

cDNA Cloning of the Housefly Pigment-Dispersing Factor (PDF) Precursor Protein and Its Peptide Comparison among the Insect Circadian Neuropeptides

AYAMI MATSUSHIMA,^a SEIJI SATO,^a YOSHIRO CHUMAN,^a YUKIMASA TAKEDA,^a SATORU YOKOTANI,^a TAKERU NOSE,^a YOSHIYA TOMINAGA,^b MIKI SHIMOHIGASHI^b and YASUYUKI SHIMOHIGASHI^{a*}

^a Laboratory of Structure–Function Biochemistry, Department of Chemistry, Faculty and Graduate School of Sciences, Kyushu University, Fukuoka 812-8581, Japan

^b Division of Biology, Faculty of Science, Fukuoka University, Fukuoka 814-0180, Japan

Received 17 April 2003

Accepted 19 May 2003

Abstract: Pigment-dispersing factor (PDF), an 18-amino acid neuropeptide, is a principal circadian neurotransmitter for the circadian rhythms of the locomotor activity in flies. Recently, two completely different types of PDF precursor were clarified; that of the cricket *Gryllus bimaculatus* and that of the last-summer cicada *Meimuna opalifera*. The *G. bimaculatus* PDF precursor is extraordinarily short and comprises a nuclear localization signal (NLS), while the *M. opalifera* PDF precursor is of ordinary length, comparable to that seen for the precursors of crustacean β -PDH homologues. Although their PDF peptide regions were exactly the same, the regions containing a signal peptide combined with a PDF-associated peptide (PAP) were remarkably different from each other. Such a grouping suggested a fundamental role for the PAP peptide in the circadian clock, perhaps associated with PDF function. In the present study, the cDNA cloning of PDF from the adult brains of the housefly *Musca domestica* was carried out and it was found that an isolated clone (527 bp) encodes a PDF precursor protein of ordinary length. The PDF peptide shows a high sequence identity (78%–94%) and similarity (89%–100%) to insect PDFs and also to the crustacean β -PDH peptides. In particular, there is only a single amino acid difference between the PDFs of *Musca* and *Drosophila*; at position 14 Ser for *Musca* PDF and Asn for *Drosophila* PDF. A characteristic Ser¹⁰ in *Drosophila* was retained in *Musca*, indicating the presence of a structural profile unique to these PDFs. The results of sequence analyses suggest that *Musca* and *Drosophila* PDFs are to be considered members of a single group that has evolved structurally. When the primary structure of the PAP regions was compared, the *Musca* PDF precursor also belonged to the same group as that to which the *Drosophila* PDF precursor belongs. Copyright © 2003 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: cDNA cloning; circadian rhythm; neuropeptide; pigment-dispersing factor

Abbreviations: AAP, abridged anchor primer; AUAP, abridged universal anchor primer; La, lamina; LNs, lateral neurons; Me, medulla; NLS, nuclear localization signal; PAP, PDF-associated peptide; PCR, polymerase chain reaction; PDF, pigment-dispersing factor; PDH, pigment-dispersing hormone; RACE, rapid amplification of cDNA ends; RLM-RACE, RNA ligase-mediated rapid amplification of cDNA ends; RT-PCR, reverse transcription PCR.

* Correspondence to: Professor Yasuyuki Shimohigashi, Laboratory of Structure–Function Biochemistry, Department of Chemistry, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan; e-mail: shimoscc@mbox.nc.kyushu-u.ac.jp

The nucleotide sequence datum reported in this article has been submitted to the DDBJ/EMBL/GenBank Data libraries under accession number AB095922.

INTRODUCTION

Pigment-dispersing factor (PDF) is an 18-amino acid insect neuropeptide with the sequence of NSE(L/I)INSL(L/G/S)LPKX(L/M)NDA-NH₂, where X represents a variable amino acid comprising either Asp, Val or Leu. PDF is involved in regulating insect circadian rhythms [1–4]. From many lines of evidence aimed at clarifying the function of PDF in the circadian system of the fruitfly *Drosophila melanogaster*, PDF has been thought to be a neuropeptide functioning as a principal circadian modulator downstream of the brain's clock machinery, organizing daily locomotion or other behaviours [5–8].

Until recently, only three *pdf* cDNAs have been elucidated for the structure of insect PDF precursors: the precursors of *D. melanogaster* [9], *Romalea microptera* (GenBank) and *Gryllus bimaculatus* [10]. The C-terminal PDF regions of these possess almost the same peptide size and sequence (number of amino acid residues (*n*) = 20 or 21). However, the region containing the signal sequence and PDF-associated peptide (PAP) region obviously differ, with 80, 66 and 20 amino acid residues, respectively. This is in contrast to the structures of the precursor proteins of the crustacean β -PDH (β -pigment-dispersing hormone) peptides. The latter are homologues of PDF and the size of their precursor proteins is almost the same, about 70 amino acids, similar to that of the *Romalea* PDF precursor. The size differences between the PDF precursors might reflect different functions of the PAP peptides. In pursuit of a further understanding of such functions it is necessary to accumulate molecular information about PDF and its precursor.

The cDNA cloning of the PDF precursor of the last-summer cicada *Meimuna opalifera* has been achieved and it was found that the signal and PAP peptide region were very similar to that of *Romalea microptera* [11]. Thus, it was hypothesized that PDF precursors are classified into at least three different types: *Drosophila*, *Meimuna-Romalea* and *Gryllus*. In the present study, the cDNA cloning of PDF from adult brains of the housefly *Musca domestica* (Diptera) was carried out. It is often said that the gene products of flies resemble those of mammals and especially the products of the fruit fly *Drosophila*, an invertebrate, are said to be intermediate between those of the vertebrate and invertebrate groups. In this regard, the peptides and proteins of the housefly *Musca* must be siblings of

those in *Drosophila*. The housefly is a useful and important alternative to the fruit fly because of the size difference, for example, for *in vivo* studies. The sequence analyses of nucleotides and the deduced amino acids of *M. domestica* PDF are reported here. In order to characterize the PDF peptide and its precursor protein, further structural analyses were carried out to clarify the sequence.

MATERIALS AND METHODS

Animals

Animals, the housefly *Musca domestica* (a kind gift from Professor Ian A. Meinertzhagen, Dalhousie University, Halifax, Canada), were maintained at 25°C under a day: night cycle of L12:D12. To collect samples for cDNA cloning, the flies were frozen in liquid nitrogen and their heads were immediately harvested. The heads were stored at –80°C until use. For *in situ* hybridization, male adults older than 24 h were used.

3' RACE for Identification of 3' end of PDF mRNA

The cDNA cloning of *Musca pdf* was carried out essentially as previously described [10]. mRNAs were extracted using a QuickPrep[®] Micro mRNA Purification Kit (Amersham Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer's instructions. The mRNAs (350 ng) were reverse-transcribed by Super Script[™] II reverse transcriptase (Life Tech., Rockville, MD, USA) using d(T)₁₇-adapter primer (5'-GGCCACGCGTCTGACTAGTAC-T₁₇-3') as previously described [12]. The resulting cDNA were subjected to the polymerase chain reaction (PCR) using the adapter primer (5'-GGCCACGCGTCTGACTAGTAC-3') and the degenerate primer PDF-F1 (5'-AAGCGCAACTCNGARMTVA-TCAACTCN-3'). In these primers, the letters M, N, R and V denote the nucleotides (A or C), (any nucleotides), (A or G) and (not T), respectively. The primer was designed based on the nucleotide sequence homology among mRNAs clarified for β -PDH and PDF [9,13–16]. PCR was performed using PLATINUM *Taq* DNA polymerase (Life Tech.) with the following conditions: 3 min at 94°C followed by 3 cycles of 94°C for 30 s, 72°C for 90 s, 3 cycles of 94°C for 30 s, 68°C for 30 s, 72°C for 60 s, 3 cycles of 94°C for 30 s, 64°C for 30 s, 72°C for 60 s, 30 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 60 s, and a final extension for 7 min at 72°C. The

PCR products were subcloned into pBluescript II SK+ and the sequence was analysed by a Thermo Sequenase Cy5.5 dye terminator cycle sequencing kit (Amersham Pharmacia Biotech). The sequence analysed was confirmed by sequencing several other clones simultaneously.

5' RACE for Identification of 5' end of PDF mRNA

To amplify the 5' end of the *M. domestica pdf* cDNA, 5' RACE was carried out using Rapid amplification of cDNA ends (version 2.0: Life Tech.) mainly according to the manufacturer's protocol. First-strand cDNA was synthesized from a mRNA by Super Script™ II reverse transcriptase (Life Tech.) with a PDF-specific antisense primer (PDF-A1) (5'-ATCTTCTTCTTATCCATAGG-3'). The product was digested by RNase H/T₁ to remove the original mRNA template. A homopolymeric tail poly(C) was elongated by using terminal deoxynucleotidyl transferase and a substrate dCTP. The resulting tailed cDNA was amplified by the Abridged Anchor Primer AAP (5'-GGCCACGCGTCTGACTAGTACGGGIIGGGIIGGGIIG-3', Life Tech.) of the kit and PDF antisense primer (PDF-A2) (5'-CTAATGCTGCATGAAA TCTATTTGCC-3'). The PCR product was further amplified with an Abridged Universal Anchor Primer AUAP (5'-GGCCACGCGTCTGACTAGTAC-3', Life Tech.) and PDF antisense nested primer (PDF-A3) (5'-GCCGGC-ATCGTTCATACTTTTGGGC-3'). Respective PCR and sequence analyses were carried out as described above, and the sequence analysed was confirmed by sequencing several different clones simultaneously.

Based on the sequence information obtained on 3' RACE and 5' RACE analyses, reverse transcription PCR (RT-PCR) was performed to amplify the full-length *pdf* cDNA by PfuTurbo® DNA polymerase (Stratagene, CA, USA). The PCR conditions were as for PCR in the 3' RACE method, except for the primers, PDF-AF (5'-TCCATTCACCCGGAATCTCTG-3') and PDF-AR (5'-CACTTCATTCAACTAATACATCAACAAC-3'). These PCR products were treated as described above for sequence analysis.

To confirm the oligonucleotide sequence of the 5' end, the RLM-RACE method was applied to the *M. domestica pdf* cDNA [17]. This RNA ligase-mediated rapid amplification of cDNA ends is used to define the transcription start site of the *pdf* gene. RLM-RACE was performed using the Gene Traer™ RLM-RACE Kit (Invitrogen, CA, USA) according to the manufacturer's instructions. After reverse

transcription using the PDF-specific primer PDF-A1, the 5' end of *pdf* cDNA was amplified using the outer adaptor primer of the kit and PDF-specific primer PDF-A2. The nested PCR was performed using the inner adaptor primer of the kit and PDF-A3. All these PCR were carried out under the same conditions as for 5' RACE and amplified products were analysed as described above.

Whole Mount *in situ* Hybridization

RNA probes. Using T7 RNA polymerase, an anti-sense RNA probe labelled with digoxigenin (DIG)-UTP was produced *in vitro* from plasmid linearized with *Bam*HI. The plasmid used consisted of almost the full-length of *pdf* cDNA (463 bp in total) in the vector pBluescript II SK+. This *pdf* cDNA is a product of PCR using PDF-AF and PDF-AR. The reaction mixture of plasmid (1 µg), 5x buffer (4 µl; 0.2 M Tris•HCl (pH 8.0), 40 mM MgCl₂, 10 mM spermidine-(HCl)₃, 125 mM NaCl), 10x DIG RNA labelling mix (2 µl; 10 mM ATP, 10 mM CTP, 10 mM GTP, 6.5 mM UTP, 3.5 mM DIG-11-UTP (pH 7.5); Roche Diagnostics, Mannheim, Germany), 5 mM dithiothreitol (DTT) (2 µl), an RNase inhibitor RNasin® (1 µl; Promega, Madison, WI, USA), and T3 RNA polymerase for sense primer or T7 RNA polymerase for antisense primer (2 µl each (100 U) stored in 20 mM potassium phosphate (pH 7.7), 0.1 M NaCl, 0.1 mM DTT and 50% glycerol: Life Tech.) was incubated at 37 °C for 2 h. The reaction was terminated by adding 0.5 M EDTA (2 µl; pH 8.0). In order to precipitate DIG-labelled cRNA product, 4 M LiCl (2 µl) and ethanol (75 µl) were added and the solution was centrifuged at 17 000 g for 20 min at 4 °C. The residual pellet was washed with 70% ethanol and dried in air. The pellet was eventually dissolved in 50% formamide (50 µl). The quality of the transcript was analysed by electrophoresis on 1.5% agarose gel stained with ethidium bromide.

In order to prepare a sense RNA probe for the assay of negative control, essentially the same method was utilized except for the use of T3 RNA polymerase instead of T7 RNA polymerase. *Hind*III was used as a restriction enzyme.

Tissue hybridization. Whole mount *in situ* hybridization was performed essentially as described by Wilkinson [18] with several significant modifications. Briefly, the brains of the housefly *Musca domestica* were dissected under microscopy and fixed in freshly made 4% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4) for 12 h at 4 °C. Tissues

were washed with the phosphate buffered saline (PBS) with 0.1% Tween 20 (PBT) at 4 °C, dehydrated with a series of MeOH solutions (25%, 50%, 75% and 2x 100%) in PBT, and stored at -20 °C until use.

All the following procedures were performed at room temperature unless a specified temperature is noted. Tissues were hydrated in reverse in a series of MeOH-PBT solutions (75%, 50%, 25% MeOH, and 2x PBT). These rehydrated tissues were permeabilized with 10 µg/ml proteinase K at 37 °C for 15 min, washed with glycine (2 mg/ml) for 5 min, and then washed twice with PBT. They were fixed in 0.2% glutaraldehyde/4% paraformaldehyde in PBT for 20 min, washed twice with PBT, and incubated in PBT at 70 °C for 50 min. After cooling on ice, the tissues were treated with 6% H₂O₂ for 1 h, washed three times with PBS, and incubated with prehybridization buffer (50% formamide, 5x SSC (pH 4.5), 1% SDS, 50 µg/ml yeast tRNA and 50 µg/ml heparin) at 70 °C for 1 h. The prehybridization buffer was then replaced with hybridization buffer, in which the prehybridization buffer contains a DIG-labelled cRNA probe, for incubation at 70 °C for 16 h.

After hybridization, the tissues were washed three times at 70 °C with solution I (50% formamide, 5x SSC (pH 4.5), and 1% SDS) for 30 min, respectively, and once at 70 °C with a 1:1 solution of I and II (solution II consists of no formamide) for 10 min. The tissues were washed a further three times with solution II for 5 min, and then incubated with solution II for 20 min. Washings were repeated twice with solution III (50% formamide, 2x SSC (pH 4.5), and 1% SDS) for 5 min at 70 °C, three times with solution III for each 30 min, and with a TBST solution in which Tris-buffered saline (TBS; 150 mM NaCl, and 100 mM Tris-HCl (pH 7.5)) contains 1% Tween 20. After blocking at room temperature for 1 h, in order to visualize mRNA-bound probe by immunocytochemical detection the tissues were treated with Fab fragments of sheep anti-DIG antibody directly conjugated to alkaline phosphatase (AP). This incubation was done for 16 h at 4 °C, and the tissues were washed with TBST for 5 min successively four times and for 1 h seven times. After washing twice with NTMT (100 mM NaCl, 100 mM Tris-HCl (pH 9.5), 50 mM MgCl₂, and 1% Tween 20) for 5 min, the antibody detection reaction was performed by incubating the tissues in the substrate solution (NTMT with 0.33 mg/ml 4-nitroblue tetrazolium chloride and 0.17 mg/ml 5-bromo-4-chloro-3-indolylphosphate) for 30 min in shaded light, and finally washed with

TBST. Eventually, whole tissues were mounted in glycerol-gelatin (Sigma, St Louis, MO, USA) and observed by light microscopy with an Olympus BX50.

RESULTS

Preparation of a *Musca domestica* PDF cDNA

Adult houseflies, *Musca domestica*, were maintained in a daily cycle of L12:D12. Immediately after the flies were frozen in liquid nitrogen, their heads were collected. To isolate mRNA, the procedure designed for direct isolation of polyadenylated RNA without purifying the total RNA was used. Isolated mRNAs were then all converted to cDNAs by the RT-PCR method, in which the reverse transcriptase catalyses the polymerization of DNA using mRNA as templates.

3' RACE to Determine the 3'-end Sequence of a *Musca domestica* PDF cDNA

With respect to the nucleotides of insect PDF precursors, no prominent similarities were found except for the N-terminal portion of the PDF peptide regions. Thus, in order to screen for the *pdf* cDNA of the housefly *M. domestica*, the 3' RACE method was carried out. A PDF-specific degenerated sense primer for 3' RACE was designed with reference to sequence alignment analyses for both amino acid and nucleotide sequences. The N-terminal portion of PDF, namely, the residue -2 ~ 7 including the dibasic processing site (KR), was chosen as a PDF-specific sense primer (PDF-S1). PCR with a primer of PDF-S1 and an adapter primer Ad lacking d(T)₁₇ yielded no distinct gel-bands, and thus the resulting PCR product was used directly as a template for further PCR under the same conditions. Of the several PCR products detected by electrophoresis as bands in the agarose gel, only one PCR product was found to code a polynucleotide of about 250 bp with the characteristic 18-amino acid peptide, PDF. The nucleotide sequence determined also encoded a PDF amidation sequence in addition to the 3'-UTR non-coding region (Figure 1). The open-reading frame encoding PDF was highly conserved among *pdf* and *pdh* cDNAs.

5' RACE to Determine the 5'-end Sequence of a *Musca domestica* PDF cDNA

To determine the structure of the 5'-end of *pdf* cDNA, the 5' RACE method was undertaken. Based

1	AACAGTTACCTGATTTATTTTCGATCAGCGAGCACATATTTCCATTACCCCGAAATCTCTGCAATTATCATC	71
	PDF-AF	
72	ATGACTAATATTTGGTTATTTTTCATTGGCCCTATTTTGGATGTCTCTGCTGCTGTGCATGTGGCCACAGCG	143
1	M T N I G Y F S L A L F W M S L L L C H V A T A	24
144	TTGCCAGCCCCGGACGAAGAGCAGTACTTTGATAAACAGCTGAATCGAGAATTGATAAATCGATGGCTTTCA	215
25	L P A P D E E Q Y F D K Q L N R E L I N R W L S	48
216	TCCATCCATAATGCCCAGATATTAATAATAATCCATGCCGCTTCTATGGAGGGGACGGTACTTGGACCGCC	287
49	S I H N A Q I L N N N P C R F Y G G D G T W T A	72
	PDF-F1	PDF-A3
288	CCTTTGCCCAAACGAAACTCGGAACTAATCAATTCGTTGTTGAGTCTGCCCAAAGTATGAACGATGCCGGC	359
73	P L P K R N S E L I N S L L S L P K S M N D A G	96
360	AAATAGATTTTCATGCAGCATTAGAAAATACCTATGGATAAGAAGAAGATAAAAGACTAAAAAATATCCCCCT	431
97	K * PDF-A2 PDF-A1	
432	ACCCCTCCTTAACAAAAATATATGTATGTATATTGAATATTGTTGTTGATGTATTAGTTGAATGAAGTGAA	503
504	AATAAATAGTAGATATTCTGTTCC (a) _n	527
	PDF-AR	

Figure 1 Nucleotide sequence and its deduced amino acid sequence for the *Musca domestica* PDF precursor. Arrows indicate the primers used for the 3' RACE and 5' RACE methods. Amino acid sequences in white, grey and black boxes correspond to the signal, PAP and PDF regions, respectively. KR (=Lys-Arg) is a processing cleavage site, while the C-terminal GK (=Gly-Lys) is an amidation site. The underlined portion with the asterisk (TAG) is a stop codon. The double-underlined portion (AATAAA) in the 3'-UTR region is a putative poly-adenylation signal site.

on the sequence information obtained from 3' RACE, the antisense primer for RT-PCR against PDF mRNA was set at the 3'-UTR region (Figure 1). First-strand cDNA was synthesized from mRNA using a PDF-specific antisense primer PDF-A1 and a reverse transcriptase. A characteristic reaction of the 5' RACE method is an addition of a poly(C) tail to the 3'-end of cDNA obtained. The resulting poly(C)-tailed cDNA was utilized as a template for PCR using an anchor primer AAP, having a poly(I) truncated with deoxyguanosine, G/poly(I)/G, that enables the primer to anneal specifically to poly(C). PCR was carried out by using AAP and PDF-specific antisense primer PDF-A2, which was assigned at the 3'-UTR region. The anchor primer AAP possesses a specific 5'-end anchor sequence, and a subsequent nested PCR was performed using antisense PDF-A3 (assigned at the PDF region) and another anchor primer AUAP lacking G/poly(I)/G. This nested PCR provided a single product that was subsequently subcloned for sequence analysis.

The sequencing analysis afforded complete sequence data including 5'-UTR, signal region and PAP peptide region (Figure 1). The cDNA sequence was further confirmed by RT-PCR using primers set for both 5'-UTR and 3'-UTR to amplify the full-length PDF cDNA. Eventually, the full length of

the cDNA clone was found to comprise 527 bp, with a precursor protein of 97 amino acid residues (Figure 1). The initiator codon ATG was assigned as shown in Figure 1, since the selected ATG codon and its adjacent nucleotide sequences fulfilled Kozak's consensus motif [19]. No additional ATG codon was found in the upstream region of any cDNA clones isolated in the RT-PCR analysis. It should be noted that the RLM-RACE method afforded exactly the same oligonucleotide sequence of *pdf* cDNA.

The program for predicting a signal peptide cleavage site, SignalP, showed that 24 amino acids from a selected Met should, according to eukaryote data, be a signal peptide. The N-terminal Met-Thr-Asn-Ile-Gly-Tyr-Phe-Ser is unique, showing no sequence similarity to other PDF and β -PDH precursors. Only *Romalea microptera* PDF precursor exhibited an initiating sequence of Met-Thr, although this is followed by the sequence Ala-Met-Ala-Val-Ser-Gly [20].

Identification of *pdf* Expressing Cells

The whole-mount *in situ* hybridization was performed to identify the cells expressing *pdf* mRNA. The *M. domestica pdf* cDNA exhibited a normal segment construction as seen for *Drosophila* PDF

preproprotein, and this *pdf* gene encoded the PDF peptide in an ordinary form. It is important to locate the cells containing mRNA and to analyse the coincidence of the results from the *in situ* hybridization for *pdf* mRNA and immunocytochemistry for PDF peptide. The signals of *in situ* hybridization in *M. domestica* brain were observed only at the single region restricted in the medulla of the optic lobes (both right and left). In each optic lobe, a single cell group called large lateral neurons was found to be positive for hybridization. As shown in Figure 2, a group of immunoreactive cell clusters were identified by anti-DIG antibody conjugated to alkaline phosphatase, showing *pdf* mRNA expression in somata. In each optic lobe, there are two clusters containing at least four cells, respectively.

DISCUSSION

The *Musca* and *Drosophila* PDF Peptides Have a Highly Conserved Primary Structure

The unique amino acid residue of *Drosophila* PDF is Ser¹⁰, since Gly¹⁰ is preserved among the known sequences of all other arthropod species. It is now known that this unusual Ser¹⁰ is retained in *Musca* PDF. In some cases, Gly → Ser substitutions are evolutionarily significant in attaining a certain characteristic structure or conformation necessary for specific interactions. As described before, peptides and proteins of the fruit

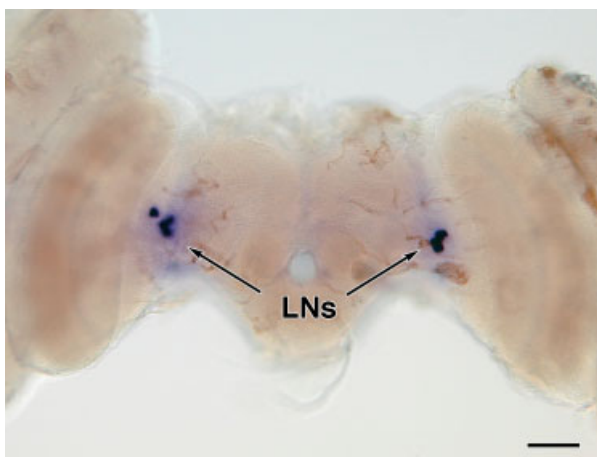


Figure 2 Spatial expression of *pdf* mRNA observed by the whole mount *in situ* hybridization. Only lateral neurons (LNs) in the anterior cortex of the medulla in the optic lobe show positive signals in this *in situ* hybridization. Scale 100 μ m.

fly *Drosophila* are often said to be intermediate between those of vertebrate and invertebrate groups. Insofar as this generalization holds true, the PDF and its precursor protein of the housefly *Musca* must be siblings of those in *Drosophila*. When the amino acid sequence of *Musca*'s PDF peptide, NSELINSLSLPKSMNDA-amide, was compared with that of *Drosophila* PDF, only a single residue substitution was found at position 14 (Table 1). The Asn¹⁴ residue in *Drosophila* PDF was replaced by Ser in *Musca* PDF. It should be noted that Ser and Asn are both hydrophilic, while all other insect PDFs and crustacean PDHs contain a hydrophobic Val, Leu, or Phe, at that position. All these features indicate that the PDF peptides in *Musca* and *Drosophila* evolutionarily are highly related to each other and different from all other PDF and PDH peptides.

Despite this difference, PDF in *Musca* is very similar to other ordinary PDFs and β -PDHs (Table 1). The amino acid sequence identity and similarity of *Musca* PDF were calculated to be 78%–94% and 89%–100%, respectively. A two-residue substitution was found between PDF in *Musca* and β -PDH of the crustacean *Penaeus vannamei* II; i.e. Ser¹⁰ \leftrightarrow Gly and Ser¹⁴ \leftrightarrow Val. On the other hand, one insect PDF and nine crustacean β -PDHs exhibited three-residue substitutions (Table 1). In addition to Ser¹⁰ \leftrightarrow Gly and Ser¹⁴ \leftrightarrow Val substitutions, Leu⁸ \leftrightarrow Ile or Leu¹¹ \leftrightarrow Ile are found when comparing *Musca* PDF and crustacean β -PDHs. The only exception is Asn¹⁶ \rightarrow Thr seen for *Penaeus japonicus* I β -PDH. As for PDF in the insect *Periplaneta americana*, the third substitution occurs at position 15 with the Met¹⁵ \rightarrow Leu replacement. Four-residue substitutions were found when comparing four insect PDFs and four crustacean β -PDHs (Table 1). In respect of the amino acid sequence, all crustacean α -PDHs are by contrast quite definitely different from *Musca* PDF.

Insect PDF Precursors Are Classified into Three Groups

It was previously hypothesized that PDF precursors are classified into three different types: the PDF precursors of *Drosophila*, *Romalea* and *Gryllus*. With the recent achievement of the cloning of the cDNA of the PDF precursor from the last-summer cicada, *Meimuna opalifera*, it was found that the PDF precursors of *Meimuna opalifera* and *Romalea microptera* are very similar to each other. Their peptide sizes are $n = 82$ and 89 , respectively, with high homology scores for sequence

Table 1 Amino Acid Sequences of Insect PDFs and Crustacean α - and β -PDHs

Number of residues different	Species	Peptide sequence	PDF or PDH	Ref.
	<i>Musca domestica</i>	NSELINSLLSLPKSMNDA-NH ₂	PDF	
1	<i>Drosophila melanogaster</i>	NSELINSLLSLPK <u>N</u> MNDA-NH ₂	PDF	[9]
2	<i>Penaeus vannamei</i> II	NSELINSLLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[16]
3	<i>Penaeus vannamei</i> I	NSELINSLLG <u>I</u> PKVMNDA-NH ₂	β -PDH	[16]
	<i>Periplaneta americana</i>	NSELINSLLG <u>L</u> PKV <u>L</u> NDA-NH ₂	PDF	[21]
	<i>Pandalus jordani</i> ^a	NSELINSLLG <u>L</u> PKV <u>M</u> TDA-NH ₂	β -PDH	[22,23]
	<i>Penaeus aztecus</i>	NSELINSLLG <u>I</u> PKVMNDA-NH ₂	β -PDH	[24]
	<i>Uca pugilator</i>	NSELINSLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[25]
	<i>Cancer magister</i>	NSELINSLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[26]
	<i>Pacifastacus leniusculus</i>	NSELINSLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[22]
	<i>Carcinus maenas</i>	NSELINSLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[13]
	<i>Callinectes sapidus</i> I	NSELINSLG <u>L</u> PKVMNDA-NH ₂	β -PDH	[15]
	<i>Penaeus japonicus</i> II	NSELINSLG <u>L</u> PK <u>F</u> MIDA-NH ₂	β -PDH	[27]
4	<i>Acheta domesticus</i>	NSE <u>I</u> INSLG <u>L</u> PKV <u>L</u> NDA-NH ₂	PDF	[28]
	<i>Romalea microptera</i>	NSE <u>I</u> INSLG <u>L</u> PK <u>L</u> LDA-NH ₂	PDF	[20]
	<i>Gryllus bimaculatus</i>	NSE <u>I</u> INSLG <u>L</u> PKV <u>L</u> NDA-NH ₂	PDF	[10]
	<i>Meimuna opalifera</i>	NSE <u>I</u> INSLG <u>L</u> PKV <u>L</u> NDA-NH ₂	PDF	[11]
	<i>Penaeus japonicus</i> I	NSELINSLG <u>I</u> PKV <u>M</u> TDA-NH ₂	β -PDH	[27]
	<i>Procambarus clarkii</i>	NSELINSLG <u>L</u> PKV <u>M</u> <u>N</u> EANH ₂	β -PDH	[29]
	<i>Orconectes limosus</i>	NSELINSLG <u>L</u> PKV <u>M</u> <u>N</u> EANH ₂	β -PDH	[14]
	<i>Orconectes immunis</i>	NSELINSLG <u>L</u> PKV <u>M</u> <u>N</u> EANH ₂	β -PDH	[22]
6	<i>Callinectes sapidus</i> II	NSELINSLG <u>I</u> S <u>A</u> L <u>M</u> N <u>E</u> ANH ₂	β -PDH	[15]
	<i>Armadillidium vulgare</i>	NSELINSLG <u>A</u> P <u>R</u> V <u>L</u> N <u>N</u> ANH ₂	β -PDH	[30]
7	<i>Pandalus jordani</i> ^a	NSG <u>M</u> INSLG <u>I</u> PKV <u>M</u> <u>A</u> DANH ₂	α -PDH	[22]
8	<i>Pandalus borealis</i>	NSG <u>M</u> INSLG <u>I</u> PRV <u>M</u> T <u>E</u> ANH ₂	α -PDH	[31]
	<i>Pandalus jordani</i> ^a	NSG <u>M</u> INSLG <u>I</u> PRV <u>M</u> T <u>E</u> ANH ₂	α -PDH	[22]

Capital alphabetical initials underlined indicate the amino acid residues that are not compatible with those of *Musca* PDF.

^a Three forms were reported for PDHs from the shrimp *Pandalus jordani*.

identity (60%) and similarity (78%). These results appeared to substantiate the hypothesis mentioned above. *Meimuna* — *Romalea* PDF precursors, or PAP regions appear to be grouped into a clear subfamily (Figure 3).

Comparing the nucleotide sequences of *pdf* cDNAs, the entire open-reading frame region of *Musca domestica* was found to be very similar to that of *Drosophila melanogaster* (ca. 60% identity). When the entire amino acid sequences were compared for the PDF precursors of the housefly and fruit fly, the homology scores of similarity and identity were calculated to be 62% and 38%, respectively, and thus very high. The number of amino acid residues, 98, for the *M. domestica* precursor was also similar to that of *D. melanogaster* ($n = 102$). Apparently, these two constitute another subfamily of PDF precursors, and this further substantiates the

hypothesis that PDF precursors can be considered as different subfamilies.

The peptide sizes of *Meimuna* — *Romalea* PDF precursors are fairly small, smaller than those of fly PDF precursors (Figure 3). The cDNA cloning of the PDF precursor of the cricket *Gryllus bimaculatus* has also been achieved. Its peptide size was likewise much smaller ($n = 43$) than those of the *Drosophila* — *Musca* ($n = 98$ – 102) and the *Meimuna* — *Romalea* groups ($n = 82$ – 89). Thus, the *Gryllus* PDF precursor can be considered a third subfamily, and at least for the moment, the PDF precursors are therefore classified into three subfamilies.

The C-terminal region of the precursor proteins is the PDF peptide *per se*. The size of the PDF peptide is fixed at $n = 18$, although the size of the peptide amide precursors can be $n = 20$ or 21 . This

Musca-Drosophila Group

M. domestica (97) MTNIGYFSLALFWMSLLLCHVATALPAPDEEQYFDKQLNRELI
D. melanogaster (102) MARFTYLVALVLLAICQWGYCGAMAMPDEERYVRKEYNRDLL
 *:.: *: :. . *:. *****:*. *: **:*:

M. domestica NRWLSSIHNAQILNN---NPCRF---YGGDGTWTAPLPKR NSELINSLLSLSPKSMNDAGK
D. melanogaster D-WFNNVGVGQFSPGQVATLCRYPLILENSLGPSPVIRKR NSELINSLLSLSPKSMNDAGK
 : *:.: . *:. . **:. .. :. *: ** *****.*****

Meimuna-Romalea Group

M. opalifera (82) MRS-A--GVMI AVLAVC---LCLCLESATSLRY
R. microptera (89) MTAMAVSGKLLTALVLS TYILGLALT-IQATQY
 * * * :.:.*.:. * *. * : : *

M. opalifera QDDKYIES--QYGPSTRELASWLEWAQKND---HAHKR NSEI INSLGLPKVLNDAGRK
R. microptera EEDKYQENEVKYG---RELASWLAQLAHKNEPAICAHKR NSEI INSLGLPKLLNDAGRK
 : :*** * . : ** ***** : * : ** : **** *****:*****

Gryllus Group

G. bimaculatus (43) MARRARFEANAAPSPLMCVHKR NSEI INSLGLPKVLNDAGRK

Figure 3 Comparison of the primary structures of insect PDF precursor proteins. The numbers in parentheses indicate the number of total amino acid residues of PDF precursors. The precursor of *Gryllus bimaculatus* has no signal peptide. * Residues in that column are identical in all sequences in the alignment. : Conserved substitutions have been observed. • Semi-conserved substitutions have been observed.

is because the precursors consist of a pre-amide sequence: $n = 2$ (GK) for *Drosophila* — *Musca* PDFs and $n = 3$ (GRK) for other insect PDFs. The region containing the signal and PAP is the peptide obtained by subtracting the PDF peptide from the entire precursor peptide. Thus, the difference in the size of the entire PDF precursors is the variance in the size of the signal/PAP regions. The peptide size of the signal/PAP regions is 76–80 for the *Drosophila* — *Musca*, 64–66 for the *Meimuna* — *Romalea*, and 20 for the *Gryllus* subfamilies. Such differences in peptide size might reflect the different functions of the PAP peptides in the different species.

The region of the signal peptides has been analysed by the computer-assisted signal peptide prediction method for eukaryotic sequences [32,33]. Assuming the result of this prediction, the peptide size of PAP was determined to be $n = 52$ for *Musca domestica* PDF precursor. Those of other insect PAPs were 56 for *Drosophila*, 34 for *Meimuna* and 32 for *Romalea*. No signal peptide was found for *Gryllus bimaculatus* PDF precursor, which, instead,

consists of a nuclear localization signal (NLS) at the N-terminal portion.

PDF peptide in *Drosophila* has recently been reported to transmit the rhythm of daily locomotor activity via output pathways from the circadian clock [6–8]. Our immunohistochemical study using a polyclonal antibody specific for *Musca* PDF reveals immunoreactivity in the axons and terminals of the *M. domestica* PDF cells, suggesting a similar role for PDF in this species too (unpublished data). To elucidate such a function for PDF, however, investigations to identify its specific receptor and to understand the details of the molecular mechanisms of peripheral circadian systems are first necessary.

pdf Expressing Cells

Pyza and Meinertzhagen have reported the results of immunocytochemistry for PDF peptide, using an antibody raised for the crustacean *Uca pugilator* PDH [3,4]. It was found that the cell bodies, namely somata in exactly the same places in the optic lobes are immunostained by both *in situ* hybridization

and immunocytochemical methods, suggesting that PDF is produced in the neurons where it functions. Figure 2 shows two groups of cell clusters of lateral neurons (LNs) in each optic lobe, and each group is likely to contain at least four cells, respectively. It should be noted that these groups appear to correspond to the large lateral neurons (l-LNs) and the small lateral neurons (s-LNs). These l-LNs and s-LNs cells in *Drosophila* express the *period* clock gene and PERIOD clock protein and produce PDF [34]. Thus, *Musca* PDF would reveal similar projections of circadian pacemaker cells within the brain.

Acknowledgements

The authors thank Professor Ian A. Meinertzhagen (Dalhousie University, Halifax, Canada) for his critical reading of this manuscript.

REFERENCES

1. Stengl M, Homberg U. Pigment dispersing hormone-immunoreactive neurons in the cockroach *Leucophaea maderae* share properties with circadian pacemaker neurons. *J Comp. Physiol. [A]* 1994; **175**: 203–213.
2. Petri B, Stengl M. Pigment-dispersing hormone shifts the phase of the circadian pacemaker of the cockroach *Leucophaea maderae*. *J. Neurosci.* 1997; **17**: 4087–4093.
3. Pyza E, Meinertzhagen IA. Neurotransmitters regulate rhythmic size changes amongst cells in the fly's optic lobe. *J. Comp. Physiol. [A]* 1996; **178**: 33–45.
4. Pyza E, Meinertzhagen IA. Neurites of *period*-expressing PDH cells in the fly's optic lobe exhibit circadian oscillations in morphology. *Eur. J. Neurosci.* 1997; **9**: 1784–1788.
5. Helfrich-Förster C, Stengl M, Homberg U. Organization of the circadian system in insects. *Chronobiol. Int.* 1998; **15**: 567–594.
6. Renn SCP, Park JH, Rosbash M, Hall JC, Taghert PH. A *pdf* neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in *Drosophila*. *Cell* 1999; **99**: 791–802.
7. Park JH, Helfrich-Förster C, Lee G, Liu L, Rosbash M, Hall JC. Differential regulation of circadian pacemaker output by separate clock genes in *Drosophila*. *Proc. Natl Acad. Sci. USA* 2000; **97**: 3608–3613.
8. Helfrich-Förster C, Täuber M, Park JH, Mühligh-Versen M, Schnewly S, Hofbauer A. Ectopic expression of the neuropeptide pigment-dispersing factor alters behavioral rhythms in *Drosophila melanogaster*. *J. Neurosci.* 2000; **20**: 3339–3353.
9. Park JH, Hall JC. Isolation and chronobiological analysis of a neuropeptide pigment-dispersing factor gene in *Drosophila melanogaster*. *J. Biol. Rhythms* 1998; **13**: 219–228.
10. Chuman Y, Matsushima A, Sato S, Tomioka K, Tominaga Y, Meinertzhagen I, Shimohigashi Y, Shimohigashi M. cDNA cloning and nuclear localization of the circadian neuropeptide designated as pigment-dispersing factor PDF in the cricket *Gryllus bimaculatus*. *J. Biochem.* 2002; **131**: 895–903.
11. Sato S, Chuman Y, Matsushima A, Tominaga Y, Shimohigashi Y, Shimohigashi M. A circadian neuropeptide, pigment-dispersing factor — PDF, in the last-summer cicada *Meimuna opalifera*: cDNA cloning and immunocytochemistry. *Zool. Sci.* 2002; **19**: 821–828.
12. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA* 1988; **85**: 8998–9002.
13. Klein JM, de Kleijn DPV, Keller R, Weidemann WM. Molecular cloning of crustacean pigment dispersing hormone precursor. *Biochem. Biophys. Res. Commun.* 1992; **189**: 1509–1514.
14. de Kleijn DPV, Linck B, Klein JM, Weidemann WM, Keller R, van Herp F. Structure and localization of mRNA encoding a pigment dispersing hormone (PDH) in the eyestalk of the crayfish *Orconectes limosus*. *FEBS Lett.* 1993; **321**: 251–255.
15. Klein JM, Mohrherr CJ, Sleutels F, Riehm JP, Rao KR. Molecular cloning of two pigment-dispersing hormone (PDH) precursors in the blue crab *Callinectes sapidus* reveals a novel member of the PDH neuropeptide family. *Biochem. Biophys. Res. Commun.* 1994; **205**: 410–416.
16. Desmoucelles-Carette C, Sellos D, Van Wormhoudt A. Molecular cloning of the precursors of pigment dispersing hormone in crustaceans. *Ann. NY Acad. Sci.* 1998; **839**: 395–396.
17. Schaefer BC. Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal. Biochem.* 1995; **227**: 255–273.
18. Wilkinson DG (ed.). *In Situ Hybridization: A Practical Approach*. IRL Press: Oxford, 1992.
19. Kozak M. Initiation of translation in prokaryotes and eukaryotes. *Gene* 1999; **234**: 187–208.
20. Rao KR, Mohrherr CJ, Riehm JP, Zahn CA, Norton S, Johnson L, Tarr GE. Primary structure of an analog of crustacean pigment-dispersing hormone from the lubber grasshopper *Romalea microptera*. *J. Biol. Chem.* 1987; **262**: 2672–2675.
21. Mohrherr CJ, Rao KR, Riehm JP. Characterization of a pigment-dispersing factor from the american cockroach. *Soc. Neurosci. Abstr.* 1991; **17**: 276.
22. Rao KR, Kleinholz LH, Riehm JP. Characterization of three forms of pigment-dispersing hormone from the

- shrimp *Pandalus jordani*. *Soc. Neurosci. Abstr.* 1989; **15**: 26.
23. Rao KR, Riehm JP. Pigment-dispersing hormones. *Ann. NY Acad. Sci.* 1993; **680**: 78–88.
24. Philips JM, Rao RK, Riehm JP, Morgan WT. Isolation and characterization of a pigment-dispersing hormone from the shrimp *Penaeus aztecus*. *Soc. Neurosci. Abstr.* 1988; **14**: 534.
25. Rao KR, Riehm JP, Zahnow CA, Kleinholz LH, Tarr GE, Johnson L, Norton S, Landau M, Semmes OJ, Sattelberg RM, Jorenby WH, Hintz MF. Characterization of pigment-dispersing hormone in the eyestalks of the fiddler crab *Uca pugilator*. *Proc. Natl Acad. Sci. USA* 1985; **82**: 5319–5322.
26. Kleinholz LH, Rao KR, Riehm JP, Tarr GE, Johnson L, Norton S. Isolation and sequence analysis of a pigment-dispersing hormone from eyestalks of the crab, *Cancer magister*. *Biol. Bull.* 1986; **170**: 135–143.
27. Yang W-J, Aida K, Nagasawa H. Characterization of chromatophorotropic neuropeptides from the kuruma prawn *Penaeus japonicus*. *Gen. Comp. Endocrinol.* 1999; **114**: 415–424.
28. Rao KR, Riehm JP. The pigment-dispersing hormone family: Chemistry, structure-activity relations, and distribution. *Biol. Bull.* 1989; **177**: 225–229.
29. MacCallum ML, Rao KR, Riehm JP, Mohrherr CJ, Morgan WT. Primary structure and relative potency of an analog of β -PDH (pigment dispersing hormone) from the crayfish *Procambarus clarkii*. *Pigment Cell Res.* 1991; **4**: 201–208.
30. Knowles AC. Master thesis. 1992; The University of West Florida, Pensacola, FL.
31. Fernlund P. Structure of a light-adapting hormone from the shrimp *Pandalus borealis*. *Biochim. Biophys. Acta* 1976; **439**: 17–25.
32. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 1997; **10**: 1–6.
33. Nielsen H, Brunak S, von Heijne G. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Eng.* 1999; **12**: 3–9.
34. Helfrich-Förster C. The period clock gene expressed in central nervous system neurons which also produce a neuropeptide that reveals the projections of circadian pacemaker cells within the brain of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* 1995; **92**: 612–616.