ELSEVIER

PEPTIDES

Peptides 20 (1999) 11–22

The molecular evolution of the allatostatin precursor in cockroaches†

XAVIER BELLÉS[a], LAURIE A. GRAHAM[b], WILLIAM G. BENDENA[b,*], QI DING[d], JOHN P. EDWARDS[c], ROBERT J. WEAVER[c], STEPHEN S. TOBE[d]

[a]*Department of Physiology and Molecular Biodiversity, Institut de Biologia Molecular de Barcelona (CID, CSIC), Jordi Girona 18, 08034 Barcelona, Spain*
[b]*Department of Biology, Queen's University, Kingston, ON, K7L 3N6, Canada*
[c]*Central Science Laboratory, MAFF, York, UK*
[d]*Department of Zoology, University of Toronto, Toronto, Ontario, M5S 3G5, Canada*

Received 10 July 1998; Accepted 30 September 1998

Abstract

Allatostatins (ASTs) of the Tyr/Phe-Xaa-Phe-Gly Leu/Ile-NH_2 family are a group of insect neuropeptides that inhibit juvenile hormone biosynthesis by the corpora allata. We have obtained genomic DNA sequences that specify the preproallatostatin precursor for the cockroaches, *Blatta orientalis*, *Blattella germanica*, *Blaberus craniifer* and *Supella longipalpa*. The sequences obtained are similar to those of *Diploptera punctata* and *Periplaneta americana* reported previously. The precursors of all these cockroach species are similar in size, and the organization of the ASTs that they contain (there are 13 or 14, depending on the species) have been conserved. With the sequences of these precursors, and using the homologous sequence in the orthopteran *Schistocerca gregaria* as an outgroup, a phylogenetic analysis using parsimony was carried out. The dendrograms obtained from these analyses, using the amino acid as well as the nucleotide sequences, are comparable with current models for cockroach phylogeny. Parsimony analysis was also used to study the genealogy of the different ASTs within the same precursor. Results suggest that the AST sequences were generated through a process of internal gene duplication which occurred before these species diverged from each other in evolutionary time. © 1999 by Elsevier Science Inc.

Cockroach; Neuropeptide; Allatostatin; Peptide evolution; Phylogeny

JUVENILE hormones (JHs) are sesquiterpenoids that play critical roles in the control of insect development and reproduction. The rate of JH biosynthesis in different insect species appears to be regulated by stimulatory and inhibitory peptides, the allatotropins and allatostatins (ASTs) respectively. These factors originate principally in cells of the brain and are transported via nerves to the corpora allata, the site of JH biosynthesis and release. Several ASTs have been purified from cockroach species including seven from *Diploptera punctata* [31,32,19], four from *Blattella germanica* [1], and two from *Periplaneta americana* [30]. The ASTs from each of these species share a core COOH-terminal sequence Tyr/Phe-Xaa-Phe-Gly Leu/Ile-NH_2. Further studies have demonstrated that these peptides have functions beyond their known role as inhibitors of JH biosynthesis. In *D. punctata*, ASTs are potent inhibitors of muscle contraction [25] and they are synthesized in, and potentially secreted from, a population of granulated hemocytes [24] suggesting additional functions. In *B. germanica*, ASTs impair vitellogenin release by the fat body, presumably by inhibiting the process of vitellogenin glycosylation [15].

Molecular characterization of the AST genes in *D. punctata* [5] and *P. americana* [4] has revealed a similar organization between the polypeptide precursors. The AST pep-

† Presented at the 1998 Winter Neuropeptide Conference.

* Corresponding author. Tel: 613-533-6121; fax: 613-533-6617; e-mail: bendenaw@biology.queensu.ca.

PII: S0196–9781(98)00155–7

tides of both species are not only similar in sequence but also have conserved locations within their respective precursor. Comparison of the precursors in different species should be phylogenetically informative, and also the analysis of different peptides within the same species should aid in our understanding of the process of peptidic diversification.

In order to have a more representative sample of the most important families of cockroaches, we have characterized the AST genes of *B. germanica* (Blattellidae), *Blatta orientalis* (Blattidae), *Supella longipalpa* (Pseudophyllodromiidae), and *Blaberus craniifer* (Blaberidae). These have been analyzed and compared with those previously reported of the Blattidae, *P. americana* [4], and the Blaberidae, *D. punctata* [5], and with the homologous AST-like precursor from the Orthopteran, *Schistocerca gregaria* [28], used as an external group. Although gene sequences of ASTs from the Dipterans *Calliphora vomitoria, Lucilia cuprina* [7] and *Aedes aegypti* [29] have been reported, they were not included in our analyses as Dipterans bear a distant relationship to cockroaches and locusts.

From a systematic-phylogenetic standpoint, the results obtained are comparable to the cockroach classifications currently in use, from the classic scheme proposed by McKittrick [16], to the more recent ones based on either morpho-anatomic data [10] or DNA sequence of mitochondrial ribosomal RNA genes [11]. In addition, the analysis of the different ASTs within each precursor provides clues as to how this diverse intragene family of peptides has evolved.

1. Methods

1.1. Animals

All cockroach species were maintained on Purina lab chow and water at 27°C 12 h:12 h light:dark cycle until use.

1.2. Isolation of genomic DNA

High molecular-mass DNA was isolated from each cockroach species by grinding 10 individuals under liquid nitrogen, then resuspending the ground powder in 6 mL of 10 mM Tris-HCl, pH 8.0, 60 mM NaCl, 10 mM EDTA, 0.15 mM spermidine, 0.15 M spermine and 5% (mass/vol) sucrose. Sample processing was then as described [4].

1.3. Polymerase chain reaction (PCR)

Alignment of *D. punctata* and *P. americana* AST sequences [4] revealed conserved nucleotides where primers were designed to amplify a core sequence from each heterologous genome. The upstream primer 5′AAGCGACTTTACGACTTC3′ (nucleotides 381–398 in the *D. punctata* AST sequence; [4]) was used with downstream primer 5′TCCTTACTGCTTCAAGTTCACTGG 3′ (nucleotides 980-1004 in the *D. punctata* AST sequence; [4]). Although a DNA fragment of approximately 600 nucleotides was expected, generated fragments varied in size with each genome tested. All fragments were gel purified (Sephaglas kit; Pharmacia) and sequenced on both strands with the Taq DyeDeoxy Terminator Cycle sequencing kit (Applied Biosystems). Inverse PCR [18] was used to obtain the flanking sequences 5′ and 3′ to the core AST PCR fragment generated from each genomic DNA template. Genomic DNA was digested with the restriction enzymes AatII, BamH1, BsaHI, EcoRI, ClaI, HindIII, HpaII, MseI, RsaI, ScaI, SpeI, SstI and XbaI, then purified by two phenol/chloroform (24:1) extractions, 2 chloroform extractions and ethanol precipitation. Approximately 200 ng of purified DNA from each reaction was circularized using T4 DNA ligase in a 40 μl reaction volume at 12°C overnight. PCR amplification of a 15 μl aliquot from each ligation reaction was done as above, using 15 min initial denaturation and 4 min extension times for 35 cycles. Restriction digests that resulted in fragments were purified and sequenced. The sequencing of multiple overlapping fragments was performed in each case. DNA and protein sequences have been deposited in the Genbank database (*B. germanica*, accession no. AF068061; *B. craniifer*, accession no. AF068062; *S. longipalpa*, accession no. AF068063 and *B. orientalis*, accession no. AF068064).

1.4. Sequence comparisons and analysis

The nucleotide and amino acid sequences of the AST precursor were used for comparisons. In addition to the sequences from *B. orientalis*, *S. longipalpa*, *B. germanica* and *B. craniifer* described in the present paper, we have used those of *D. punctata* [5] and *P. americana* [4]. As an external group, we have chosen the Orthopteran *S. gregaria*, whose cDNA coding for the corresponding ASTs has been reported [28].

Software from Genetics Computer Group (GCG, version 9.1), University of Wisconsin [3] was used for sequence alignments. Percentage of similarity and of identity between sequences was estimated with the application BESTFIT. Sequence alignments prior to parsimony analyses were carried out with PILEUP, and were displayed with SHADEBOX and FIG.

Parsimony analyses were carried out with the Phylogeny Inference Package (PHYLIP, version 3.57c) [8], using amino acid sequences (PROTPARS) or nucleotide sequences (DNAPARS) [26]. For the analysis of the AST sequences within the same precursor, we followed the jumble option, restarting the process 20 times. Bootstrap analyses were carried out with the application SEQBOOT in the PHYLIP package, and the procedure was repeated 100 times unless stated otherwise. The resulting trees were used to construct a consensus tree (estimated with CONSENSE).

2. Results

2.1. Isolation and sequencing of the cockroach AST gene coding region

The coding region for *B. craniifer*, *B. orientalis*, *B. germanica* and *S. longipalpa* preproAST was obtained through

Fig. 1. Nucleotide sequences of the AST precursor of *Blatta orientalis*, *Supella longipalpa*, *Blattella germanica and Blaberus craniifer*. Sequences were aligned using the program Clustal W. Sequences that are identical in all four species are indicated with an asterisk. A downward arrow indicates the position of the start codon and termination codons are shown within the sequence in non-capital letters.

the PCR of genomic DNA templates (Fig. 1). Previous comparison of *D. punctata* and *P. americana* genomic and cDNA sequences [4] revealed that the sequence was contiguous in the region encoding the preproAST precursor. This appears to be a conserved feature as alignment of genomic sequences (Fig. 1) suggests that introns are not

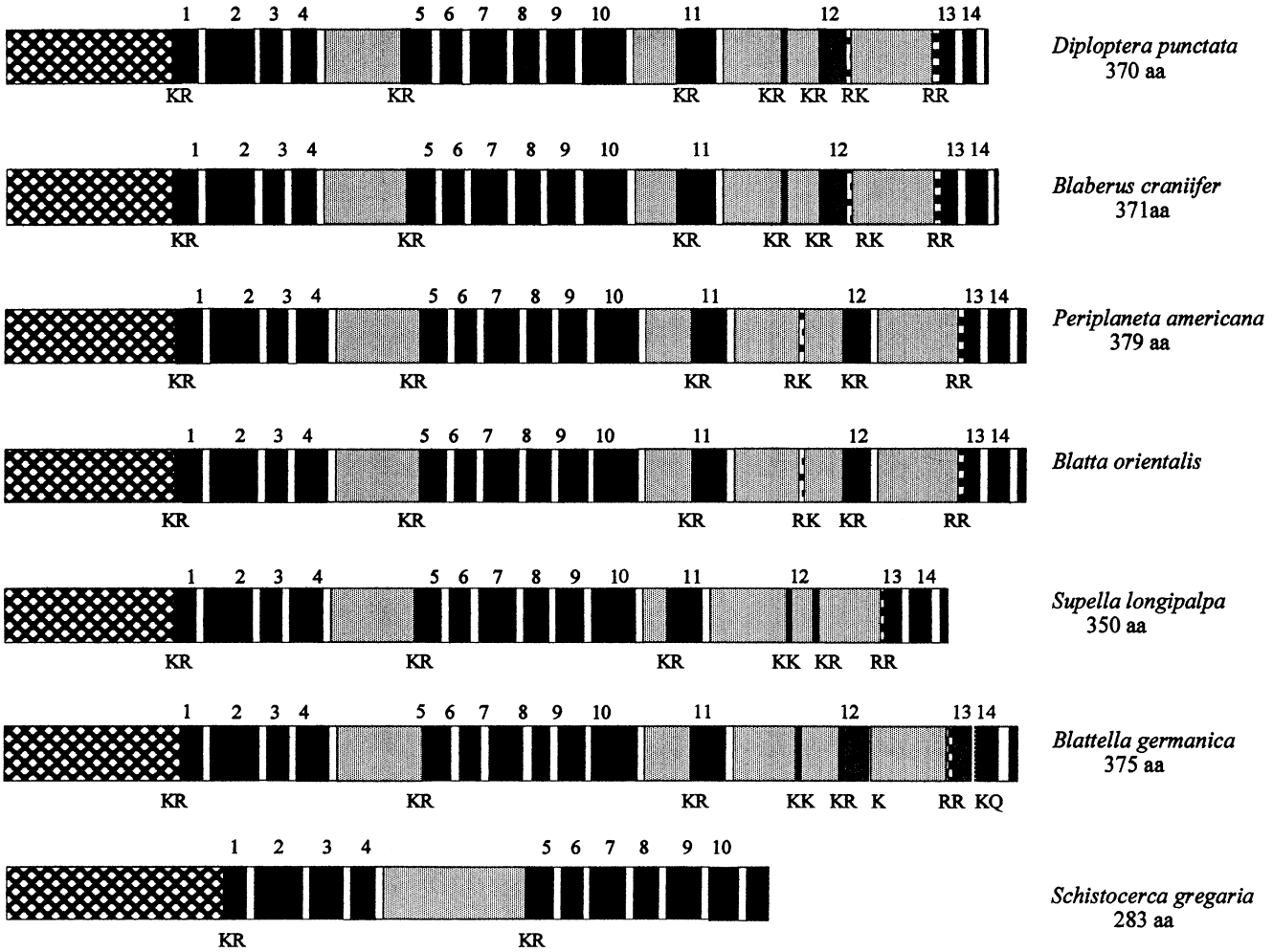

Fig. 2. Schematic organization of the cockroach AST precursor. The precursors begin with a hydrophobic leader region ▨ that is presumably cleaved by signal endoproteases. The individual ASTs that are numbered according to their position relative to the NH_2-terminus in the precursor are shown. ■ Acidic regions are indicated. ▨ Sequences (GKR) required for COOH-terminal amidation and processing are also indicated. □ The positions and sequence of mono- and dibasic endoproteolytic cleavage sites are indicated for each species.

present within the coding regions of the other pre-proASTs.

2.2. Structure and organization of the cockroach AST precursor

The AST precursors of all species are similar in size (350 to 379 amino acids). The lowest degree of amino acid identity exists in the NH_2-terminal region preceding the first dibasic cleavage site (Fig. 2 and 3). The position of the acidic spacer regions that appear to separate the AST peptides into distinct groups are also conserved between species. Although the acidic character has been maintained in these domains, the sequence identity is limited (Figs. 2 and 3). The AST peptides follow the first dibasic endoproteolytic cleavage site (Fig. 2). There are thirteen potential ASTs in *D. punctata*, *S. longipalpa*, *B. craniifer*, and *B. germanica* and fourteen in *P. americana* and *B. orientalis*. We have used a homologous system of numeration for the ASTs within the precursor of each species (Fig. 2), allowing direct comparison of amino acid sequence information at equivalent peptide positions within each precursor without requiring the assumed production of a functional peptide.

Four peptides, AST 1, 2, 3, and 6 are identical in both sequence and position for each species examined. Interspecies comparison of the remaining ASTs reveals that there have been conservative amino acid substitutions that have maintained either the hydrophobic or aromatic nature of the position. ASTs at positions 4, 7 and 8 have at least one species-specific N-terminal extension. In all species, Lys-Arg is adjacent to the NH2-terminus of AST 1, 5, 11 and 12. These putative peptides follow either the leader region or an acidic spacer region (Fig. 2). In all species, AST 13 is

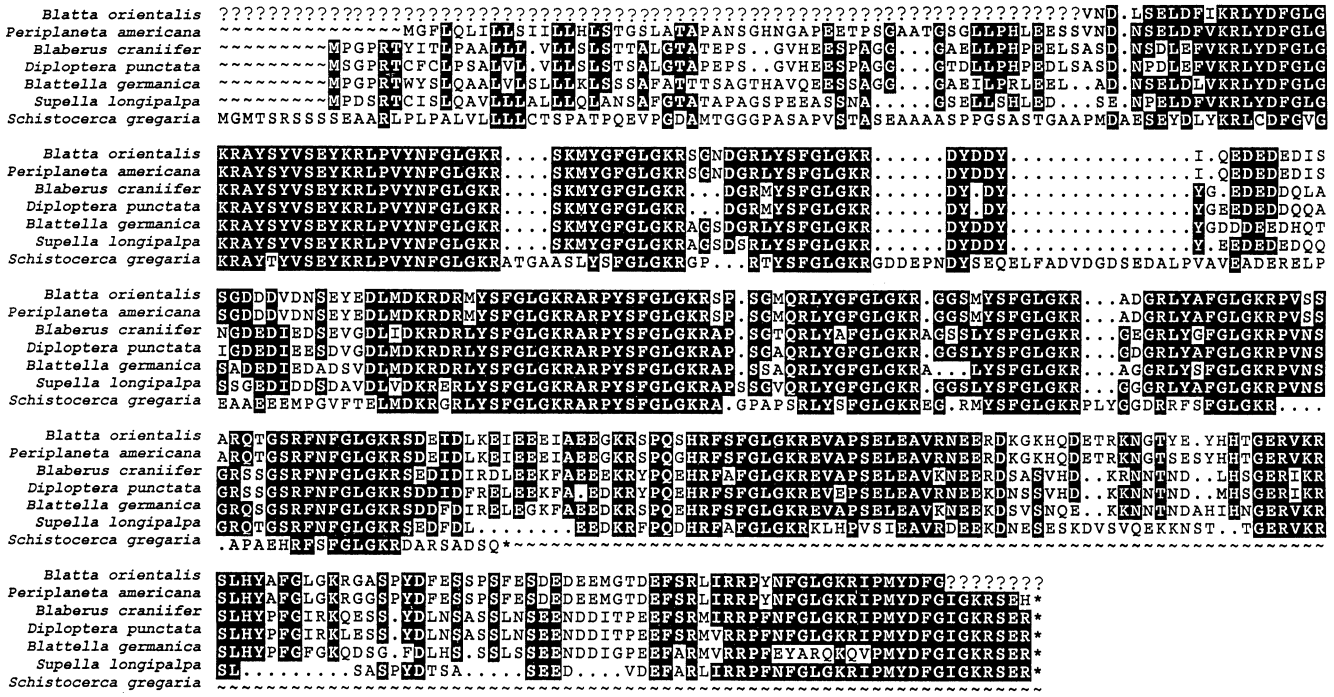

Fig. 3. Alignments of the preproAST amino acid sequences in the six cockroach species and *Schistocerca gregaria*. The PILEUP program from the GCG package was used, and results were displayed by SHADEBOX, from the same package.

preceded by a unique Arg-Arg endoproteolytic cleavage site.

ASTs require 34 α-amidation of the COOH-terminal amino acid for biologic activity [25]. Immediately following the COOH-terminal amino acid of each potentially functional AST is the sequence Gly Lys-Arg required for amidation and processing [25]. AST 12 is variant between species as the appropriate Gly Lys-Arg required for amidation and processing that is present in *P. americana* and *B. orientalis*, has been altered in *B. craniifer* and *D. punctata* to the sequence Arg-Lys. Endoproteolytic cleavage of this site would result in a non-amidated peptide (Fig. 3). AST12 is absent within the precursor of *S. longipalpa*. As well, the COOH-terminal sequence and processing signal of *B. germanica* AST12 is different in that the peptide would terminate in Phe-Gly Phe-NH2 if the single Lys or Lys-Gln is recognized as a processing substrate. Similarly, the utilization of Lys or Lys-Gln following *B. germanica* AST13 would release a seven amino acid non-amidated peptide (Figs. 3 and 5). Alternatively, in the absence of cleavage, the combination of AST13 and 14 would result in a 17 amino acid peptide. It is uncertain whether these sites are utilized in vivo.

2.3. Sequence alignments and parsimony analysis of cockroach AST precursor

Alignments were carried out with nucleotide and amino acid sequences, using the six species of cockroaches and *S. gregaria*. In all cases, we found a high degree of similarity and identity among the cockroach species and also between cockroaches and *S. gregaria*. Table 1 summarizes the percentage of identity and similarity using the amino acid sequences, and Fig. 3 shows the alignment of these sequences.

For parsimony analysis, *S. gregaria* was always used as the outgroup. Initially, we analyzed the amino acid sequences, using the alignment shown in Fig. 3, and only one most parsimonious tree requiring 1,098 steps (Fig. 4left) was obtained. The consensus tree constructed from boot-

TABLE 1

	Blo	Pea	Blg	Sul	Blc	Dip	Scg
Blo	–	98.4	73.1	76.4	73.6	74.9	66.7
Pea	98.7	–	69.0	73.5	72.2	73.3	59.4
Blg	82.5	77.5	–	72.0	78.6	78.4	71.0
Sul	81.3	78.5	77.2	–	73.4	73.8	62.2
Blc	83.4	80.9	84.4	79.0	–	90.2	66.5
Dip	84.0	81.7	83.3	78.4	93.8	–	66.5
Scg	72.5	64.3	74.6	66.2	71.2	73.5	–

Degree of identity (on the right) and similarity (on the left) between the 6 species of cockroach herein studied (*Blatta orientalis*: blo, *Periplaneta americana*: pea, *Blattella germanica*: blg, *Supella longipalpa*: sul, *Blaberus craniifer*: blc, *Diploptera punctata*: dip),and between the cockroaches and the orthopteran *Schistocera gregaria* (scg).

The BESTFIT program from the GCG package was used.

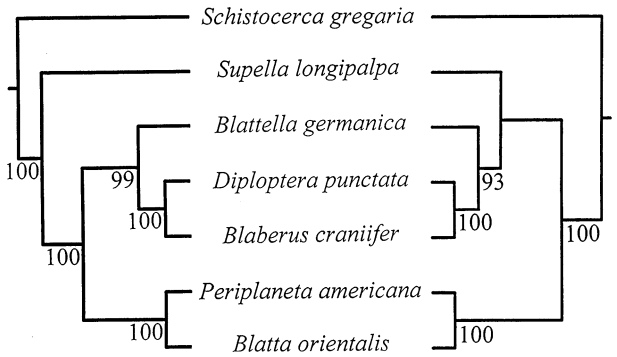

Fig. 4. Parsimony analysis of the amino acid sequence (left) and nucleotide sequence (right) of preproAST in six cockroach species, using *Schistocerca gregaria* as an outgroup. The PROTPARS and DNAPARS programs from the PHYLIP package were used respectively. The trees represented are the most parsimonious, requiring 1,098 steps (left) and 2,114 steps (right). Values from bootstrap analysis (100 replications using SEQBOOT, from the PHYLIP package) are also indicated.

strap analysis had an identical topology and the bootstrap values were high (99–100) in all nodes.

Additionally, we used the corresponding nucleotide sequences (also aligned with the PILEUP program, not shown), and we similarly obtained only one most parsimonious tree requiring 2,114 steps (Fig. 4, right). In this case the consensus tree from bootstrap analysis had a slightly different topology [interestingly, it was identical to that resulting from the amino acid sequences analysis (Fig. 4, left)], and only the bootstrap values corresponding to common nodes have been indicated.

2.4. Molecular evolution of cockroach ASTs

The organization of the AST polypeptide precursor in the six species of cockroaches (Fig. 2) is strikingly similar. Also striking is the high degree of similarity of the AST sequences corresponding to equivalent positions in the precursor (Fig. 5). When all of the individual ASTs were compared, the PROTPARS program gave 50 equally parsimonious (165 steps) trees, the consensus of which is shown in Fig. 6. All trees showed that the splitting of the sequences occurred before the species diverged since all 50 trees tend to group the equivalent sequences of different species (Fig. 6).

The data suggest that the fourteen AST sequences in the precursor were generated through a process of genetic duplication, before the species separated from each other in evolutionary time. In order to study the duplication event, a parsimony analysis of the fourteen AST or AST-like sequences was undertaken for each species. For this purpose, we again used the PROTPARS and DNAPARS programs, for analysis of the amino acid and nucleotide sequences, respectively. Since we obtained similar results, with either the amino acid or the nucleotide sequences, particularly for the more external branches of the trees, only the results from the amino acid sequence analyses will be described in more detail.

With the sequences from *B. orientalis* and *P. americana* we obtained three equally parsimonious trees requiring 97 steps, three trees of 103 steps for *B. germanica*, fifteen trees of 120 steps for *S. longipalpa*, six trees of 102 steps for *B. craniifer*, and nine trees of 101 steps for *D. punctata*. The comparison of the 39 most parsimonious trees of the six species (Fig. 7 shows a tree example for each species) indicates that the topology is similar in all cases, and identical trees can be found for more than one species (as in the pairs *B. orientalis-P. americana*, or in *B. craniifer-D. punctata*). The bootstrap analysis (Fig. 8) shows that, in general,

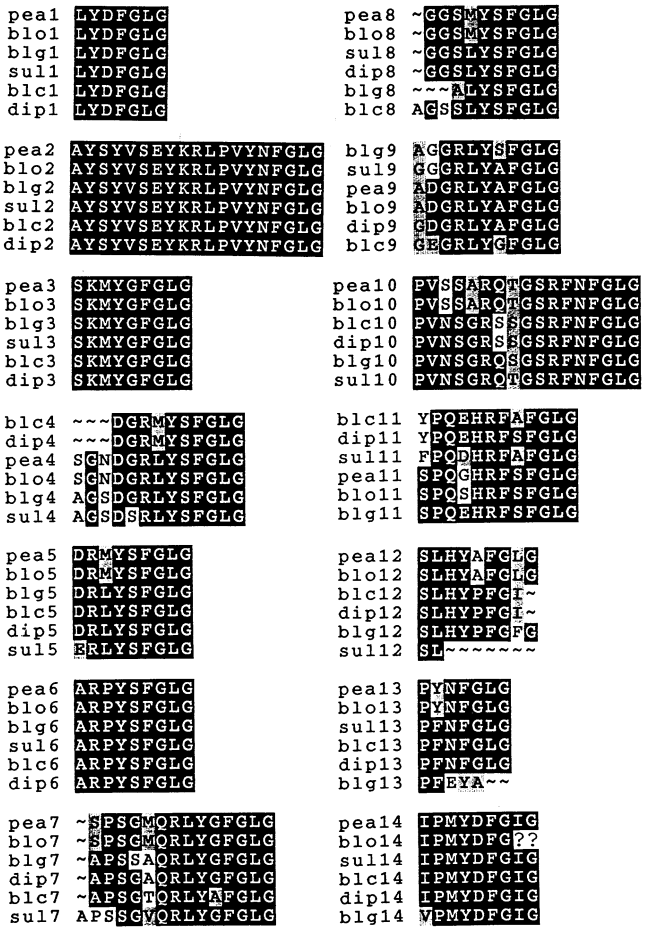

Fig. 5. Alignment of the AST-like sequences corresponding to equivalent positions in the precursor in the six cockroach species (*Blatta orientalis*: blo, *Periplaneta americana*: pea, *Blattella germanica*: blg, *Supella longipalpa*: sul, *Blaberus craniifer*: blc, *Diploptera punctata*: dip). Gly as been used instead of NH_2. In the case of sul12 and blg13 the sequence in the equivalent position was used, although it is clearly different from a typical AST sequence.

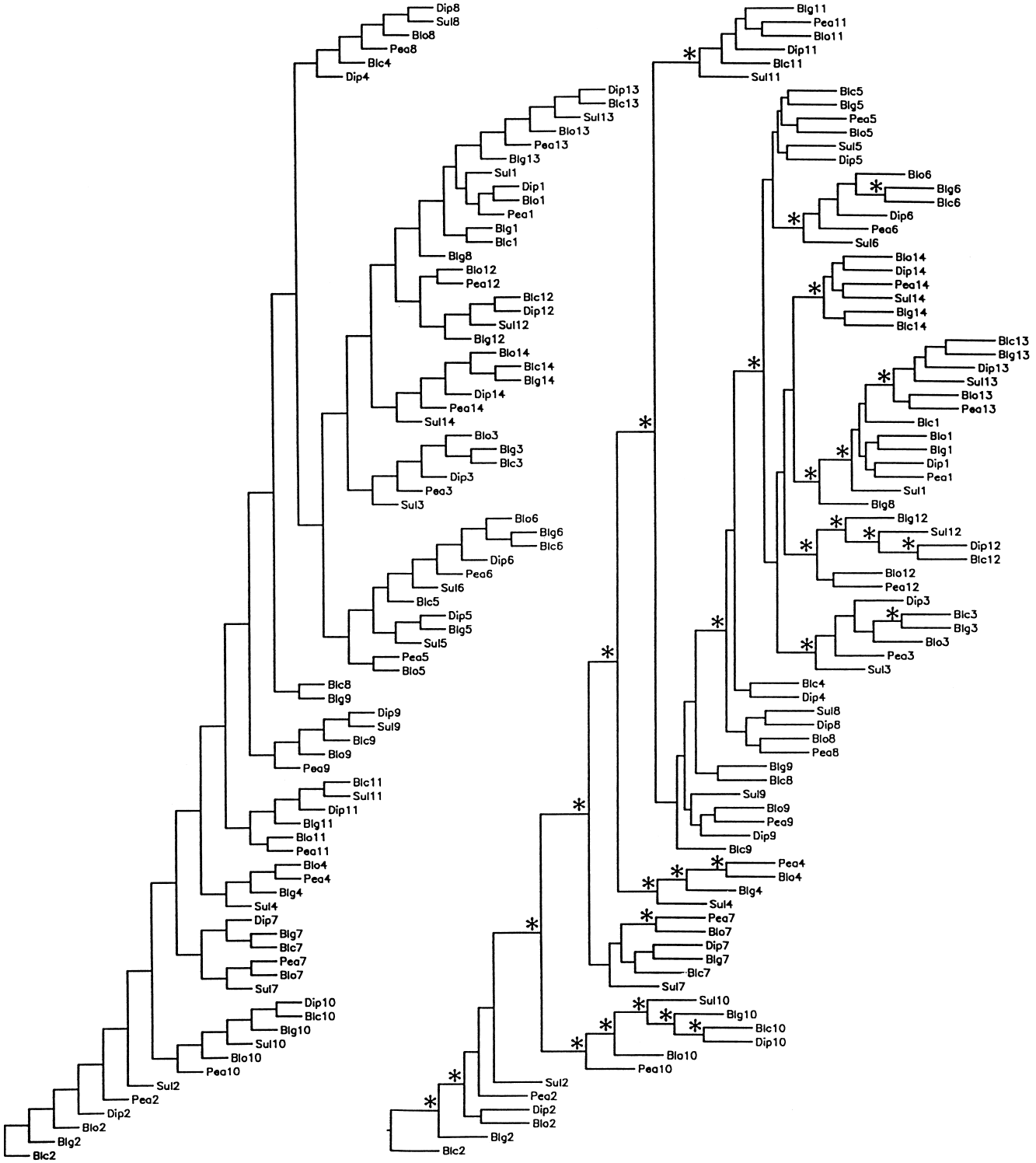

Fig. 6. Parsimony analysis of the 84 AST-like sequences of the six species of cockroaches herein studied (see sequences in Fig. 5). The PROTPARS and CONSENSE programs from the PHYLIP package were used. The tree represented on the left is a consensus of the 50 most parsimonious (requiring 165 steps) trees. That on the right is the consensus tree from the bootstrap analysis (50 replications using SEQBOOT, from PHYLIP package), where bootstrap values greater than 50% are indicated with asterisks. Abbreviations of the species are as in Fig. 5.

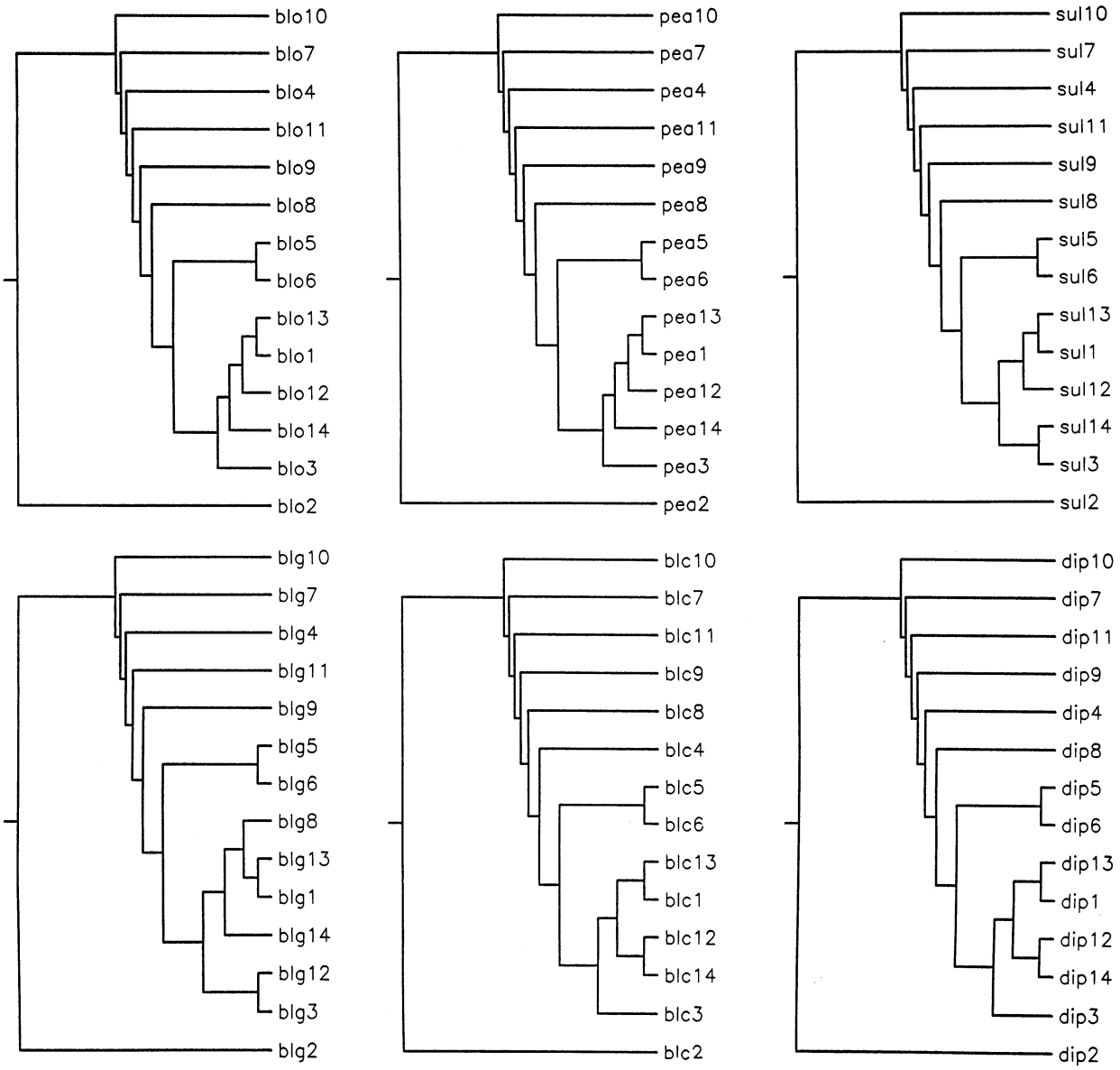

Fig. 7. Parsimony analysis of the fourteen AST-like sequences in each one of the 6 species of cockroaches herein studied (see sequences in Fig. 5). The PROTPARS program from the PHYLIP package was used. Different equally parsimonious trees were obtained for each species (*B. orientalis* and *P. americana*: 3 trees of 97 steps; *B. germanica*: 3 trees of 103 steps; *S. longipalpa*: 15 trees of 120 steps; *B. craniifer*: 6 trees of 102 steps; *D. punctata*: 9 trees of 101 steps). The trees represented have been chosen among the most parsimonious to show the similarities in topology. Abbreviations of the species are as in Fig. 5.

the more external branches are the most consistent. However, discrepancies are found in AST 8 in *B. germanica*, and AST 4 in *B. craniifer* and *D. punctata*, which occupy more internal positions in comparison to topologies corresponding to *B. orientalis*, *P. americana* and *S. longipalpa*. This suggests that the method places too much emphasis on size differences between the ASTs.

3. Discussion

The acquisition and alignment of the sequences for the preproAST in six cockroach species has revealed several common features. Firstly, the precursors are remarkably similar in size, and the organization of the peptides within the precursor is conserved. The separation of peptides into groups by acidic domains is maintained. ASTs 1, 2, 3 and 6 are identical in all species examined in both sequence and position within the precursor. The significance of this conservation with respect to biologic activity is unclear. However, AST 2 is the most effective inhibitor of in vitro JH biosynthesis in *P. americana* ($ED_{50} = 7.0 \times 10^{-10}$) and *D. punctata* ($1.4 \times 10^{-11}$) [2]. AST 1 and 3 are poor inhibitors

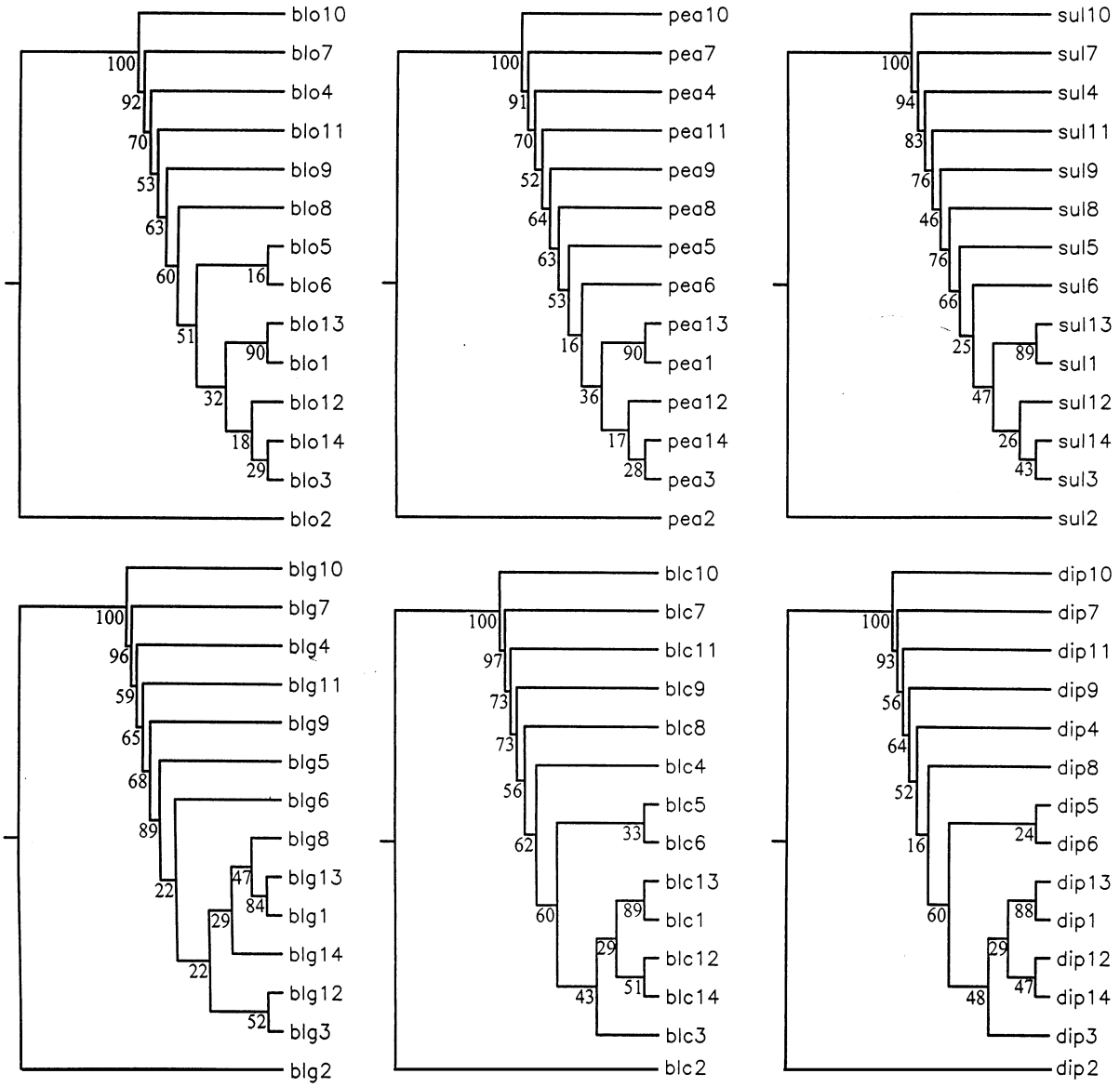

Fig. 8. Bootstrap analysis of the fourteen AST-like sequences in each one of the six species of cockroaches herein studied (see Fig. 7). The SEQBOOT, PROTPARS and CONSENSE programs from the PHYLIP package were used (100 replications). Abbreviations of the species are as in Fig. 5.

of JH biosynthesis in both species. *D. punctata* ASTs 1, 2, 3 and 6 rank poorly as inhibitors of proctolin-stimulated hindgut muscle contraction [2]. Differences between equivalent ASTs in different species occur primarily in NH_2-terminal sequence. These subtle changes may have profound effects on activity by altering receptor affinity or specificity [4]. AST 7 in *D. punctata* and *P. americana* differs by 2 amino acids near the NH_2-terminus and shows 2–3 orders of magnitude greater inhibition of JH biosynthesis in the same species than in the reciprocal species [30]. Developmental sensitivity to the two species-specific AST7 sequences was also found to vary in the two cockroaches [30]. The greatest variation appears to reside in AST 12, in which the sequences required to generate a functional AST are present only for *P. americana* and *B. orientalis*. For the remaining cockroaches, the processing signals appear to be altered such that a functional peptide may not be expressed. *S. longipalpa* is an exception, as the coding region rather than just the processing signal for AST12 appears to have been removed.

The comparative analysis of the AST precursor in the six cockroaches studied, has allowed some phylogenetic inferences. In this sense, and as antecedents, the most widely used phylogeny of cockroaches is that proposed by McKit-

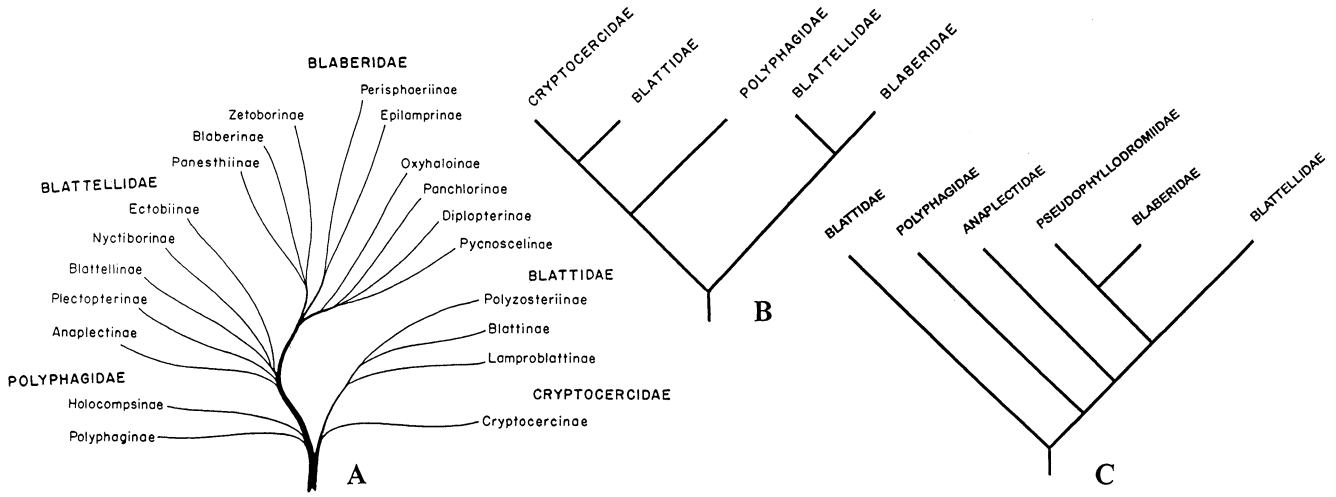

Fig. 9. Phylogenetical ordering of cockroach families based on: A. morpho-ethological characters [16]; B. the sequence of mitochondrial 12S rRNA gene [11] and C. morpho-anatomic characters [10].

trick [16], based on morpho-ethological characters. She divided the Blattaria into two superfamilies: Blaberoidea and Blattoidea, and five families: Polyphagidae, Blattellidae and Blaberidae (included in the Blaberoidea), and Blattidae and Cryptocercidae (Blattoidea). Recent phylogenetic analysis based on the sequence of the mitochondrial 12S rRNA gene [11] has provided general support for this scheme [16]. Conversely, cladistic approaches using morpho-anatomic characters [10] have divided the Blattaria into 6 families: Blattidae, Polyphagidae, Anaplectidae, Pseudophyllodromiidae, Blaberidae and Blattellidae. The later phylogeny [10] agrees in part with that of McKittrick [16], but isolates the Pseudophyllodromiidae (= Plecopterinae, considered as subfamily of Blattellidae by McKittrick) as an independent family, and considers Cryptocercidae as synonymous with Polyphagidae (Fig. 9).

Our phylogenetic analysis using the amino acid sequences of the AST precursor gave a single most parsimonious tree (Fig. 4, left), whereas the corresponding nucleotide sequences also gave a single, slightly different most parsimonious tree (Fig. 4, right). The only difference between the two trees is the position of *S. longipalpa*, which appears as a sister group of the remaining cockroaches in the amino acid-based cladogram, and as the sister group of the other three Blaberoidea in the nucleotide-based cladogram. This second tree is more congruent with the currently accepted cockroach phylogeny in that the two species of Blattoidea are clustered together, and appear as the sister group of the four Blaberoidea (Fig. 9). Although our analyses were limited to a few representatives of Blattidae (*P. americana* and *B. orientalis*), Pseudophyllodromiidae (*S. longipalpa*), Blattellidae (*B. germanica*) and Blaberidae (*B. craniifer* and *D. punctata*), the data provided by the sequences of the AST precursor seem phylogenetically informative.

Another interesting point emerging from the analysis is the segregated position of *S. longipalpa* in each of the two trees (Fig. 4), this supports the notion of an independent family proposed by Grandcolas [10] for Pseudophyllodromiidae. However, our results (nucleotide-based cladogram; Fig. 4, right) reveal that *S. longipalpa* appears as the sister-group of the Blaberidae + Blattellidae, whereas in the phylogeny proposed by Grandcolas [10], the Pseudophyllodromiidae clusters with the Blaberidae, and both appear as the sister group of Blattellidae (Fig. 9). Unfortunately, the phylogenetic analysis inferred from mitochondrial 12S rRNA gene sequences [11] did not include any species of Pseudophyllodromiidae. In any case, the topology of our nucleotide-based dendogram (Fig. 4, right) explains the evolution toward viviparity in a more parsimonious way. Accordingly, oothecal rotation preceding oothecal retraction [21] would have appeared only once in evolution, in the branch leading to Blaberidae + Blattellidae, whereas in the scheme of Grandcolas [10] it appears twice (in Blaberidae and in Blattellidae) by convergence, since the Pseudophyllodromiidae do not rotate the ootheca.

The comparison of the structure and organization of the AST DNA in the six cockroaches species, the alignments of the preproAST sequences, and the results of the parsimony analysis all suggest that the fourteen AST sequences in the precursor were generated through a process of internal gene duplication [12,13] which occurred before the species diverged. This means that the different AST sequences in the same gene are paralogous (derived from a duplication event), whereas the corresponding genes in different species are orthologous (derived from a speciation event) [9]. More-

over, knowing the precursor sequence in six species of cockroaches provides the opportunity to study the duplication event. If, as postulated, the history of duplication had been the same in all six species, then the parsimony analysis of the fourteen AST-like sequences in each of these species would result in trees with the same topology. Although our analyses did not give a single common tree for all cockroach species, many trees had similar or identical topologies for some species, which supports the hypothesis of a common process of internal gene duplication for all species.

ASTs provide a good example of a set of peptides with similar sequences which derive from a single precursor encoded by a single gene. This represents the most basic category of DNA duplication, internal gene duplication [12,13], leading to intragene families of peptides. One of the best known cases of an intragene family is that of the FMRFamide peptides, whose genetic organization have been studied in molluscs (*Aplysia*: [22,27]; *Lymnaea*: [14]), nematodes (*Caenorhabditis*: [20]) and insects (*Drosophila* [17,23]; *Lucilia and Calliphora*: [6]). However, even in the case of the thoroughly studied FMRFamides, the physiological significance of such a diversity of related peptides in a single precursor [see, for example, 22]) remains uncertain.

Acknowledgements

Financial support from DGICYT, Spain (project PB95–0062) (XB), and Natural Sciences and Engineering Research Council of Canada Grants OPG0036481 (WGB) and A9407 (SST) is gratefully acknowledged.

References

[1] Belles X, Maestro J-L, Piulachs M-D, Johnsen AH, Duve H, Thorpe A. Allatostatic neuropeptides from the cockroach *Blattella germanica* (L.) (Dictyoptera, Blattellidae). Identification, immunolocalization and activity. Regul. Pep. 1994; 53:237–247.

[2] Bendena WG, Garside CS, Yu CG, Tobe SS. Allatostatins. Diversity in structure and function of an insect neuropeptide family. In: Neuropeptides in Development and Aging. Annals of the New York Acad. Sci. 814:1997; 53–66.

[3] Devereux J, Haeberli P, Smithies O. A comprehensive set of sequences analysis programs for the VAX. Nucleic Acids Res. 1984;12:387–395.

[4] Ding Q, Donly BC, Tobe SS, Bendena WG. Comparison of the allatostatin neuropeptide precursors in the distantly related cockroaches *Periplaneta americana* and *Diploptera punctata*. Eur. J. Biochem. 1995;234:737–746.

[5] Donly BC, Ding Q, Tobe SS, Bendena WG. Molecular cloning of the gene for the allatostatin family of neuropeptides from the cockroach *Diploptera punctata*. Proc. Natl. Acad. Sci. USA 1993;90:8807–8811.

[6] Duve H, Johnsen A, East P, Thorpe A. Comparative aspects of the FMRFamides of blowflies: isolation of the peptides, genes and functions. In: Davey KG, Peter RE, Tobe SS. Perspectives in Comparative Endocrinology. National Research Council of Canada, Ottawa. 91–96; 1994.

[7] East P, Tregenza K, Duve H, Thorpe A. Identification of the dipteran Leu-callatostatin peptide family: Characterization of the prohormone gene from *Calliphora vomitoria* and *Lucilia cuprina*. Reg. Peptides 1996;67:1–9.

[8] Felsenstein J. PHYLIP (Phylogeny Inference Package) Version 3.57c. Distributed by the author. Department of Genetics, University of Washington, Seattle; 1995.

[9] Fitch WM. Distinguishing homologous from analogous proteins. Syst. Zool. 1970;19:99–113.

[10] Grandcolas P. The phylogeny of cockroach families: a cladistic appraisal of morpho-anatomical data. Can. J. Zool. 1996;74:508–527.

[11] Kambhampati S. Phylogenetic relationship among cockroach families inferred from mitochondrial 12S rRNA gene sequence. Syst. Entomol. 1996;21:89–98.

[12] Li W-H. Molecular Evolution. Sinauer Associates Inc., Sunderland, Massachusetts; 1997.

[13] Li W-H, Graur D. Fundamentals of molecular evolution. Sinauer Associates Inc., Sunderland, Massachusetts; 1991.

[14] Linacre A, Kellett E, Saunders S, Bright K, Benjamin PR, Burke JF. Cardioactive neuropeptide Phe-Met-Arg-Phe-NH2 (FMRFamide) and novel related peptides are encoded in multiple copies by a single gene in the snail *Lymnaea stagnalis*. J. Neurosci. 1990;10:412–419.

[15] Martin D, Piulachs M. D, Belles X. Inhibition of vitellogenin production by allatostatin in the German cockroach. Mol. Cell. Endocrinol. 1996;121:191–196.

[16] McKittrick FA. Evolutionary studies of cockroaches. Cornell University Agricultural Experiment Station, Memoir 1964; 389:1–197.

[17] Nambu JR, Murphy–Erdosh C, Andrews PC, Feistner GJ, Scheller RH. Isolation and characterization of a *Drosophila* neuropeptide gene. Neuron 1988;1:55–61.

[18] Ochman H, Medhora MM, Garza D, Hartl DL. Amplification of flanking sequences by inverse PCR. In: Gelfand MA, Sninsky DH, White TJ, Eds. PCR protocols: A guide to methods and applications. Academic Press, San Diego, CA. 1990;219–227.

[19] Pratt GE, Farnsworth DE, Fok KF, Siegel NR, McCormack AL, Shabinowitz J, Hunt DF, Feyereisen R. Identity of a second type of allatostatin from cockroach brains: An octopeptide amide with a tyrosine rich address sequence. Proc. Natl. Acad. Sci. USA 1991;88:2412–2416.

[20] Rosoff ML, Bürglin TR, Li C. Alternatively spliced transcripts of the flp-1 gene encode distinct FMRFamide-like peptides in *Caenorhabditis elegans*. J. Neurosci. 1992;12: 2356–2361.

[21] Roth LM. Evolution and taxonomic significance of reproduction in Blattaria. Annu. Rev. Entomol. 1970;15:75–96.

[22] Schaefer M, Picciotto MR, Kreiner T, Kaldany R-R, Taussig R, Scheller RH. *Aplysia* neurons express a gene encoding multiple FMRFamide neuropeptides. Cell 1985;41: 457–467.

[23] Schneider LF, Taghert PH. Isolation and characterization of a *Drosophila* gene that encodes multiple neuropeptides related to Phe-Met-Arg-Phe-NH_2 (FMRFamide). Proc. Natl. Acad. Sci. USA 1988;85:1993–1997.

[24] Skinner JR, Fairbairn SE, Woodhead AP, Bendena WG, Stay B. Allatostatin in hemocytes of the cockroach *Diploptera punctata*. Cell Tissue Res. 1997;290:119–128.

[25] Stay B, Tobe SS, Bendena WG. Allatostatins: identification, primary structures, functions and distribution. Adv. Insect Physiol. 1994;25:267–337.

[26] Swofford DL, Olsen GJ. Phylogeny reconstruction. In: Hillis DM, Moritz C, Eds. Molecular Systematics. Sinauer, Sunderland, Massachusetts 411–501; 1990.
[27] Taussig R, Scheller RH. The *Aplysia* FMRFamide gene encodes sequences related to mammalian brain peptides. DNA 1986;5:453–461.
[28] Vanden Broeck J, Veelaert D, Bendena WG, Tobe SS, De Loof A. Molecular cloning of the precursor cDNA for schistostatins, locust allatostatin-like peptides with myoinhibiting properties. Mol. Cell. Endocrinol. 1996;122:191–198.
[29] Veenstra, JA, Noriega FG, Graf R, Feyereisen R. Identification of three allatostatins and their cDNA from the mosquito *Aedes aegypti*. Peptides 1997;18:937–942.
[30] Weaver RJ, Freeman ZA, Pickering MG, Edwards JP. Identification of two allatostatins from the CNS of the cockroach *Periplaneta americana*: Novel members of a family of neuropeptide inhibitors of insect juvenile hormone biosynthesis. Comp. Biochem. Physiol. C. Comp. Pharmacol. 1994;107:119–127.
[31] Woodhead AP, Stay B, Seidel SL, Khan MA, Tobe SS. Primary structure of four allatostatins: Neuropeptide inhibitors of juvenile hormone biosynthesis. Proc. Natl. Acad. Sci. USA 86:5997–6001.
[32] Woodhead AP, Khan MA, Stay B, Tobe SS. Two new allatostatins from the brains of *Diploptera punctata*. Insect Biochem. Mol. Biol. 1994;24:257–263.