

Molecular cloning of the gene for the allatostatin family of neuropeptides from the cockroach *Diploptera punctata*

(juvenile hormone/endoproteolytic processing/neurobiology)

B. C. DONLY*, Q. DING†, S. S. TOBE†, AND W. G. BENDENA*

*Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada; and †Department of Zoology, University of Toronto, Toronto, ON Canada M5S 1A1

Communicated by Mary Lou Pardue, June 16, 1993

ABSTRACT Allatostatins (ASTs) are insect neuropeptides that inhibit juvenile hormone biosynthesis by the corpora allata. We have isolated a cDNA from the cockroach *Diploptera punctata* that encodes a 41.5-kDa precursor polypeptide containing the AST family of peptides. Translation of the cDNA revealed a 370-amino acid pre-pro-peptide consisting of 13 AST-type peptides and appropriate processing sites for endoproteolytic cleavage and amidation. The 13 potential AST sequences are characterized by the C-terminal AST core-structure Phe-Gly-Leu-NH₂, with only one exception. Separating the clustered ASTs in the precursor, three acidic spacer regions are found. Contained within the largest of these are two potentially related peptides that may also be processed. Southern blot analysis revealed the presence of a single copy of the AST gene per haploid genome, as well as the probability that the gene may be present in at least two allelic forms. *In situ* hybridization indicated the AST-encoding gene is expressed in neurosecretory cells of *D. punctata* brain.

Juvenile hormones (JHs) are sesquiterpenoid molecules produced by the corpora allata (CA) that play a central role in the metamorphosis and reproduction of most insect species. The presence of JH is required for maintenance of larval characteristics, whereas its absence results in the transition from larva to adult. In the adult female of most species, JH is also required for oocyte development (1).

JH biosynthesis is probably regulated, in part, by a family of brain-derived neuropeptide amides called allatostatins (ASTs), which inhibit JH production by the CA (2). Seven ASTs ranging in size from 8 to 18 amino acids have been purified from the cockroach *Diploptera punctata* (3–5). All share the common C-terminal sequence Tyr-Xaa-Phe-Gly-Leu-NH₂, which has been demonstrated *in vitro* to be the minimum sequence capable of inhibiting JH biosynthesis (6, 7). The cAMP-independent signal-transduction mechanism(s) through which ASTs exert their effect on the JH biosynthetic pathway is unclear (8). However, membrane receptors for ASTs have recently been identified (9). ASTs have also been identified and purified from muscle tissue, suggesting that these molecules may have other roles apart from regulation of JH production (10).

To date, a limited number of insect neuropeptide genes have been characterized (11–15), despite the potential for application to pest control and management (16, 17). The recent characterization of several AST peptides from *D. punctata* has provided the opportunity for application of molecular approaches to isolate the gene responsible for production of these factors. Also, in the event that a family of small peptides might be coded for in the form of a single large precursor, as for the *Drosophila* Phe-Met-Arg-Phe-NH₂ (FMRFamide) group (18), gene characterization provides the

opportunity to quickly identify additional members of the peptide family. We report here the characterization of a cDNA from the cockroach *D. punctata* that encodes such a precursor for the AST family of neuropeptides.† The deduced precursor contains at least 13 potential ASTs, including all seven of the peptides previously isolated.

MATERIALS AND METHODS

DNA Synthesis and Sequencing. Oligonucleotides were synthesized on a Biosearch model 8750 DNA synthesizer at the Core Facility for Protein/DNA Chemistry (Queen's University). Templates for DNA sequencing were prepared by PCR amplification of single-stranded cDNA followed by purification on 1% low-melting-temperature agarose gels. Resulting templates were sequenced by the dideoxynucleotide chain-termination method using either of two procedures: (i) Sequenase version 2.0 DNA sequencing kit (United States Biochemical) or (ii) *Taq* DyeDeoxy Terminator Cycle sequencing kit from Applied Biosystems (Core Facility for Protein/DNA Chemistry).

RNA Isolation and cDNA Synthesis. Brains were dissected from the heads of day 2–3 virgin female cockroaches (*D. punctata*) and frozen immediately in liquid nitrogen. Poly(A)⁺ RNA was directly extracted from 1000 brains by using the Pharmacia QuickPrep mRNA purification system and then stored under ethanol at –20°C.

Synthesis of first- and second-strand cDNA from the RNA was done by using the Zap-cDNA synthesis system (Stratagene). The resulting vector-ready cDNA was then ligated to a compatible plasmid, pBluescript II SK⁻ (Stratagene), using T4 DNA ligase (19). The circularized cDNA was diluted 5-fold with water and used as a pool for subsequent PCR.

PCR Methods. Standard PCR was done in 100- μ l volumes containing 2.5 μ l of cDNA (above), 1 \times PCR buffer (Perkin-Elmer/Cetus), 200 μ M dNTPs, bovine serum albumin at 100 μ g/ml, 2.5 units of Amplitaq DNA polymerase (Perkin-Elmer/Cetus), and, when using degenerate primers, one Ampliwax PCR Gem. Unique oligonucleotide primers were added to 0.25 μ M concentration, whereas degenerate ones were used at 2.5 μ M. To produce larger quantities of clean DNA product for sequencing template, PCR reactions were routinely run as nested pairs, using a second set of oligonucleotide primers inside the first to improve specificity.

Temperature-cycling profiles varied according to the oligonucleotide primers being used but were generally based on the following model: 95°C for 2.5 min, 50–65°C for 1 min, and 72°C for 1 min per kb of expected product, followed by 34 cycles of 95°C for 45 sec, 50–65°C for 1 min, and 72°C as in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AST, allatostatin; CA, corpora allata; JH, juvenile hormone.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. U00444).

the first cycle. A final 5-min extension at 72°C was also included.

Southern Blotting. High-molecular-weight DNA was purified from adult *D. punctata* essentially as outlined in ref. 20. Aliquots containing 10 µg of extracted DNA were digested with selected restriction enzymes and then separated on a 0.7% agarose gel. The DNA was denatured and transferred under vacuum to Hybond-N⁺ nylon membrane.

Hybridization was done at 65°C in 5× standard saline phosphate/EDTA (SSPE)/10× Denhardt's solution/0.5% SDS/denatured salmon sperm DNA at 250 µg/ml for 16 hr (19). The hybridization probe was a random-primed ³²P-labeled segment of the AST gene corresponding to bases 76–658 (see Fig. 2). The first wash contained 2× SSPE/0.1% SDS and was done at 60°C, followed by two more washes in 0.5× SSPE/0.1% SDS also at 60°C.

In Situ Hybridization. Brains from day 5 mated *D. punctata* females were fixed in 4% (wt/vol) paraformaldehyde and then desheathed by submersion in a solution of collagenase (Sigma, type II, 1 mg/ml) for 15 min before hybridization as described (21). The hybridization probe was the same AST gene segment described above labeled with digoxigenin (21).

RESULTS

PCR Using Degenerate Oligonucleotides. To isolate the coding sequence of an internal portion of the AST-encoding gene, oligonucleotides were designed to amplify between the ends of the largest AST peptide sequence, the octadecapeptide ASB2 (4). The degenerate sequences are diagrammed in Fig. 1 (primers 1 and 2), in relation to the amino acid sequence of the peptide. Each oligonucleotide carries a 5' extension that provides restriction recognition sequences for cloning the fragments produced.

PCR amplification was done by using as template a cDNA pool synthesized from RNA extracted from the brains of virgin female cockroaches. A product of 66 bp was isolated and cloned; a number of the resulting clones were then sequenced. A consensus sequence for the amplified region was derived from these and used to design a third oligonucleotide having only one degenerate position (primer 3 in Fig. 1), with sufficient specificity to allow 5'-end-specific PCR of the cDNA.

One-sided PCR was used to amplify the 5' end of the sequence because the cDNA target had first been ligated to a plasmid vector. The plasmid provides a known reference point beyond the cDNA terminus where opposing primers can anneal (22). Therefore, one-sided PCR was performed by using two nested primer pairs: primers 2 and 3 (Fig. 1) in conjunction with two vector-based primers T3 and SK (Stratagene), respectively. In retrospect, our consensus primer (primer 3) was found to have three mismatches with the authentic gene sequence. However, annealing of the primer

was sufficient to amplify a DNA fragment of ≈380 bp, which upon sequencing was found to encode the appropriate N terminus for the peptide. This sequence was then used to design further pairs of specific oligonucleotide primers, enabling a similar application of one-sided PCR to derive fragments containing the 3' portion of the gene.

The AST Gene Sequence. The sequence of the AST cDNA was derived by direct sequencing of PCR-amplified material on both strands. The entire sequence produced by this method is presented in Fig. 2. A variant of this sequence was also found in the cDNA (data not shown), which diverges from that shown in Fig. 2 at nt 1669. The alternative species is slightly shorter (1900 bp) and differs only in the extreme 3' region, beyond nt 1669.

Analysis of the cDNA sequence reveals a single large open reading frame beginning at residue 96 and ending at residue 1206 (Fig. 2), which encodes an apparent precursor polypeptide of 370 amino acids. This pre-pro-AST is predicted to be 41.5 kDa in size, beginning with a methionine codon at nt position 96, which is positioned in a suboptimal translational context (23). Several potential sites for cleavage of a signal sequence are found clustered between amino acids 24 and 28 (24). The remainder of the open reading frame predicts a complex set of at least 13 AST-type peptides, each of which is bounded by potential dibasic endoproteolytic cleavage sites and ends with a glycine, which is the signal for C-terminal amidation by peptidyl-glycine α-amidating monooxygenase (PAM) (25) (Fig. 3). Seven of these potential peptides have previously been purified, providing evidence that these cleavage sites are used *in vivo*. The peptide sequences are displayed in Table 1 in the order in which they appear in the gene sequence and have been numbered accordingly. We believe this to be the most practical convention for future AST nomenclature.

All but one of the peptides includes the C-terminal amino acid sequence Phe-Gly-Leu, whereas 10 of the group are further alike in encoding tyrosine in the fifth position from the terminus (Table 1). The last of these peptides, AST-13, is unique in having isoleucine as the C-terminal residue. The N-terminal regions of the predicted peptides are rather more variable, suggesting the likelihood that address sequences for individual targeting, function, or degradation rate may be located here.

An unusual feature of the organization of the AST peptides in the precursor is their tendency to be clustered in groups, separated by spacer regions of extremely high acidity in the protein sequence (Fig. 3). Although such acidic regions have never been documented for an insect neuropeptide gene, a precedent has been found in the *Aplysia* FMRFamide gene (26). These regions have been postulated to serve a neutralization function, counterbalancing the high frequency of basic residues found in the neuropeptides and their accom-

Amino Acid	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
	Ala	Tyr	Ser	Tyr	Val	Ser	Glu	Tyr	Lys	Arg	Leu	Pro	Val	Tyr	Asn	Phe	Gly	Leu		
DNA	GCN	TAT	TCN	TAT	GTN	TCN	GAA	TAT	AAA	CGN	CTN	CCN	GTN	TAT	AAT	TTT	GGN	CTN		
			C	AG	C		AG	G	C	G	A	T		C	C	C		T		
Primer 1	5'	GGAATTC	GCN	TAT	TCN	TAT	GT	3'												
			C	AG	C															
Primer 2											3'	GGN	CAN	ATA	TTA	AAA	CCN	GA	CTCGAG	5'
														G	G	G		A		
Primer 3				3'	ATA	CAC	AGC	CTC	ATG	TTC	GCTCGAG	5'								
					G															

FIG. 1. Design of oligonucleotide primers for PCR of internal and 5' segments of the AST-encoding gene. The amino acid sequence for the largest peptide (ASB2 from ref. 4) is shown along with the reverse-translated DNA sequence. Primers 1 and 2 are degenerate oligonucleotides used to amplify an internal segment of the sequence. Primer 3 represents the resulting consensus sequence obtained and was used as a nested primer with primer 1 for one-sided PCR. N represents A, C, G, or T.

5'	CCGCTCACCTAAACGTTTCATTACTTCGTCGCTGACGCTGAGACACGACGCCTCACAAAGTGGATTATTGCCCCGCTGCTCCAATACAGAAACCA	95
ATG TCA GGC CCG AGG ACG TGC TTC TGT CTA CCG TCG GCT CTT GTA CTA GTA CTG CTG AGT CTC AGC ACT TCG GCA	170	
MET Ser Gly Pro Arg Thr Cys Phe Cys Leu Pro Ser Ala Leu Val Leu Val Leu Leu Ser Leu Ser Thr Ser Ala	25	
CTA GGG ACA GCC CCT GAG CCT TCT GGA GTC CAT GAA GAA TCT CCT GCT GGA GGT GGA ACA GAT CTT CTT CCG CAT	245	
Leu Gly Thr Ala Pro Glu Pro Ser Gly Val His Glu Glu Ser Pro Ala Gly Gly Gly Thr Asp Leu Leu Pro His	50	
CCA GAA GAT CTA TCA GCA TCA GAC AAT CCC GAC CTA GAG TTC GTG AAG CGA CTT TAC GAC TTC GGG CTC GGA AAA	320	
Pro Glu Asp Leu Ser Ala Ser Asp Asn Pro Asp Leu Glu Phe Val <u>Lys Arg</u> Leu Tyr Asp Phe Gly Leu <u>Gly Lys</u>	75	
CGC GCC TAC AGT TAT GTT TCA GAG TAC AAA CGC TTA CCA GTC TAC AAC TTC GGT CTC GGC AAG AGA AGC AAA ATG	395	
<u>Arg</u> Ala Tyr Ser Tyr Val Ser Glu Tyr Lys Arg Leu Pro Val Tyr Asn Phe Gly Leu <u>Gly Lys Arg</u> Ser Lys Met	100	
TAC GGA TTT GGT CTT GGA AAG AGA GAC GGC AGA ATG TAT TCC TTT GGC TTG GGC AAA CGC GAC TAC GAT TAC TAT	470	
Tyr Gly Phe Gly Leu <u>Gly Lys Arg</u> Asp Gly Arg Met Tyr Ser Phe Gly Leu <u>Gly Lys Arg</u> Asp Tyr Asp Tyr Tyr	125	
GGC GAA GAA GAT GAA GAT GAT CAG CAA GCA ATT GGC GAT GAA GAT ATC GAA GAA TCT GAT GTG GGA GAC CTT ATG	545	
Gly Glu Glu Asp Glu Asp Asp Gln Ala Ile Gly Asp Glu Asp Ile Glu Ser Asp Val Gly Asp Leu Met	150	
GAC AAG CGC GAT AGA TTG TAT TCC TTT GGC TTG GGC AAA AGA GCA AGA CCT TAC AGC TTT GGA TTG GGC AAA CGA	620	
Asp <u>Lys Arg</u> Asp Arg Leu Tyr Ser Phe Gly Leu <u>Gly Lys Arg</u> Ala Arg Pro Tyr Ser Phe Gly Leu <u>Gly Lys Arg</u>	175	
GCA CCA TCT GGA GCA CAG CGT CTA TAT GGA TTT GGA CTT GGC AAG CGA GGA GGT TCC CTT TAC AGT TTT GGT CTT	695	
Ala Pro Ser Gly Ala Gln Arg Leu Tyr Gly Phe Gly Leu <u>Gly Lys Arg</u> Gly Gly Ser Leu Tyr Ser Phe Gly Leu	200	
GGA AAA CGA GGC GAC GGC CGG CTA TAT GCT TTT GGT CTG GGT AAA AGA CCA GTC AAC TCT GGA AGG TCC TCT GGA	770	
<u>Gly Lys Arg</u> Gly Asp Gly Arg Leu Tyr Ala Phe Gly Leu <u>Gly Lys Arg</u> Pro Val Asn Ser Gly Arg Ser Ser Gly	225	
AGC AGA TTC AAC TTC GGT CTT GGA AAG AGA TCA GAT GAC ATT GAC TTC AGA GAA CTT GAA GAG AAA TTT GCA GAA	845	
Ser Arg Phe Asn Phe Gly Leu <u>Gly Lys Arg</u> Ser Asp Asp Ile Asp Phe Arg Glu Leu Glu Glu Lys Phe Ala Glu	250	
GAC AAG AGA TAT CCT CAA GAA CAC AGG TTT TCT TTC GGT CTT GGA AAG CGT GAA GTG GAA CCC AGT GAA CTT GAA	920	
Asp <u>Lys Arg</u> Tyr Pro Gln Glu His Arg Phe Ser Phe Gly Leu <u>Gly Lys Arg</u> Glu Val Glu Pro Ser Glu Leu Glu	275	
GCT GTA AGG AAT GAA GAA AAA GAC AAC TCG TCG GTC CAT GAT AAG AAA AAC AAC ACA AAT GAT ATG CAC TCT GGC	995	
Ala Val Arg Asn Glu Glu Lys Asp Asn Ser Ser Val His Asp <u>Lys Lys</u> Asn Asn Thr Asn Asp Met His Ser Gly	300	
GAA CGA ATC AAG AGG AGT CTG CAT TAT CCT TTT GGA ATC AGG AAG CTA GAA TCT TCC TAC GAT CTC AAT TCT GCG	1070	
Glu Arg Ile <u>Lys Arg</u> Ser Leu His Tyr Pro Phe Gly Ile <u>Arg Lys</u> Leu Glu Ser Ser Tyr Asp Leu Asn Ser Ala	325	
TCA TCG CTC AAT AGC GAA GAA AAT GAC GAC ATT ACC CCT GAG GAA TTT TCC AGG ATG GTA CGT CGA CCT TTC AAC	1145	
Ser Ser Leu Asn Ser Glu Glu Asn Asp Asp Ile Thr Pro Glu Glu Phe Ser Arg Met Val <u>Arg Arg</u> Pro Phe Asn	350	
TTT GGT CTT GGT AAA CGA ATT CCT ATG TAT GAC TTC GGT ATA GGA AAG AGA TCA GAG CGT TAA AATTTGGTCTATTG	1223	
Phe Gly Leu <u>Gly Lys Arg</u> Ile Pro Met Tyr Asp Phe Gly Ile <u>Gly Lys Arg</u> Ser Glu Arg ---		
TCTTAATTTTCAACACCACTTAACCTTAATCACTGGGCAAAGAAATAATCTCTCTCCATGCGCTATTGAAATAAAAAACAAACCTTAAATTCAGAAATCT	1322	
AAATGTACAGAATACAAACTTGGGAATAATTTATGCCAACCGATGGTATTAACTTAACTCTGAAATCATAGACCATAAACTGGATTGTTTCCTTTCA	1421	
TTACTATAGATCGTGGTTTCATTAATTTGTTATACAATACATCCTCGTGTGTTGTCATAGAACCAATAATCTCTGTTACATTTCTCTATGTGAATATGT	1520	
ACTGCTGGATTTTATTAAGCGATTCAAAAGTATCCTGCAGAACAGCGAGAAAGTGCATGGCATATGTAAATTAATTTAATATAGGCCTAAAAATCA	1619	
TATATTTTAAAAATAACATATTACTATAAATAAATATATTTGAAATCTTCTCCAGGTCACCATGGAACCTCCGGCCGGAAAGGCTTCGGGAAT	1718	
ACACACATTTATGTTGAGTATGTAAGTAATCACTGAAGAAAAAATATATAGCTTTTAAACAAAGAAATACTTAACAAAGAAATACGGGAAAT	1817	
CATTATAATTAAGGAGAAATTAATGTAACAACTAATCTAAAATTCATATTTACATTGGAAAAATATACATAACATCGCTTAAACAGAAATAGCCAACGT	1916	
TTCCGAAACACGCTAGGCAAGAAATTAATAAAAAA 3'	1950	

FIG. 2. Nucleotide sequence of AST cDNA and deduced amino acid sequence of the peptide precursor. The numbering for each sequence is shown at right. The amino acid sequences for the 13 proposed AST peptides are shown in boldface type. Dibasic amino acid pairs are shaded, and glycine residues required for amidation are underlined. Potential polyadenylation signals are marked by thick bars. The most likely site for cleavage of a signal sequence is indicated by stars.

panying processing sites. In the AST precursor, the total numbers of acidic and basic amino acid residues are equal.

The third and largest acidic spacer region displays additional interesting features. Within this region three potential dibasic cleavage sites are found, processing of which could produce two additional peptides of 12 and 8 amino acids. However, neither of these sequences includes the necessary

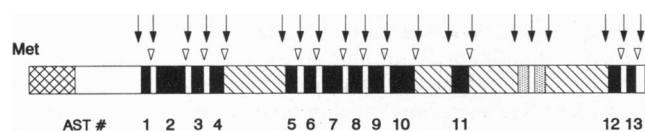


FIG. 3. Diagram of the AST polypeptide precursor from *D. punctata*. Numbering of AST peptides is shown below; the peptides themselves are represented as filled rectangles. Acidic spacer regions are shown filled by diagonal lines. The potential nonamidated peptides within the third spacer are lightly shaded. Black arrows designate sites of dibasic amino acid pairs for likely endoproteolytic cleavage. [The dibasic pair of amino acids in the AST-2 peptide may not be cleaved (4)]. Open triangles mark the presence of glycine as a potential substrate for amidation. The presence of a probable leader sequence after the initiating methionine is designated as a cross-hatched rectangle.

glycine as an amidation signal, but both do contain isoleucine as their C-terminal amino acid, in common with AST-13. Searches of the GenBank and SwissProt data bases for homologous sequences using these two peptides produced no significant matches.

The predicted AST precursor terminates at a TAA stop codon at position 1206 (Fig. 2). The untranslated nucleotide sequence that follows is extremely A + T rich and extends for >700 bp. Potential signals for polyadenylation are located near the 3' end of this region.

Genomic DNA Analysis. Copy number of the AST-encoding gene was examined by Southern analysis of *D. punctata* genomic DNA. Fig. 4 shows hybridization of a fragment of cDNA spanning residues 76–658 of the AST gene to restriction digests of total DNA. Digestion with the 4-base recognition sequence enzymes *Mse* I and *Hpa* II showed hybridization with smaller fragments consistent with the sizes predicted from the cDNA sequence (lanes M and H). An *Xho* I digest (lane X) showed hybrids at 14.2 kb and 11.4 kb. A combination digest with *Eco*RI and *Xho* I together reduced the sizes of these fragments by ≈ 2.6 kb each (lane E/X). A recognition site for *Eco*RI is found in the cDNA sequence at residue 1163 (Fig. 2). The combination of *Eco*RI and *Sal* I

Table 1. *D. punctata* AST peptides deduced from cDNA

AST	Previous designation*	Primary structure	Isolation
1		L Y D F G L -NH ₂	
2	ASB2	A Y S Y V S E Y K R L P V Y N F G L -NH ₂	†
3		S K M Y G F G L -NH ₂	
4	VII	D G R M Y S F G L -NH ₂	†
5	IV	D R L Y S F G L -NH ₂	†
6		A R P Y S F G L -NH ₂	
7	I	A P S G A Q R L Y G F G L -NH ₂	†
8	III	G G S L Y S F G L -NH ₂	†
9	II	G D G R L Y A F G L -NH ₂	†
10		P V N S G R S S G S R F N F G L -NH ₂	
11	VI	Y P Q E H R F S F G L -NH ₂	†
12		P F N F G L -NH ₂	
13		I P M Y D F G I -NH ₂	

*Previous designations are from refs. 3–5.

†AST peptides that have been isolated.

produced a single hybrid of ≈ 1.9 kb. These results suggest that the gene encoding AST is present in a single copy in the haploid *D. punctata* genome and, further, that between the upstream *Sal* I and *Xho* I sites is an allelic variation in the population such that one allele contains an extra 1.9 kb.

AST Gene Expression in *Diploptera* Brain. *In situ* hybridization analysis using brains from day 5 mated female *D. punctata* showed that the AST mRNA is strongly expressed by four large medial cells in the pars intercerebralis of the protocerebrum (Fig. 5).

DISCUSSION

We have determined the sequence of a cDNA encoding pre-pro-AST in the cockroach *D. punctata*. Because our initial attempts to amplify the AST gene from brain cDNA libraries using PCR were inconclusive, we chose, instead, to make use of double-stranded cDNA as the source material. As a result, sequence data were obtained directly from PCR-amplified DNA. To eliminate the risk of *Taq*-generated

sequence errors, data were collected from both strands of the DNA, using independently amplified templates.

The occurrence of two products in PCR amplifications of the 3' end of the AST cDNA suggests that the AST mRNA might undergo alternative splicing. The affected region is contained entirely within the 3' untranslated sequence downstream of nt 1669 (Fig. 2). Both of the possible 3' combinations are extremely A + T rich. Highly A + U-rich sequences in the 3' regions of some mRNAs have been shown to affect message stability (27). It is, therefore, tempting to speculate that important information regulating relative mRNA degradation rates resides in these different 3' arrangements of the AST mRNA.

The deduced sequence of the AST precursor contains all the appropriate signals for an exported protein capable of being processed into 13 amidated AST-type peptides. The peptide sequences confirm the structures of seven previously isolated ASTs and predict at least six new candidates. These potential ASTs are largely consistent with the basic structural-core motif for AST activity of Tyr-Xaa-Phe-Gly-Leu-NH₂, with two notable exceptions. (i) In ASTs 10–12, phenylalanine is substituted for tyrosine in the fifth position from the C terminus; the consequences of this alteration on activity remain to be studied. (ii) The final peptide (AST-13) contains isoleucine in place of leucine at the C terminus. This change may, in fact, have a relatively limited impact on the effectiveness of the peptide in inhibiting JH production. An analogous peptide variant of AST-2 (with isoleucine at the C terminus) was reported by Pratt *et al.* (4) to inhibit JH production in CA from 10-day-old female *D. punctata* to a similar degree as the authentic product ending with leucine.

Curiously, both potential peptides located within the third acidic spacer region also have C-terminal isoleucine residues. However, neither of these peptides would be expected to be amidated. Previous studies using synthetic ASTs modeled on isolated peptides have found amidation an absolute requirement for significant JH inhibition (ref. 7 and S.S.T., unpublished work). Making use of anti-AST antibodies, ASTs have been identified in nerves serving other insect tissues, particularly those projecting to the pulsatile organ muscle in *D. punctata* (10). This result suggests that other functions for this family of bioactive peptides will most probably be found. These two nonamidated AST candidates should prove useful in the search for such functions.

In *D. punctata*, retrograde filling of axons from the CA with nickel chloride (28) demonstrated that CA neurons are innervated by medial cells in the pars intercerebralis and lateral cells in the pars lateralis. Medial cells in the brain extend to the CA through the nervi corporis cardiaci I and lateral cells extend to the CA through the paired nervi

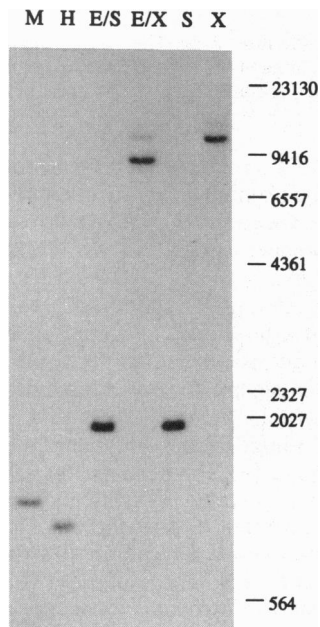


FIG. 4. Southern hybridization analysis of total cellular DNA (10 μ g per lane) using a ³²P-labeled fragment of the AST gene (nt 76–658 in Fig. 2). Genomic DNA samples were digested with various restriction enzymes (M, *Mse* I; H, *Hpa* II; E, *Eco*RI; S, *Sal* I; X, *Xho* I). DNA size markers (bases) are indicated at right.

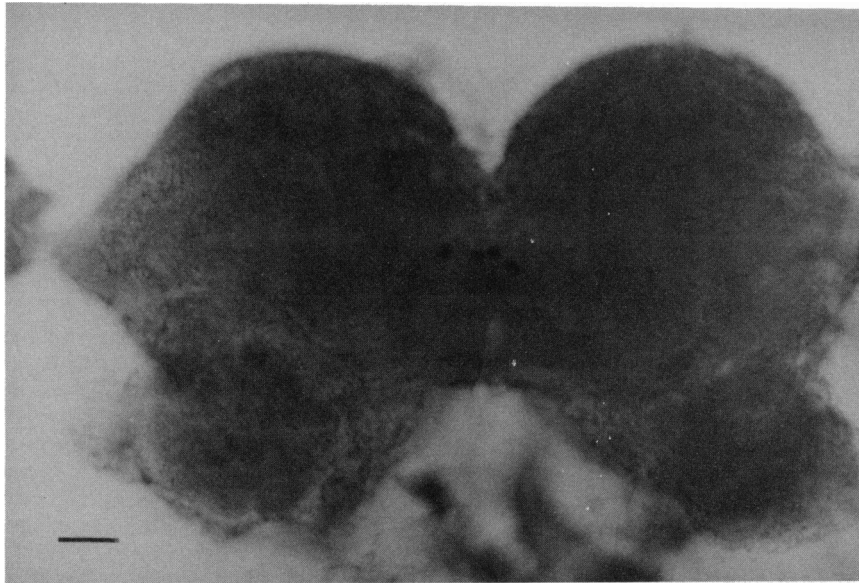


FIG. 5. Cellular localization of expression of the *D. punctata* AST gene in whole (day 5 mated) female brains (frontal view). Four strongly hybridizing medial neurosecretory cells were found in all brains examined. (Bar = 40 μ m.)

corporis cardiaci II. Destruction of either medial or lateral cell regions or the axons that extend from them (29) relieves the inhibition of JH biosynthesis imposed on the CA of virgin females, thus suggesting that ASTs are produced from both cell types. Recently, the immunoreactivity of the *D. punctata* brain was described by using monoclonal antibody raised to AST-7 (30). Lateral neurosecretory cells and associated nervi corporis cardiaci II are the major source of AST immunoreactivity extending to the CA. However, in agreement with our *in situ* hybridization observations, four medial cells are strongly immunopositive. These immunopositive medial cells terminate within the protocerebrum in areas where lateral cells form arborizations, rather than extending to the CA. Although AST mRNA expression was not evident in lateral cells of the brain by whole-mount *in situ* hybridization, preliminary studies using brain sections have indicated that cells within the lateral and tritocerebral regions also express AST mRNA at low levels compared with the four positive medial cells (unpublished work).

We thank Tian Yu for dissections and Tina Inalsingh for sequencing support. This work has been funded in part by Insect Biotech Canada, one of 15 Networks of Centres of Excellence supported by the government of Canada and by National Sciences and Engineering Research Council Grants OGP0036481 (W.G.B.) and A9407 (S.S.T.).

1. Tobe, S. S. & Stay, B. (1985) *Adv. Insect Physiol.* **18**, 305–432.
2. Tobe, S. S. (1980) in *Insect Biology in the Future: VBW 80*, eds. Locke, M. & Smith, D. S. (Academic, New York), pp. 345–367.
3. Woodhead, A. P., Stay, B., Seidel, S. L., Khan, M. A. & Tobe, S. S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5997–6001.
4. Pratt, G. E., Farnsworth, D. E., Fok, K. F., Siegel, N. R., McCormack, A. L., Shabanowitz, J., Hunt, D. F. & Feyereisen, R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2412–2416.
5. Woodhead, A. P., Khan, M. A., Stay, B. & Tobe, S. S. (1993) *Insect Biochem. Mol. Biol.*, in press.
6. Stay, B., Woodhead, A. P., Joshi, S. & Tobe, S. S. (1991) in *Insect Neuropeptides: Chemistry, Biology and Action*, eds. Menn, J. J., Kelly, T. J. & Masler, E. P. (Am. Chem. Soc., Washington, DC), pp. 164–176.
7. Pratt, G. E., Farnsworth, D. E., Siegel, N. R., Fok, K. F. & Feyereisen, R. (1991) in *Insect Neuropeptides: Chemistry, Biology and Action*, eds. Menn, J. J., Kelly, T. J. & Masler, E. P. (Am. Chem. Soc., Washington, DC), pp. 177–192.
8. Cusson, M., Yagi, K. J., Xue-Chen, G. & Tobe, S. S. (1992) *Mol. Cell. Endocrinol.* **89**, 121–125.
9. Cusson, M., Prestwich, G. D., Stay, B. & Tobe, S. S. (1991) *Biochem. Biophys. Res. Commun.* **181**, 736–742.
10. Woodhead, A. P., Stoltzman, C. A. & Stay, B. (1992) *Arch. Insect Biochem. Physiol.* **20**, 253–263.
11. Kawakami, A., Kataoka, H., Oka, T., Mizoguchi, A., Kimura-Kawakami, M., Adachi, T., Iwami, M., Nagasawa, H., Suzuki, A. & Ishizaki, H. (1990) *Science* **247**, 1333–1335.
12. Horodyski, F. M., Riddiford, L. M. & Truman, J. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8123–8127.
13. Digan, M. E., Roberts, D. N., Enderlin, F. E., Woodworth, A. R. & Kramer, S. J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11074–11078.
14. Bradfield, J. Y. & Keely, L. L. (1989) *J. Biol. Chem.* **264**, 12791–12793.
15. Davis, M.-T. B., Vakharia, V. N., Henry, J., Kempe, T. G. & Raina, A. K. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 142–146.
16. Masler, E. P., Kelly, T. J. & Menn, J. J. (1993) *Arch. Insect Biochem. Physiol.* **22**, 87–111.
17. Keeley, L. L. & Hayes, T. K. (1987) *Insect Biochem.* **17**, 639–651.
18. Schneider, L. E. & Taghert, P. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1993–1997.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
20. Ayme-Southgate, A., Lasko, P. F., French, C. K. & Pardue, M. L. (1989) *J. Cell Biol.* **108**, 521–531.
21. Tautz, D. & Pfeifle, C. (1989) *Chromosoma* **98**, 81–85.
22. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
23. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
24. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
25. Bradbury, A. F., Finnie, M. D. A. & Smyth, D. G. (1982) *Nature (London)* **298**, 686–688.
26. Schaefer, M., Picciotto, M. R., Kreiner, T., Kaldany, R.-R., Taussig, R. & Scheller, R. H. (1985) *Cell* **41**, 457–467.
27. Shaw, G. & Kamen, R. (1986) *Cell* **46**, 659–667.
28. Lococo, D. J. & Tobe, S. S. (1984) *Int. J. Insect Morphol. Embryol.* **13**, 65–76.
29. Ruegg, R. P., Lococo, D. J. & Tobe, S. S. (1983) *Experientia* **39**, 1329–1334.
30. Stay, B., Chan, K. K. & Woodhead, A. P. (1992) *Cell Tissue Res.* **270**, 15–23.