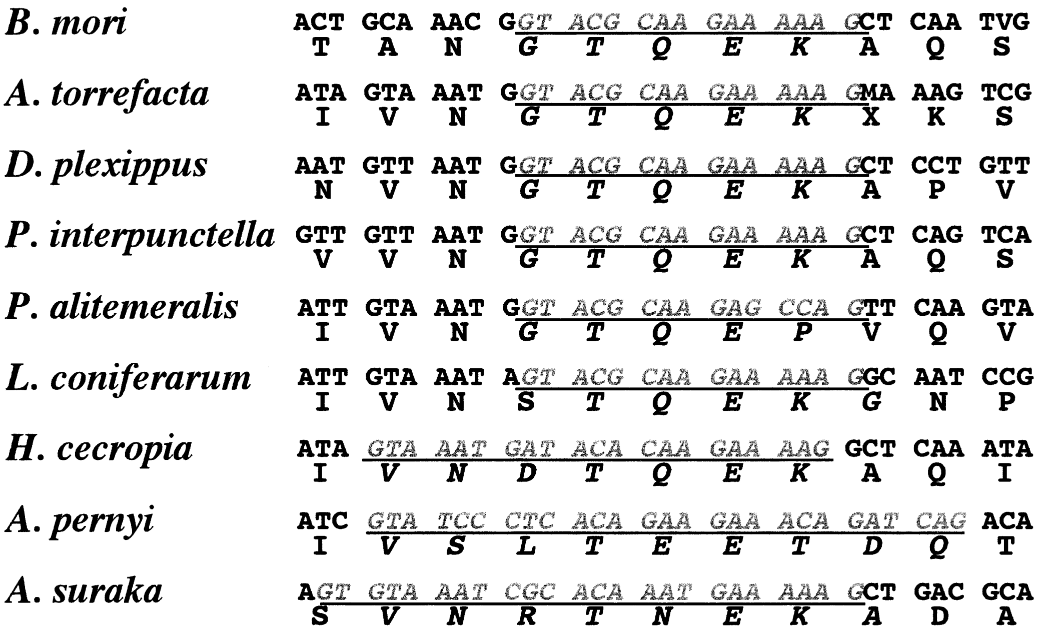


Fig. 8. Nucleotide sequence and its deduce amino acid sequence of various Lepidoptera PERIOD proteins around GTQEK-like site. The nucleotide sequences depicted in Fig. 1 are analyzed in the putative alternative splicing site (gray letters). It is probable that a PERIOD isoform lacking GTQEK or GTQEK-like sequence is present in each insect.

method (Nishikawa and Noguchi, 1991; Ito *et al.*, 1997), GTQEK in a proline-truncated tetradecapeptide PKTANGTQEKAQSP frame was predicted likely to be in a loop structure. The particular PAS site of *B. mori* PERIOD isoform β, PKTANAQSP lacking GTQEK is also in a loop. These loop structures may be exposed to the molecular surface when the protein is not bound in a complex. The presence or absence of GTQEK might therefore affect the preferred interaction with other proteins, for example TIMELESS. The side chains of T(=Thr), Q(=Gln), E(=Glu), and K(=Lys) are all hydrophilic and act simultaneously as both proton donor and acceptor. It is not clear at the moment whether or not GTQEK is a key structure to differentiate the complementary binding protein.

Questions to be answered for α and β PERIOD isoforms

Although we may have identified two distinct mRNA, or protein isomers of PERIOD α and β, there are still a number of questions to be answered. The first question is whether or not the presence of these isoforms is a simple consequence of individual differences between silk moths. Since we found a splice site for pentapeptide GTQEK in the exon of the silkworm genome DNA sequence, it is likely that there is a specific alternative splicing mechanism to produce either *period* α mRNA or *period* β mRNA. If such a mechanism is dependent upon the individual silk moth, there might be no significant functional difference between PERIOD α and β. In order to solve these questions, it will be necessary to quantify the ratio of isoform α and isoform β in PERIOD protein forms or *period* mRNA forms. So far, we have not succeeded in such quantifications in spite of attempts to differentiate the ratio between *period* α and *period* β mRNAs. As shown in Fig. 3B, the amount of *period* β mRNA was judged to be a few times larger than that of *period* α mRNA. To quantify this difference more firmly, we may need an elite system to amplify each mRNA isoform accurately and quantitatively.

The RNase protection assay performed for *period* mRNA exhibited a distinct temporal pattern of mRNA abundance of the *B. mori period* gene with an approximately 3.5–fold daily oscillation in the amount of mRNA. The question therefore arose as to whether the isomers exhibited the same temporal pattern of such oscillation. Using a semi-quantified PCR method, we attempted to quantify the *period* mRNA isoforms, but were unable to measure the amount of each isoform accurately, so that the ratio varied considerably, resulting in our inability to see a prominent oscillation. The *period* β mRNA is several times more abundant than *period* α mRNA throughout the entire day. Again, we need a selected system to amplify each mRNA isoform precisely and quantitatively.

Another important issue is whether these two *period* mRNA isoforms are expressed in the same neurons. We first attempted *in situ* hybridization for adult silk moth brains by using probes prepared from various sites of the cDNA clone. However, this ordinary hybridization is not yet successful for reasons that are not yet clear. On the other hand, isoforms of the PERIOD proteins should be detected by sophisticated monoclonal antibodies that can differentiate the insertion/deletion of pentapeptide GTQEK, so that immunocytochemical study could then clarify their localization in different cells. Thus, we are now preparing antibodies for further investigation of the PERIOD protein. Nonetheless, the present study has shown, for the first time, the presence of two *period* transcripts with different PAS domain sequences present in a single insect species. It will now be important to analyze or specify the functions of the two isoforms, whether these may be different or differently localized, especially in order to explore the essential molecular mechanism of circadian rhythms in this species.

ACKNOWLEDGMENTS

We thank Prof. Ian A. Meinertzhagen (Dalhousie University, Halifax, Canada) for his kind comments on the manuscript and correction of the English.

REFERENCES

Chang DC, McWatters HG, Williams JA, Gotter AL, Levine JD, Reppert SM (2003) Constructing a feedback loop with circadian clock molecules from the silkmoth, *Antheraea pernyi*. J Biol Chem 278: 38149–38158

Chang DC, Reppert SM (2003) A novel C-terminal domain of *Drosophila* PERIOD inhibits dCLOCK:CYCLE-mediated transcription. Curr Biol 13: 758–762

Colot HV, Hall JC, Rosbash M (1988) Interspecific comparison of the *period* gene of *Drosophila* reveals large blocks of non-conserved coding DNA. ENBO J 7: 3929–3937

Crews ST, Thomas JB, Goodman CS (1988) The Drosophila *single-minded* gene encodes a nuclear protein with sequence similarity to the *per* gene product. Cell 52: 143–151

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) "Touchdown" PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19: 4008

Dunlap JC (1999) Molecular bases for circadian clocks. Cell 96: 271–290

Frohman MA, Dush MK, Martin GR (1998) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8998–9002

Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. Nature 390: 308–311

Gekakis N, Saez L, Delahaye-Brown A-M, Myers MP, Sehgal A, Young MW, Weitz CJ (1995) Isolation of *timeless* by PER protein interaction: defective interaction between *timeless* protein and long-period mutant PER^L. Science 270: 811–815

Getzoff ED (2002) PASsing a signal: low carbs, less protein. Chem Biol 9: 1165–1166

Huang ZJ, Edery I, Rosbash M (1993) PAS is a dimerization domain common to *Drosophila* Period and several transcription factors. Nature 364: 259–262

Ito M, Matsuo Y, Nishikawa K (1997) Prediction of protein secondary structure using the 3D-1D compatibility algorithm. Comput Appl Biosci 13: 415–424

Itoh MT, Hattori A, Nomura T, Sumi Y, Suzuki T (1995) Melatonin and arylalkylamine N-acetyltransferase activity in the silkmoth, *Bombyx mori*. Mol Cell Endocrinol 115: 59–64

Konopka RJ, Benzer S (1971) Clock mutants of *Drosophila melanogaster*. Proc Natl Acad Sci USA 68: 2112–2116

Lee C, Bae K, Edery I (1999) PER and TIM inhibit the DNA binding activity of a *Drosophila* CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: A basis for circadian transcription. Mol Cell Biol 19: 5316–5325

Lin GG, Liou R, Lee H (2002) The period gene of the German cockroach and its novel linking power between vertebrate and invertebrate. Chronobiol Int 19: 1023–1040

Maniatis T, Tasic B (2002) Alternative pre-mRNA splicing and promote expansion in metazoans. Nature 418: 236–243

Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin-I T, Abe H, Shimada T, Morishita S, Sasaki T (2004) The genome sequence of silkworm, *Bombyx mori*. DNA Res 11: 27–35

Nicole M, Gaillard J, Prudhomme JC (1987) Nucleotide sequence of the coding region of two actin genes in *Bombyx mori*. Nucleic Acids Res 15: 2781

Nishikawa K, Noguchi T (1991) Predicting protein secondary structure based on amino acid sequence. Methods in Enzymology 202: 31–44

Norton PA (1994) Alternative pre-mRNA splicing: factors involved in splice site selection. J Cell Sci 107: 1–7

Regier JC, Fang QQ, Mitter C, Peigler RS, Friedlander, Solis MA (1998) Evolution and phylogenetic utility of the *period* gene in Lepidoptera. Mol Biol Evol 15: 1172–1182

Rusak B, Abe H, Mason R, Piggins HD, Ying SW (1993) Neurophysiological analysis of circadian rhythm entrainment. J Biol Rhythms 8 Suppl: S39–45

Saez L, Young MW (1996) Reguration of nuclear entry of the Drosophila clock proteins period and timeless. Neuron 17: 911–920

Sauman I, Tsai T, Roca AL, Reppert SM (1996) Period protein is necessary for circadian control of egg hatching behavior in the silkmoth Antheraea pernyi. Neuron 17: 901–909

Shimizu I, Kawai Y, Taniguchi M, Aoki S (2001) Circadian rhythm and cDNA cloning of the clock gene *period* in the honeybee *Apis cerana japonica*. Zool Sci 18: 779–789

Shirasu N, Shimohigashi Y, Tominaga Y, Shimohigashi M (2003) Molecular cogs of the insect circadian clock. Zool Sci 20: 947–955

Soriano V (1981) The circadian rhythm embraces the variability that occurs within 24 hours. Int J Neurol 15: 7–16

Toma DP, Bloch G, Moore D, Robinson GE (2000) Changes in period mRNA levels in the brain and division of labor in honey bee colonies. Proc Natl Acad Sci USA 97: 6914–6919

Vielhaber EL, Duricka D, Ullman KS, Virshup DM (2001) Nuclear export of mammalian PERIOD proteins. J Biol Chem 276: 45921–45927

Vosshall LB, Price JL, Shegal A, Saez L, Young MW (1994) Block in nuclear localization of *period* protein by a second clock mutation, *timeless.* Science 263: 1606–1609

(Received May 13, 2004 / Accepted July 13, 2004)