cDNA Cloning and Stage-Dependent Expression of Arylphorin Gene from Chinese Oak Silkworm, *Antheraea pernyi*

Sang Mong Lee, Jae Sam Hwang, Nam Sook Park, Yong Gyun Kim, Keun Ki Kim, Hong Joo Son, Hyun Chul Park and Byung Rae Jin

Department of Life Science and Environmental Biochemistry, College of Natural Resources and Life Science, Pusan National University, Miryang 627-703, Korea

1Department of Agricultural Biology, National Academy of Agricultural Science, Rural Development Administration, Suwon 441-100, Korea

2College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Korea

Received July 6, 2010 / Accepted August 10, 2010

The cDNA cloning and developmental profiles of the mRNA for *A. pernyi* arylphorin was determined. The complete *A. pernyi* arylphorin cDNA sequence comprised 2,234 bp (without the poly A⁺ tail), including an open reading frame of 2,112 bp beginning with a methionine ATG at bp 34. The *A. pernyi* arylphorin contained 704 amino acids which are highly enriched in aromatic amino acids, phenylalanine and tyrosine. The calculated molecular mass of the *A. pernyi* arylphorin from the ORF was 83,439 Da. The deduced amino acid sequence of *A. pernyi* arylphorin showed 78, 71, 62 and 64% identity with those of *H. cecropia*, *M. sexta* α subunit, *M. sexta* β subunit and *B. mori* storage protein. In Northern blot analysis, the *A. pernyi* arylphorin mRNA only in the fat body of the 5th instar larvae was responsible for gene expression of the protein, and the synthetic activity of the mRNA was detected strongly in the early larvae, but not in the middle or late-stage larvae. In addition, a very weak signal in mRNA activity was detected in pupal stages, but this was considered to be inactive mRNA after reviewing the results of the labeling experiment of this protein.

**Key words**: Insect, Chinese oak silkworm, cDNA, arylphorin, *Antheraea pernyi*, storage protein

Introduction

The hemolymph proteins in insects show dynamic changes in their species, titer and stage-dependent synthesis during the developmental stages and metamorphic transformation. Investigations on the insect hemolymph proteins are very important to understand the physiology of the insect, even though there are many other important researches in hemolymph or tissue. A lot of newly obtained information on the structure, site of synthesis, rate of synthesis, developmental timing of synthesis, and gene structure for several hemolymph proteins are contributed for understanding the physiological functions of these proteins in insects. Therefore, a variety of insect hemolymph proteins such as storage proteins [29-32], lipophorin [4,5], vitellogenin and yolk proteins [3,8,9,14,34], antibacterial proteins [11], lectins [6], protease inhibitors [13], enzymes [2], peptides [16], chromoproteins [10], and specific transport proteins [7] were investigated and reported up to the present time. The term 'storage protein' implies uptake from the hemolymph and storage in the fat body tissue. A clear definition of these insect storage proteins has been suggested as following brief description [26]. First, these proteins occur only in the larval stages where they accumulate in the hemolymph. Second, they are synthesized predominantly, if not exclusively, by the larval fat body. Third, their concentration increases enormously in the final larval instar. Remnants of the insect storage proteins may carry over into early adult life, but within a few days after adult eclosion no trace of these proteins can be detected. Over the above biological definition, additional two biochemical criteria can be added. These families of storage proteins reported to the present day have molecular weight in the range of around 500,000 Da and are composed of six subunits. Furthermore, nearly all of these proteins are highly aromatic, containing a high proportion of tyrosine and phenylalanine residues [21]. A lot of studies have been concentrated on the characterization of the above insect storage hemolymph proteins from a variety of insect species since the first discovery of the prototype storage protein (calliphorin) from *Calliphora erythrocephala* [1,18]. Especially, arylphorin is the best-studied storage protein in insects. The arylphorin is characterized by higher content of aromatic amino acids, e.g. tyrosine and phenylalanine (18
The synthesis and utilization of the insect storage proteins are very important events linked to the metamorphosis of holometabolous insects. Insect storage proteins are synthesized in fat body, secreted into the larval hemolymph and taken by the fat body shortly before pupation.

In the end, they can be supplied as amino acids reserve necessary for the completion of adult development.

In a previous research, we also reported on purification and characterization of arylphorin from the Chinese oak silkworm, Antheraea pernyi [19,25]. The native molecular weight of arylphorin was 450 kDa with an 80 kDa single subunit, suggesting hexamer. The protein contained high amount (18.3%) of aromatic amino acids, phenylalanine (9.7%) and tyrosine (8.6%). Therefore, the protein was identified as a kind of a storage protein referred to as an arylphorin. The protein was stained by Schiff’s reagent, suggesting a glycoprotein containing 4.9% (w/w) sugar and mannose, of which major components were N-acetylglucosamine. These results showed that the A. pernyi arylphorin from our previous study was hexamer associated with the six subunits consisting of an 80 kDa single subunit. The Chinese oak silkworm, Antheraea pernyi, is a lepidopteran and an industrially important wild silkworm in China and Far Eastern Asia. However, there is no report to describe cDNA cloning and stage-dependent expression of A. pernyi arylphorin. Accordingly, cDNA sequence information for the storage protein identified as arylphorin in the previous study of A. pernyi will be valuable for the understanding metamorphosis for the adult development of the wild oak silk moth. Therefore the present study describes cDNA cloning and stage-dependent expression of the A. pernyi arylphorin.

Materials and Methods

Insects

The Chinese wild oak silkworm, A. pernyi was reared indoors with the fresh leaves of Japanese oak tree, Quercus acutissima, at room temperature under the laboratory condition of a natural photoperiod.

N-terminal amino acid sequencing

Protein samples were passively transferred on to a polyvinylidene difluoride membrane after electrophoresis using 10 mM CAPS pH 4.5 as the transfer buffer. Amino acid sequencing was performed with an automatic amino acid sequencer (Korea Basic Science Institute).

Sequence of N-terminal peptide and cDNA encoding the arylphorin gene

For the result of SDS-PAGE as shown in previous research [25], the molecular weight of arylphorin protein was determined to be about 80 kDa using molecular weight marker. By Edman degradation procedure, 8 amino acids sequence of the N-terminal was determined to be SVHITTP. Based on that sequence, degenerate primer (5’-AGC GTG GTG CAT CCG CCG CCG CAT-3’ and 5’-AGC GTK GTK CAY CCG CCG CCG CAY-3’) was designed. Arylphorin cDNA fragment was amplified by 3’-RACE and 5’-RACE-PCR.

RNA isolation and Northern blot analysis

RNAs were isolated from the whole body, fat body, mid gut and silk gland of the A. pernyi larvae by using the Total RNA Extraction Kit (Promega, Madison, WI). Total RNAs (10 μg/lane) from the A. pernyi larvae were denatured by glyoxalation (McMaster and Carmichael, 1977), transferred onto a nylon blotting membrane (Schleicher & Schuell, Dassel, Germany) and hybridized at 42°C with a probe in a buffer containing 2 X PIPES, 50% formamide, 1% sodium dodecyl sulphate (SDS) and blocking agent (Boehringer Mannheim, Mannheim, Germany). The probe used to detect the arylphorin gene transcript was a 2,112 bp A. pernyi arylphorin gene amplified by PCR in this study and labeled with [32P]dCTP (Amersham, Arlington Heights, IL) using the Prime-It II Random Primer Labeling Kit (Stratagene, La Jolla, CA). After hybridization, the membrane filter was washed three times for 30 min. each in 0.1% SDS and 0.2×SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) at 65°C, and finally exposed to X-ray film.

Reverse transcription polymerase chain reaction (RT-PCR)

Poly(A)+RNA was isolated from fat-body of A. pernyi the using Poly(A)+ Extraction Kit MicroFastTrack (Invitrogen) according to the protocol provided by the manufacturer and reverse transcribed with oligo (dT) primer using First-Strand cDNA Synthesis Kit (SuperScript RT, GIBCO BRL).

Degenerate oligonucleotides primers were designed on the basis of the N-terminus peptide sequences of the
Fig. 1. cDNA sequence and deduced amino acid sequence of the arylphorin. A), Location of the primers used in the study; B), Nucleotide sequence and deduced amino acid from arylphorin cDNA. N-term-Dp1 (N-t-Dp1), Degenerate primer designed from N-terminus peptide sequences of A. pernyi Gp1 and Gp2. Gene specific primers selected from a published A. pernyi partial gene [28]; Ap1, Adapter primer (Adapter sequence attached to the ends of the cDNA enabled it to be used in 5' and 3' RACE); RACE, Rapid amplified cDNA ends. Primer sequences used in the study are as follows. Ap1, 5'-CTAATACGACTCACTATAGGGCTCGAGCGGC CGCCCGGGCAAGGT-3'; N-t-Dp1, 5'-GTGTTGTACACCCACCACCTCA-3'; GP1, 5'-CACATGTAACATCTTCAACCCGG-3'; GP2, 5'-GGAACGTCGTGGTGAAATGTATCATC-3'.
Fig. 2. Multiple alignment of the *A. pernyi* arylphorin amino acids with other insect arylphorin types. The predicted amino acid sequences of *A. pernyi* arylphorin were aligned with arylphorins of *Hyalophora cecropia* (GenBank sequence ID AF032396), *Manduca sexta* α subunit (GenBank sequence ID M28396), *M. sexta* β subunit (GenBank sequence ID M28396) and *Bombyx mori* (GenBank sequence ID A34287) using the Clustal W algorithm. The deduced amino acid sequences of arylphorin from *A. pernyi* showed 78%, 71%, 62% and 64% identity with those of *H. cecropia* *M. sexta* α subunit, *M. sexta* β subunit and *B. mori* storage protein.

arylphorin. These primers were used for RT-PCR. The annealing temperature was 55°C in 30 cycles. PCR products were isolated by electrophoresis and cloned into the pGEM T Easy Vector (Promega).
gene fragment of *A. pernyi*. These primers were used in conjunc-
tion with the anchor primer API (Clontech) to amplify
the 5' and 3' ends of the arylphorin gene from cDNA. The
PCR conditions were 94°C for 3 min, followed by 30 cycles
of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min and a
final elongation at 72°C for 5 min. The PCR products were
isolated by electrophoresis and cloned into the pGEM™-Easy
Vector. The DNA sequence was determined by cycle se-
quencing using a Prism Sequencing Kit (Perkin-Elmer)
and an automated DNA sequencer (Model ABI 377).

**Results**

Sequences of cDNA encoding the arylphorin

We sequenced a cDNA clone encoding the arylphorin
which was isolated from the fat body RNA by RT-PCR and
5' and 3' RACE, and its inferred amino-acid sequences are
shown in Fig. 1. As shown in the complete nucleotide se-
quence in Fig. 1, there is a 5'-untranslated region of 33 nu-
cleotide sequences followed by an initiating ATG codon. The
TAA termination codon occured at nucleotide 2,145, thus
translation of the sequence from nucleotides 34 to 2,145
would produce a protein of amino acids with a calculated
molecular mass of 83,439. The complete *A. pernyi* arylphorin
cDNA sequence comprised 2,234 bp (without poly A* ”*tail"
including an open reading frame of 2,112bp beginning with
a methionine ATG at bp 34. The *A. pernyi* arylphorin con-
tained 704 amino acids which are highly enriched in ar-
omatic amino acids, phenylalanine and tyrosine. The
AAATTAAA consensus polyadenylation signal is present at 70
nucleotides downstream from stop codon (TAA). The poly
(A)* tail is 16 nucleotides downstream from the recognition
sequence, which is in agreement with the fact that the signal
as most open present at 11-30 nucleotide upstream from the
poly (A)* tail. In addition, there are four N-glycosylation
sites (Asn-X-Thr/Ser), at 84 to 86 (NYT), 223 to 225 (NYS),
371 to 373 (NDS), 651 to 653 (NLT) in the amino acid
sequence. Also the deduced amino acid sequence of ar-
ylphorin from *A. pernyi* showed 78%, 71%, 62% and 64%
identity with those of *Hyalomorpha cecropia*, *Manduca sexta* a
subunit, *M. sexta β* subunit and *Bombyx mori* storage protein
(Fig. 2). In addition, phylogenetic background was analysed
by statistical analysis (Fig. 3). The *A. pernyi* arylphorin was
grouped more related into the categories of the lepidopter-
an insect such as *B. mori*, *B. mandarina*, *A. yamamai* and *M.
sexta*.

Fig. 3. Phylogenetic tree based on the full or partial cDNA se-
quence of arylphorins. The genetic distance and phyl-
ogenetic tree were constructed using UPGMA clustering
method by MEGA programme. The sequences used in
the tree were downloaded from GenBank file. *C. kicidud-
valli*, GenBank sequence ID D44888; *A. aemmonis*, GenBank
sequence ID D44882; *A. aegypti*, GenBank sequence ID
U86080; *B. mandarina*, [17]; *A. yamamai* [17]. The se-
quences for arylphorin genes of *B. mandarina*, *A. yam-
mai*, *C. kiciduvalli*, *A. aemmonis* and *A. aegypti* used in the
analysis are partial ones. See Fig. 2 for further legend.

Developmental expression of the arylphorin gene

The developmental expression of arylphorin was ana-
lysed by Northern blot analysis during the larval and pupal
developmental stages. The presence of the mRNA for the
arylphorin in total RNA extracts in whole body, fat body,
midgut, middle silk gland and posterior silk gland from the
wild silk moth, *A. pernyi*, during the different developmental
stages was tested by hybridization with the arylphor-
in-cDNA probe (Fig. 4). For this, whole body, fat body, mid-
gut, middle silk gland and posterior silk gland were used
for the extraction of the total RNA in the various de-
velopmental stages. As seen in Fig. 4, the transcripts were
detected in the whole body and fat body extract during the
periods from the 3rd instar larvae to the middle pupae, and
the amount of the messenger always showed higher amount
in fat body than in whole body extracts due to the relative
abundant concentration of the arylphorin-mRNA in the fat
body tissues.

The transcripts detected firstly from the 3rd instar larvae
were reached a higher abundance at early and middle
larval stages in the 5th instar, and gradually decreased
thereafter, that is, reached a minimum abundance during
the period from the late larvae to early and middle pupae.
Finally, the transcript was detectable at trace level at the
late pupae.
which shows higher molecular mass compared with 80 kDa determined by SDS-PAGE analysis [25]. The gene structure of the arylphorin in genomic DNA was not elucidated, but subsequent investigation will be followed in the near future. At present, many cDNAs for insect arylphorins were cloned and sequenced in several insects [12,15,20,22,23,33]. We cloned the complete cDNA for arylphorin identified as A. pernyi storage protein from its several molecular properties. We compared the deduced amino acid sequences with those of other lepidopteran insects. The sequence identity is 62~78% among those lepidopteran arylphorin sequences (H. crista, [29]; B. mori, [15]; M. sexta, [33]). Whether this is in agreement with biochemical and immunological results on the relatedness of the four arylphorins including that of A. pernyi in the present study will be discussed in the near future, but, the Antheraea pernyi arylphorin showed same immunological identity with those of the wild silkworms, Antheraea yamamai, Samia Cynthia pryeri and Actias gnoma in the previous study [24]. From the results, cDNA cloned and sequenced for arylphorin of A. pernyi was found to have very similar structure at the molecular and immunological level with those of the lepidopteran insect species mentioned above. Surprisingly, even though the mRNA is still present in the pupae, the corresponding protein was not synthesized during this pupal stage (data not shown in the present study) [27]. This result is very interesting because the arylphorin is not synthesized during the pupal stages or detected at trace level (data not shown in the present study) [27]. Even though the presence of the arylphorin is proved during the developmental periods from the 3rd larvae to the late pupae by PAGE and Western blot analysis, it is elucidated by protein labeling using [14C]-leucine that the synthesis of the arylphorin is limited to the developmental stage of only 5th instar larvae (of course, it is certain that the synthesis is possible during the period of the 1st to 4th instar larvae). Also, from these results, it is possible that the mRNA disappears very slowly during the larval to pupal developmental stages, and the transcript (mRNA for arylphorin) present in the pupal stage can be an inactive form of the corresponding mRNA.

Discussion

cDNA encoding the arylphorin of the A. pernyi storage protein was cloned from the fat body mRNA of the A. pernyi larvae in the 5th instar stage (Fig. 1). The cDNA for arylphorin comprised 2,234 bp and was deduced as 704 amino acids. From the deduced amino acid sequences, the molecular weight of arylphorin was exactly calculated as 83,439 Da.
References


초록: 작잠(Antheraea pernyi) 아릴포린(Arylphorin) 유전자의 cDNA 클로닝 및 아릴포린 유전자의 발육시기 의존성 발현양상

상수리 잎을 먹고 자라는 야생견사곤충의 일종인 작잠의 저장단백질인 아릴포린 유전자의 cDNA를 클로닝하고 발육경과에 따른 유전자발현의 양상을 조사 검토하였다. 작잠 아릴포린 유전자의 cDNA는 2,112 bp의 ORF (open reading frame)를 포함하여 2,234 bp임을 밝혔다. 작잠 아릴포린 유전자의 cDNA염기서열로부터 아미노산 서열을 검토한 결과 방향족 아미노산인 베타알라닌(phenylalanine)과 티로신(tyrosine)의 성분이 높은 아미노산 서열구조를 보였으며 ORF로부터 계산한 단백질의 분자량은 83,439 Da 이었다. 작잠 아릴포린 저장단백질의 아미노산서열을 다른 곤충의 서열과 그 상동성을 비교 분석한 결과 세크로피아잠(H. cecropia)과는 78% 담배나방의 알파단량체(M. sexta-α subunit)와는 71% 담배나방의 베타단량체(M. sexta β-subunit)와는 62% 그리고 가지(M. nuxi)와의 알파단량체(M. sexta-α subunit)와는 71%, 담배나방의 베타단량체(M. sexta β-subunit)와는 62% 그리고 가지(M. nuxi)와의


