Evolution and function of the insect hexamerins*

THORSTEN BURMESTER

Institute of Zoology, University of Mainz, Müllerweg 6, D-55099 Mainz, Germany; e-mail: thorsten@uzomai.biologie.uni-mainz.de

Key words. Insect, hexamerin, arylphorin, storage protein, evolution, cuticle, hormone transport, immune response, hexamerin receptor

Abstract. Hexamerins are hemocyanin-related haemolymph proteins that are widespread in insects and may accumulate to extraordinarily high concentrations in larval stages. Hexamerins were originally described as storage proteins that provide amino acids and energy for non-feeding periods. However, in recent years other specific functions like cuticle formation, transport of hormones and other organic compounds, or humoral immune defense have been proposed. During evolution, hexamerins diversified according to the divergence of the insect orders. Within the orders, there is a notable structural diversification of these proteins, which probably reflects specific functions. In this paper, the different possible roles of the hexamerins are reviewed and discussed in the context of hexamerin phylogeny.

INTRODUCTION

The hexamerins of the insects belong to a growing protein superfamily that also comprises the arthropod hemocyanins and prophenoloxidases, and the hexamerin receptors discovered in the Diptera (Telfer & Kunkel, 1991; Beintema et al., 1994; Burmester & Scheller, 1996). A typical hexamerin consists of six identical or closely related subunits with molecular weights of about 80,000 Daltons each, giving rise to a native molecule of about 500,000 Daltons (Scheller et al., 1990; Telfer & Kunkel, 1991). There are at least two notable exceptions from the standard hexamerin model: in the Diptera, a dodecameric form has been reported (Markl et al., 1992), and in the higher Hymenoptera there are Glx-rich hexamerins with subunits of 105 to 110 kDa (Wheeler & Martinez, 1995; Danty et al., 1998).

Hexamerins have been discovered in all insects investigated so far. In some developmental stages, hexamerins may accumulate in the haemolymph to extraordinarily high concentrations, reaching up to 50% of the total salt-extractable proteins in larval stages (refer to Telfer & Kunkel, 1991). When the first sequence data of the hexamerins became available (Fujii et al., 1989; Willott et al., 1989), a close relationship of these proteins to the arthropod hemocyanins confirmed a notion that had been already put forward at that time on the basis of structural and immunological similarities (Telfer & Massey, 1987; Markl & Winter, 1989).

Phylogenetic analyses have shown that the hexamerins evolved from crustacean hemocyanins, while the chelicerate hemocyanins form a separate branch (Beintema et al., 1994; Burmester & Scheller, 1996; Burmester et al., 1998b). Based on molecular phylogenetic and comparative studies of development, it has been postulated that the Crustacea, and not, as commonly assumed, the Myriapoda present the closest living relatives of the insects (e.g., Turbeville et al., 1991; Averof & Akam, 1995; Friedrich & Tautz, 1995). Recent calculations indicate that the divergence of the clades which lead on one side to the hexamerins, and the hemocyanins of the decapod crustaceans on the other, occurred less than 450 to 480 million years ago. This raises the possibility that the Hexapoda share a common ancestry with the crustacean class of the Malacostraca (Burmester, unpubl.).

The question of the roles of hexamerins in living insects has intrigued and puzzled scientists since the discovery of these proteins in the late sixties (Munn & Greville, 1969; Munn et al., 1969). In view of their massive accumulation during larval life and their later disappearance during pupal and adult development, these authors suggested that hexamerins likely serve as some form of storage for amino acids and energy during subsequent non-feeding stages (Munn & Greville, 1969). This explanation became widely accepted as the standard description of hexamerin function, although several other specific roles of at least some hexamerins have been frequently proposed (see below). In previous reviews, Telfer & Kunkel (1991) and Haunerland (1996) summarised the occurrence and the assumed functions of some hexamerins and structurally or functionally related proteins. However, until recently our understanding of hexamerin evolution was very limited. Many more hexamerin sequences have become available (Table 1), and their phylogenetic relationships have been studied in detail (Burmester et al., 1998b). This has allowed the recognition of inter-order hexamerin relationships, as well as the assignment of distinct hexamerin classes that are conserved within single insect orders. In the following review, the present state of knowledge of both the evolution and the function of the hexamerins will be reviewed and integrated.

* This paper is an expanded version of a lecture presented at the 6th European Congress of Entomology held in České Budějovice, Czech Republic, August 1998.
MATERIAL AND METHODS

Protein sequences were retrieved from the relevant databases using the GenBank or EMBL World Wide Web interface. A complete list of all available complete hexamerin sequences is given in Table 1. The tools provided with the Sequence Analysis Software Package 8.0 or 9.0 from the Genetic Computer Group (GCG) were used for sequence translation and manipulation. Some of the published sequences were corrected for likely frameshift errors as described (Massey, 1995; Burmester et al., 1998b) (see Table 1). Sequence alignment was carried out by hand using a previously published multiple sequence alignment as guideline (Burmester et al., 1998b). The final alignment is available from the author upon request.

Phylogenetic inference was carried out with the help of the PHYLIP 3.5c software package (Felsenstein, 1993). The maximum parsimony method implemented in the PROTPARS program of this software package turned out to be more reliable than other phylogenetic methods and was used throughout to construct phylogenetic trees. The N- and C-terminal sequences covering the signal peptides and long C-terminal extensions were excluded as described. The statistical significance of these trees was tested by bootstrap analysis (Felsenstein, 1985) with 100 replications (program SEQBOOT) and majority-role consensus trees were obtained with the program CONSENSE.

The phylogenetic trees of the lepidopteran and dipteran hexamerins were linearised under the assumption of a molecular clock and that the divergence of the Brachycera from the culicid Nematocera took place about 210 MYA (see Burmester et al., 1998b). The final alignment is approximated under the consideration of the two closest consensus trees were obtained with the program CONSENSE. The phylogenetic program used. Biological and statistical nodes.

THE HEXAMERIN SUPERFAMILY

The arthropod phenoloxidases, arthropod hemocyanins, insect hexamerins, and dipteran hexamerin receptors share significant sequence similarities that indicate a common ancestry of these functionally very different proteins (Beintema et al., 1994; Burmester & Scheller, 1996; Burmester et al., 1998b). In these studies, the basic phylogenetic relationships in this protein superfamily have been described (Fig. 1). The general arrangement of the different members of this protein superfamily is strongly supported in statistical analysis, regardless of the phylogenetic program used. Biological and statistical arguments point to an ancient position of the phenoloxidases, which can be therefore considered as outgroup that allows the rooting of the tree (Burmester & Scheller, 1996; Sánchez et al., 1998). The hemocyanins obviously diverged from the phenoloxidases very early in arthropod evolution, probably already in the pre-Cambrian period before the different subphyla separated (Burmester, unpubl.). As mentioned above, there is a very strong statistical support for a sistergroup-position of the crustacean hemocyanins (probably including the only known insect hemocyanin) and the insect hexamerins. The hexamerin receptors that have been identified so far only in the brachyceran Diptera most likely branched off from early hexamerins (see below).

HEXAMERIN EVOLUTION

Occurrence of hexamerins and nomenclature

The existence of hexamerin-like proteins has been demonstrated by biochemical means in Orthoptera, Blattodea (Dictyoptera), Heteroptera, Coleoptera, Hymenoptera, Lepidoptera and Diptera (Munn & Greville, 1969; see Telfer & Kunkel, 1991 for a detailed review of these data) (Fig. 2). To my knowledge, no report specifically shows the absence of hexamerins in any insect species. Most hexamerins have been identified in larval stages, although there are reports on the presence and expression of some hexamerins in later development (e.g., Beneš et al., 1990; Martinez & Wheeler, 1993, 1994).

Due to the lack of an easily recognisable function of these proteins, the nomenclature of these proteins has been very chaotic. Initially, the mass-occurrence of these proteins in some insects led to species-names like "calliphorin", "manducin", "lucilin" etc. (Munn et al., 1969; Thomson et al., 1976; Kramer et al., 1980). Also the more general term "storage proteins" or similar expressions came into use (e.g. Tojo et al. 1980; Miller & Silhacek, 1982). Other authors emphasised the specific occurrence of the proteins in the larval developmental stages, therefore, some hexamerins have been referred to as larval serum proteins (LSP: Roberts et al., 1977), larval haemolymph proteins (LHP: Beverley & Wilson, 1982), or main larval haemolymph proteins (MLSP: Mintz & Rina, 1986). Because of its particularly high content of aromatic amino acids, a hexamerin of Hyalophora cecropia (Lepidoptera) has been termed "arylphorin" (Telfer et al., 1983). This designation has been adapted for hexamerins of other insect orders with a similar amino acid composition (e.g. Scheller et al., 1990; de Kort & Koopmanschap, 1994; Jamroz et al., 1996). However, the "arylphorins" of different insect orders are not particularly related. Other related haemolymph proteins do not fit any of these classifications. Therefore, Telfer & Kunkel (1991) proposed that all related hexameric storage proteins of insects should be called "hexamerins". In the following, this general term will be used throughout, except when referring to lepidopteran arylphorins or dipteran LSP-1 and -2 proteins.

Hexamerin sequences and evolution

The complete amino acid sequences of 41 hexamerins are now available (Table 1). However, these data are strongly biased towards the Diptera (17 sequences) and Lepidoptera (15 sequences), while the other insect orders are clearly underrepresented. Similarity scores range from 21 to 100% sequence identities on the amino acid level. The identical reports represent different alleles from early hexamerins (see below).
Table 1. List of available complete hexamerin sequences. The abbreviations are the same as used in the text and figures.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Species</th>
<th>Abbreviation</th>
<th>Acc. no.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>JH-binding hexamerin</td>
<td>Locusta migratoria</td>
<td>LmiJHBP</td>
<td>U74469</td>
<td>Braun &amp; Wyatt, 1996</td>
</tr>
<tr>
<td>Hexamerin (Allergen C12)</td>
<td>Periplaneta americana</td>
<td>PamHexC12</td>
<td>L40818</td>
<td>Wu et al., 1996</td>
</tr>
<tr>
<td>Hexamerin</td>
<td>Blaberus discoidalis</td>
<td>BdiHex</td>
<td>U31328</td>
<td>Jamroz et al., 1996</td>
</tr>
<tr>
<td>Cyanoprotein α</td>
<td>Riportius clavatus</td>
<td>RclCyanA</td>
<td>D87272</td>
<td>Miura et al., 1996</td>
</tr>
<tr>
<td>Cyanoprotein β</td>
<td>Riportius clavatus</td>
<td>RclCyanB</td>
<td>D87273</td>
<td>Miura et al., 1996</td>
</tr>
<tr>
<td>Hexamerin</td>
<td>Bracon hebetor</td>
<td>BheHEX</td>
<td>I25974</td>
<td>Quistad &amp; Leisy, 1996</td>
</tr>
<tr>
<td>Diapause protein 1</td>
<td>Leptinotarsa decemlineata</td>
<td>LeDP19</td>
<td>X76080</td>
<td>de Kort &amp; Koopmanschap, 1994; X86074</td>
</tr>
<tr>
<td>LHP82</td>
<td>Galleria mellonella</td>
<td>GmeLHP82</td>
<td>L21997</td>
<td>Memmel et al., 1994</td>
</tr>
<tr>
<td>Arylporin</td>
<td>Galleria mellonella</td>
<td>GmeAryl</td>
<td>M73793</td>
<td>Memmel et al., 1992</td>
</tr>
<tr>
<td>Riboflavin-binding hexamerin</td>
<td>Hyalophora cecropia</td>
<td>HceRbH</td>
<td>AF032397</td>
<td>Massey, 1995</td>
</tr>
<tr>
<td>Moderately Met-rich hexamerin</td>
<td>Hyalophora cecropia</td>
<td>HceMtHF</td>
<td>AF032398</td>
<td>Massey, 1995</td>
</tr>
<tr>
<td>Very Met-rich hexamerin</td>
<td>Hyalophora cecropia</td>
<td>HceMtHS</td>
<td>AF032399</td>
<td>Massey, 1995</td>
</tr>
<tr>
<td>Arylporin</td>
<td>Hyalophora cecropia</td>
<td>HceAryl</td>
<td>AF032396</td>
<td>Massey, 1995</td>
</tr>
<tr>
<td>Sex-specific storage protein 1</td>
<td>Bombbyx mori</td>
<td>BmoSSP1</td>
<td>P09179</td>
<td>Sakurai et al., 1988</td>
</tr>
<tr>
<td>Sex-specific storage protein 2</td>
<td>Bombbyx mori</td>
<td>BmoSSP2</td>
<td>P20613</td>
<td>Fujii et al., 1989</td>
</tr>
<tr>
<td>Arylphorin α</td>
<td>Manduca sexta</td>
<td>MseAryla</td>
<td>P14296</td>
<td>Willott et al., 1989</td>
</tr>
<tr>
<td>Arylphorin β</td>
<td>Manduca sexta</td>
<td>MseArylb</td>
<td>P14297</td>
<td>Willott et al., 1989</td>
</tr>
<tr>
<td>Met-rich storage protein</td>
<td>Manduca sexta</td>
<td>MseMRSP</td>
<td>L7609</td>
<td>Warg et al., 1993</td>
</tr>
<tr>
<td>Storage protein-1</td>
<td>Hyphantria cunea</td>
<td>HcuSP1</td>
<td>U60988</td>
<td>Mi et al., 1998</td>
</tr>
<tr>
<td>Acidic JH-suppressible protein</td>
<td>Trichoplusia ni</td>
<td>TniAHHSP</td>
<td>P22327</td>
<td>Jones et al., 1990</td>
</tr>
<tr>
<td>Basic JH-suppressible protein α</td>
<td>Trichoplusia ni</td>
<td>TniBJHSPa</td>
<td>L03280</td>
<td>Jones et al., 1993</td>
</tr>
<tr>
<td>Basic JH-suppressible protein β</td>
<td>Trichoplusia ni</td>
<td>TniBJHSPb</td>
<td>L03281</td>
<td>Jones et al., 1993</td>
</tr>
<tr>
<td>Larval serum protein 1 α (partial)</td>
<td>Drosophila melanogaster</td>
<td>DmeLSP1a</td>
<td>X03872</td>
<td>Delaney et al., 1986</td>
</tr>
<tr>
<td>Larval serum protein 1 β (partial)</td>
<td>Drosophila melanogaster</td>
<td>DmeLSP1b</td>
<td>U63556</td>
<td>Massey et al., 1997</td>
</tr>
<tr>
<td>Larval serum protein 1 γ (partial)</td>
<td>Drosophila melanogaster</td>
<td>DmeLSP1g</td>
<td>AF016033</td>
<td>Bauer &amp; Aquadro, 1997</td>
</tr>
<tr>
<td>Larval serum protein 1 y (partial)</td>
<td>Drosophila melanogaster</td>
<td>DsiLSP1g</td>
<td>AF016034</td>
<td>Bauer &amp; Aquadro, 1997</td>
</tr>
<tr>
<td>Larval serum protein 2</td>
<td>Drosophila melanogaster</td>
<td>DmeLSP2</td>
<td>X97770</td>
<td>Mousseron-Grall et al., 1997</td>
</tr>
<tr>
<td>Larval serum protein 2</td>
<td>Musca domestica</td>
<td>MdoLSP2</td>
<td>U72651</td>
<td>de Capurro et al., 1997</td>
</tr>
<tr>
<td>LSP-1/Arylporin</td>
<td>Calliphora vicina</td>
<td>CviAryl1</td>
<td>P28513</td>
<td>Naumann &amp; Scheller, 1991</td>
</tr>
<tr>
<td>LSP-1/Arylporin (partial)</td>
<td>Calliphora vicina</td>
<td>CviAryl2-5</td>
<td>X63340-X63344</td>
<td>Fischer &amp; Scheller, 1992</td>
</tr>
<tr>
<td>LSP-1/Arylporin (partial)</td>
<td>Sarcophaga peregrina</td>
<td>SpeSP1</td>
<td>A24941</td>
<td>Matsumoto et al., 1986</td>
</tr>
<tr>
<td>Larval serum protein 2</td>
<td>Calliphora vicina</td>
<td>CviSP2</td>
<td>U89789</td>
<td>Burmester et al., 1998a</td>
</tr>
<tr>
<td>Hexamerin 1.1</td>
<td>Anopheles gambiae</td>
<td>AgaHEX1</td>
<td>U51225</td>
<td>Zakschkin et al., 1997</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles gambiae</td>
<td>AgaHEX3</td>
<td>AF020872</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles arabiensis</td>
<td>AarHEXA</td>
<td>AF020873</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles quadriannulatus</td>
<td>AquHEXA</td>
<td>AF020874</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles merus</td>
<td>AmeHEX1A</td>
<td>AF020875</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles merus</td>
<td>AmeHEX1A</td>
<td>AF020876</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles melas</td>
<td>AmeHEX1A</td>
<td>AF020877</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin A</td>
<td>Anopheles melas</td>
<td>AmeHEX1A</td>
<td>AF020878</td>
<td>A. Caccone et al., unpubl.</td>
</tr>
<tr>
<td>Hexamerin 1 γ</td>
<td>Aedes aegypti</td>
<td>AaeHEX1g</td>
<td>U86079</td>
<td>Gordazde et al., 1999</td>
</tr>
<tr>
<td>Hexamerin 2 α</td>
<td>Aedes aegypti</td>
<td>AaeHEX2a</td>
<td>U86080</td>
<td>Gordazde et al., 1999</td>
</tr>
</tbody>
</table>
identical tree was obtained (with, however, lower bootstrap supports), proving that the analysis is not biased due to alignment errors (data not shown). A maximum likelihood approach again revealed the same tree, while distance matrix methods were not sufficient to resolve the relationships among the hexamerins (data not shown), probably due to homoplasic accumulation of similar amino acids in different clades (see below).

While prophenoloxidases and hemocyanins have been highly conserved throughout evolution, the hexamerins display an extraordinary diversity that has been explained by a relaxation of structural constraints possibly due to a lack of the oxygen-binding role of these proteins (Burmester & Scheller, 1996). Phylogenetic analyses show that the hexamerins share a common ancestry with the hemocyanins of the Crustacea (Beintema et al., 1994; Burmester & Scheller, 1996). Recently, a probable hemocyanin from an insect, *Schistocerca americana* (Orthoptera) has been reported (Sánchez et al., 1998). In parsimony analysis, there is about 75% bootstrap support that this protein groups with the hemocyanins of the Crustacea (using different alignments and methods, Sánchez and colleagues found a similar value), suggesting that the hexamerins already diverged in the Crustacea before the taxon split.

The basic pattern of hexamerin evolution within the insects follows the generally accepted scheme of the phylogeny of this taxon (Orthoptera + (Blattodea + (Heteroptera + Holometabola))) with significant statistical support (Figs 1, 3). Therefore, on the super-order level, the hexamerins did not diversify according to some specific functions at the first place, but followed the evolution of taxa. This does not necessarily mean that the last common species ancestor of all these clades had only a single hexamerin type; however, it is likely that only the descendants of a single protein survived in evolution and gave rise to hexamerin diversity in the extant species.

While the general phylogeny of hexamerins has been discussed in detail in an earlier paper (Burmester et al., 1998b), these authors were not able to conclude with sufficient confidence the likely position of one particular class of lepidopteran hexamerins that serves as a carrier for riboflavin. After the inclusion of more sequences and slight improvement of the multiple sequence alignment used for cladistic analysis, the riboflavin-binding hexamerins occupy a rather ancient position within the clade that combines the lepidopteran and dipteran hexamerins (93% bootstrap support), but diverging before the Lepidoptera and Diptera split (71% bootstrap value) (Fig. 3). This means that while these proteins are present in the Lepidoptera, within the dipteran stem-line they were probably lost. The ancient position of the riboflavin-binding hexamerins relative to the other dipteran or lepidopteran hexamerins is also supported by the exon-intron pattern, which demonstrates a rather conservative gene structure of the GmeLHP82, while the other lepidopteran and dipteran hexamerins are related more closely (Burmester et al., 1998b).
Fig. 3. Most parsimonious tree of insect hexamerins. A multiple sequence alignment of the hexamerin and hemocyanin sequences was analysed by the parsimony method (see Burmester et al., 1998b and Table 1). Nine hexamerins from the genus \textit{Anopheles} (Caccone et al., unpubl.) were omitted from the analysis because of their high degree of identity to \textit{A. gambiae} HEX-1.1. Signal peptides and C-terminal extensions were excluded from the analysis as described (Burmester et al., 1998b); the final alignment covers 848 positions. The tree is rooted by using the chelicerate hemocyanin sequences, which are collapsed in this figure to a single branch, as an outgroup. The total length of the tree is 13,339 substitutions; the numbers beneath the branches are the bootstrap values of 100 replications. Abbreviations are the same as given in Table 1.

The phylogenetic tree demonstrated a notable diversity of the hexamerins in the \textit{Lepidoptera} and \textit{Diptera} taxa (Fig. 3), which will be discussed below (Figs 4, 5). This does not mean that there is less hexamerin differentiation in the other insect orders, where the present knowledge on the hexamerins is limited. For example, in the honeybee, \textit{Apis mellifera}, N-terminal sequencing shows the existence of four distinct hexamerins that are utilised in a caste-specific manner (Danty et al., 1998). However, because our knowledge on hexamerins is most complete in the \textit{Diptera} and \textit{Lepidoptera}, their diversity in these taxa will be described in the following in more detail.

\textbf{Hexamerins in the Lepidoptera}

Hexamerins have been identified in this taxon by using very different approaches, resulting in the use of specific designations. For example, in a series of papers G. Jones and co-workers described three hexamerins in \textit{Trichoplusia ni} that are specifically repressed by juvenile hormone, which were consequently named "juvenile-hormone suppressible proteins" (Jones et al., 1988, 1990, 1993). Other hexamerins have been termed according to their particular amino acid content (e.g., Pan & Telfer, 1992, 1996) or their ability to bind to riboflavin (Magee et al., 1994). However, in spite of these different names, phylogenetic analysis demonstrates that nevertheless some of these hexamerins obviously represent orthologous proteins (Figs 3, 4).

In the \textit{Lepidoptera}, there are three different hexamerin classes that have diversified up to 290 MYA (Fig. 4). As mentioned above, the riboflavin-binding hexamerins (RhHex) likely diverged before the \textit{Lepidoptera} and \textit{Diptera} split. While these proteins were obviously lost in the \textit{Diptera}, they likely fulfilled an important role in the \textit{Lepidoptera} that cannot be taken over by another hexamerin. However, this function is likely not vital for the general performance of a butterfly or moth because it is absent in some \textit{Lepidoptera} (Pan & Telfer, 1992). Different kinds of amino acids have accumulated in the highly aromatic arylphorins on one hand, and the methionine-rich hexamerins on the other. These two hexamerin classes split very soon after the separation of the \textit{Diptera} and \textit{Lepidoptera}, about 250 MYA. It would be interesting whether similar protein classes can be found in the sistergroup of the \textit{Lepidoptera}, the \textit{Trichoptera} (see Fig. 2). In the higher extant \textit{Lepidoptera}, there are two different types of methionine-rich hexamerins that can be discriminated according to their apparent molecular weights in SDS-PAGE and their amino acid composition (Tojo et al., 1980; Ryan et al., 1985; Tojo & Yoshiga, 1994). These proteins probably diverged in the Jurassic period, about 162 MYA (Fig. 4). In general, the methionine-rich hexamerins are more abundant in the female than in the male, and it has been speculated that these proteins support female reproduction and egg developing by enhancing the pool of sulphur-containing amino acids at the time of vitellogenesis (Pan & Telfer, 1996).
proteins to infer the evolution of the brachyceran Diptera. Hexamerins in the Diptera

There may be more in some Nematocera (Korochkina et al., 1997b), probably due to the recent di-
vergence of the species of that genus less than 5 MYA. The LSP-1 and LSP-2 hexamerins likely diverged in the dipteran stem-line about 240 MYA (Burmester et al., 1998a) (Figs 3, 5). The LSP-1 proteins are highly aromatic hexamerins that also possess, in contrast to the lepidopteran arylphorins, a high content of methionine (see Telfer & Kunkel, 1991). There are generally fewer aromatic amino acids in the LSP-2 hexamerins. The LSP-1 and LSP-2 hexamerins likely diverged in the dipteran stem-line about 240 MYA (Burmester et al., 1998b; Fig. 5). In brachyceran species, it seems that there is only a single LSP-2 gene per haploid genome, while there may be more in some Nematocera (Korochkina et al., 1997a, b). In both Nematocera and acalyptrate Brachycera, LSP-1 proteins are encoded by several genes that are, at least in Drosophila, scattered in the genome (Akam et al., 1978; Brock & Roberts, 1983; Chrysanthis et al., 1994; Zakharkin et al., 1997). In contrast, in the calypterate Brachycera, the Lsp-1-like genes are clustered as a multigene family at a single site in the genome (Thomson et al., 1976; Schenkel et al., 1985; Tahara et al., 1984). The amplification of these genes from a single Lsp-1 gene occurred before the divergence of the Sarcophagidae and Calliphoridae, probably accompanying the radiation of the calypterate Diptera at the beginning of the Caenozoic period some 60 MYA. The available phylogenetic methods are not sufficient to resolve the relationships among the different Anopheles Hexamerin A proteins (homologues to the Hex-1 proteins described by Korochkina et al., 1997b), probably due to the recent divergence of the species of that genus less than 5 MYA.

Use of hexamerins in molecular systematics

Beverley & Wilson (1982, 1984, 1985) used the LSP-2 proteins to infer the evolution of the brachyceran Diptera (Schizophora) by immunological methods. LSP-1 and LSP-2 diversity in the Drosophilidae was also investigated by Brock & Roberts (1983). Shimada et al. (1995)

REGULATION OF HEXAMERIN EXPRESSION

As most hexamerins are regulated in a stage- and tissue-specific manner, certain signals are required that control the expression of these proteins. While there is little knowledge of the factors that restrict hexamerin synthesis to some particular cells (but see Benes et al., 1996), the developmental control is executed at least in part by insect hormones. It is generally assumed that insect larval hormones, moulting, puparium formation and the onset of metamorphosis are controlled mainly by juvenile hormone (JH) and ecdysteroids (ecdysone). Although one cannot expect a uniform pattern of the regulation of hexamerin synthesis by these hormones, some basic mechanisms seem to be conserved even among different orders.

In Blattodea and Lepidoptera, hexamerin expression is strongly suppressed by JH (Jones et al., 1988, 1993; Memmel & Kumaran, 1988; Memmel et al., 1994; Jamroz et al., 1996). The action of ecdysone on hexamerin expression is not as clear. Several authors reported that ecdysteroids accomplish a stimulating effect on transcription of the LSP-2 gene in Drosophila melanogaster (Jowett & Postlethwait, 1981; Lepesant et al.,

Fig. 5. Phylogeny of the dipteran hexamerins. This tree was constructed with an alignment of either total or partial LSP-1 and LSP-2 sequences as described in Fig. 4. Sequences that have been corrected (see "Materials and Methods") denoted by an asterisk, the abbreviations are given in Table 1.

Hexamerins in the Diptera

Two different hexamerin types have been found in the flies (Diptera), which can be arranged within two different clades, represented by the LSP-2-like hexamerins on one hand, and the LSP-1-like proteins on the other (Mousseron-Grall et al., 1997; Massey et al., 1997; Burmester et al., 1997a, b). In both Nematocera and acalyptrate Brachycera, LSP-1 proteins are encoded by several genes that are, at least in Drosophila, scattered in the genome (Akam et al., 1978; Brock & Roberts, 1983; Chrysanthis et al., 1994; Zakharkin et al., 1997). In contrast, in the calypterate Brachycera, the Lsp-1-like genes are clustered as a multigene family at a single site in the genome (Thomson et al., 1976; Schenkel et al., 1985; Tahara et al., 1984). The amplification of these genes from a single Lsp-1 gene occurred before the divergence of the Sarcophagidae and Calliphoridae, probably accompanying the radiation of the calypterate Diptera at the beginning of the Caenozoic period some 60 MYA. The available phylogenetic methods are not sufficient to resolve the relationships among the different Anopheles Hexamerin A proteins (homologues to the Hex-1 proteins described by Korochkina et al., 1997b), probably due to the recent divergence of the species of that genus less than 5 MYA.

Use of hexamerins in molecular systematics

Beverley & Wilson (1982, 1984, 1985) used the LSP-2 proteins to infer the evolution of the brachyceran Diptera (Schizophora) by immunological methods. LSP-1 and LSP-2 diversity in the Drosophilidae was also investigated by Brock & Roberts (1983). Shimada et al. (1995) studied the relationship among selected species of the Bombycidae by using partial sequences of the arylphorin gene. In a recent paper, Burmester et al. (1998b) used the available hexamerin sequences to trace the phylogeny of the winged insects under the assumption of a molecular clock.

REGULATION OF HEXAMERIN EXPRESSION

As most hexamerins are regulated in a stage- and tissue-specific manner, certain signals are required that control the expression of these proteins. While there is little knowledge of the factors that restrict hexamerin synthesis to some particular cells (but see Benes et al., 1996), the developmental control is executed at least in part by insect hormones. It is generally assumed that insect larval hormones, moulting, puparium formation and the onset of metamorphosis are controlled mainly by juvenile hormone (JH) and ecdysteroids (ecdysone). Although one cannot expect a uniform pattern of the regulation of hexamerin synthesis by these hormones, some basic mechanisms seem to be conserved even among different orders.

In Blattodea and Lepidoptera, hexamerin expression is strongly suppressed by JH (Jones et al., 1988, 1993; Memmel & Kumaran, 1988; Memmel et al., 1994; Jamroz et al., 1996). The action of ecdysone on hexamerin expression is not as clear. Several authors reported that ecdysteroids accomplish a stimulating effect on transcription of the LSP-2 gene in Drosophila melanogaster (Jowett & Postlethwait, 1981; Lepesant et al.,

Fig. 6. Evolution of hexamerin function. Known functions of the hexamerins and hemocyanins are compiled and arranged according to the phylogeny of these proteins. See text for details.

REGULATION OF HEXAMERIN EXPRESSION

As most hexamerins are regulated in a stage- and tissue-specific manner, certain signals are required that control the expression of these proteins. While there is little knowledge of the factors that restrict hexamerin synthesis to some particular cells (but see Benes et al., 1996), the developmental control is executed at least in part by insect hormones. It is generally assumed that insect larval development, moulting, puparium formation and the onset of metamorphosis are controlled mainly by juvenile hormone (JH) and ecdysteroids (ecdysone). Although one cannot expect a uniform pattern of the regulation of hexamerin synthesis by these hormones, some basic mechanisms seem to be conserved even among different orders.

In Blattodea and Lepidoptera, hexamerin expression is strongly suppressed by JH (Jones et al., 1988, 1993; Memmel & Kumaran, 1988; Memmel et al., 1994; Jamroz et al., 1996). The action of ecdysone on hexamerin expression is not as clear. Several authors reported that ecdysteroids accomplish a stimulating effect on transcription of the LSP-2 gene in Drosophila melanogaster (Jowett & Postlethwait, 1981; Lepesant et al.,
1982; Powell et al., 1984). Moreover, the presence of ecdysone response elements in the enhancer region of this gene could be experimentally demonstrated (Antoniewski et al., 1995; Beneš et al., 1996). Other studies of Manduca sexta proposed no actual influence of ecdysone on hexamerin expression (i.e., Caglayan & Gilbert, 1987). Pau et al. (1979) postulated an inhibitory effect of high concentrations of ecdysone on LSP-1 synthesis in Calliphora vicina in vitro. Similar results were obtained in Galleria mellonella (Ray et al., 1987) and Manduca sexta (Webb & Riddiford, 1988). Schenkel et al. (1983) reported that translatable LSP-1 mRNA is present in the form of mRNPs after cessation of hexamerin biosynthesis. These authors proposed that a small rise in titre of ecdysone might be responsible for the cessation of arylphorin translation in vivo. This result has gained support by the observation that the translation, but not the transcription, of Calliphora LSP-1 may be suppressed by an intermediate ecdysone titre, but neither by a low nor by a high level of this hormone (Burmester et al., 1995).

**HEXAMERIN FUNCTION**

**Hexamerins as storage proteins**

The accumulation of energy and amino acids that support the organism during later non-feeding periods is an essential process in insects. During the development of both Holometabola and Hemimetabola, the periods of feeding and starvation are largely predictable, and may have promoted the evolution of storage mechanisms that help the insect to survive. The assumption that hexamerins mainly act as storage proteins arose because of the specific accumulation of these proteins during the larval stages, and disappearance of these proteins in later pupal or adult development (e.g., Munn & Greville, 1969; de Kort & Koopmanschap, 1987; Chrysanthis et al., 1994). Experimental evidence comes from observation that in Calliphora vicina, radioactive amino acids from labelled hexamerins are incorporated in adult tissues. Wheeler & Buck (1995) demonstrated a differential expression and use of hexamerins in the sexes of the autogenous mosquito, Aedes atropalpus. They suggested that in the females the amino acids stored in the hexamerins are transferred to vitellogenins supporting egg development. A specific storage function of hexamerins (and other proteins as well) is also the most likely explanation for caste-specific accumulation in colony founding and adult development of the ants (Martinez & Wheeler, 1993, 1994; Wheeler & Buck, 1995). Pan & Telfer (1992, 1996) have investigated the utilisation of hexamerins in the Lepidoptera (Hyalophora cecropia and Actias luna) in detail. They demonstrated that these proteins are differentially cleared from the haemolymph by the fat body, and that there is also a difference in the times of hexamerin degradation and usage, probably linked to specific needs in different stages and sexes.

Genetic analysis in Drosophila has offered an understanding of hexamerin function using specific mutants. Roberts and colleagues showed that flies that are deficient in all three Lsp-1 genes, as a Lsp-1 null mutant, are viable (Roberts et al., 1991a). It should be mentioned nonetheless that, according to electron microscopic observations, there are some LSP-1 molecules present in this null mutant (Markl et al., 1992). Surprisingly, even under physiological stress conditions no effects were detected on the viability of such mutants, which still possess Lsp-2 (Roberts et al., 1991b). However, a significant decrease of fecundity was observed, mirrored by a slightly abnormal mating performance, a reduction of the number of eggs and severe abnormalities in the egg structure. Although there is no straightforward explanation for the difference in mating behaviour, the other aberrations can be clearly attributed to the reduced amount of stored amino acids due to the lack of the most prominent larval proteins.

However, one should be aware that the degradation of some proteins and the reutilisation of their amino acids do not justify the assumption of an exclusive storage function. It may be well possible that at least some hexamerins are required in some stages for some particular other roles, but are subject to the normal turn-over processes in later development.

**Binding of hormones and other small organic compounds**

Some hexamerins bind to small organic metabolites like riboflavin (Magee et al., 1994) or biliverdin (Miura et al., 1994) with high affinity. It has been suggested that these hexamerins serve as specific transporters for these molecules, but the exact physiological role of these interactions is uncertain. In the Orthoptera, a hexamerin has been described that binds to juvenile hormone (JH) with a considerably high affinity of about Kd = 10^-9 M (Ismail & Gillott, 1995; Braun & Wyatt, 1996). The sequence of such a hexamerin is known from Locusta migratoria, and there are some indications that JH binds to the sequences of the first domain (Braun & Wyatt, 1996). Haunerland & Bowers (1986) demonstrated the binding of insecticides to the arylphorin of Heliothis zea, probably reflecting either the general non-specific affinity of this protein to small organic compounds, or some role of hexamerins in the detoxification of xenobiotics.

In vertebrate blood, the serum albumins are known to act as transport vehicles for steroid hormones (e.g. Fischer et al., 1993). Similarly, some hexamerins, notably LSP-1 of Calliphora vicina (= "calliphorin"), bind to ecdysteroids in the larval insect (Enderle et al., 1983). The affinity to ecdysteroids appears to be rather low (Kd > 3x10^-5 M), but within the range that has been reported for the steroid-serum albumin interaction. The low affinity may be counteracted by the extraordinarily high concentration of the hexamerins in the larval hemolymph, and it has been calculated that even such a low affinity is sufficient to keep up to 90% of the ecdysteroids bound to hexamerins (Enderle et al., 1983). Hexamerins may be, therefore, considered as either additional regulators or protectors of this hormone, although the exact mechanisms are still obscure.
Involvement of hexamerins in cuticle formation

The tanning of the insect cuticle involves the incorporation of proteins into a complicated matrix and the subsequent cross-linking of these proteins by various diphenols (for review see Anderson, 1979). Already in the early seventies, it was observed that in *Periplaneta americana* some of these proteins derive from the haemolymph (Fox et al., 1972). The presence of hexamerins in the larval integument of *Calliphora vicina* and *Ceratitis capitata* (Diptera) was demonstrated by immunological methods (Scheller et al., 1980; König et al., 1986; Tsakas et al., 1991; Chrysanthis et al., 1994). In the Diptera, hexamerins are not synthesised by the epidermis, therefore their presence can be attributed to specific incorporation (Schenkel & Scheller, 1986). In vitro studies show that aromatic hexamerins may be efficiently cross-linked by diphenolic tanning agents like tyrosinase (Grün & Peter, 1983). However, it should be noted that at least in *Drosophila melanogaster*, LSP-1 is not essential for the formation of the pupal cuticle, because in the Lsp-1 null mutant larvae pupate normally (Roberts et al., 1991a, b). A more detailed review of the possible role of hexamerins in cuticle formation is presented by Peter & Scheller (1991).

Possible role of hexamerins in humoral immune defense

The protection of the insect from infections and parasites is essential for the survival of the animal. A number of humoral response factors have been identified (for review see Hoffmann, 1995). In recent years, several lines of evidence imply that the arylphorins of the Lepidoptera may be specifically involved in immune protection and may act as cytotoxic effectors, which are specifically induced by bacterial infections (Phripps et al., 1994; Beresford et al., 1997).

Probable related to some function in host protection, arylphorin translation is blocked upon infection with a double-stranded RNA poly-DNA-virus, which is injected by the wasp *Campeolitis sonorensis* into *Heliothis virescens* (Lepidoptera) along with the parasite’s egg (Shelby & Webb, 1994, 1997). A similar effect has been reported in *Manduca sexta* parasitised with *Cotesia congregata* (Beckage & Kanost, 1993). In contrast, Hayakawa (1994) reported that the arylphorin level is enhanced in larvae of *Pseudaletia separata* infected by the wasp *Cotesia kariyai*. The author suggested that arylphorin suppresses hemocyte degranulation and subsequent immune reactions that lead to the encapsulation of the parasite’s egg. So far, nothing is known about the molecular mechanism of hexamerin action during a humoral immune response. However, the high content of aromatic amino acids in the arylphorins may enhance the cross-linking capabilities of this protein, and, therefore, turn it into an ideal tool for the isolation of parasites. Other reported effects may be explained by the initiation of the humoral immune reaction cascade due to the clotting of the arylphorins.

Hexamerins as braconid toxins?

A single hexamerin sequence is known from the braconid Hymenoptera (Quistad & Leisy, 1996). Most surprisingly, this hexamerin has been discovered as a component of the venom of *Bracon hebetor*, which is injected into the host (Quistad et al., 1994). However, this sequence is covered by a US patent, and a thorough analysis of the function of this protein in the venom has not been published so far.

**HEXAMERIN UPTAKE AND RECEPTORS**

At the end of larval development, fat body cells change their function from synthesis to storage. The presence of protein storage granules in larval tissues has been already observed by Bishop (1922, 1923) in *Apis mellifera*, later also in *Drosophila melanogaster* (von Gaudecker, 1963; Butterworth, 1965). The work of Locke & Collins (1965, 1966, 1967, 1968) demonstrated that these granules are formed by proteins that have been taken up from the haemolymph. These proteins were later identified as hexamerins (Tojo et al., 1981; Locke et al., 1982). Selective uptake of hexamerins by the fat body was confirmed by injection of labelled haemolymph proteins into the living animal or by in vitro incubation of isolated fat bodies (Martin et al., 1971; Miller & Silhacek, 1982; Ueno & Natori, 1982; Ueno et al., 1983; Ryan et al., 1985; Mariotti & deBianchi, 1986; Caglayan & Gilbert, 1987; Burmester & Scheller, 1992; Pan & Telfer, 1992, 1996).

The transport of hexamerins across the fat body cell membrane requires the existence of a specific receptor. Such proteins have been demonstrated so far in the Lepidoptera (Wang & Haunerland, 1994a, b; Kirankumar et al., 1997) and Diptera (Ueno et al., 1983; Ueno & Natori, 1984; Burmester & Scheller, 1992, 1995a) by biochemical means. Sequences of the receptors are only known from the dipteran clade, notably from the blowfly *Calliphora vicina* (Burmester & Scheller, 1995b), the fleshfly *Sarcophaga peregrina* (Chung et al., 1995), and the fruitfly *Drosophila melanogaster* (Maschat et al., 1990; Burmester et al., 1999). Surprisingly, a significant similarity of these proteins to the hexamerins and other members of this protein superfamily was observed (Burmester & Scheller, 1996). Although phylogenetic analysis implies that these proteins diverged early in insect evolution, probably before the radiation of the Pterygota (Burmester & Scheller, 1996; Burmester et al., 1998b), receptor sequence data are still missing outside the Diptera.

**EVOLUTION OF HEXAMERIN FUNCTION**

Hexamerins evolved from ancient crustacean hemocyanins but lost the ability to bind oxygen. The regulation of the respiratory function of hemocyanin involves a highly co-operative molecule (for review, see Markl & Decker, 1992). This requirement probably restricts the mutability of the hemocyanin subunit sequences. In contrast, the hexamerins are much more variable due to a relaxation of the selective pressure (see above). Nevertheless, the high degree of conservation within the hemocyanins does not rule out that there may be other
functions of these proteins as well. For example, it has been demonstrated that in the Decapoda, hemocyanin concentration in the hemolymph is associated with the molting cycles, suggesting a specific utilization during starvation (cf. Depledge & Bjerrregaard, 1989), similar to the storage function of some hexamerins in metamorphosis. While the amino acid compositions of the hemocyanins do not significantly vary from the average protein (Telfer & Kunkel, 1991), some hexamerins are enriched in aromatic amino acids (up to 26% Phe + Tyr) or methionine (up to 10%). Therefore, although a storage function is likely a rather ancient role of these proteins, the specific requirements of a particular insect are reflected in an accumulation of particular amino acids in some hexamerins. Phylogenetic analyses demonstrate that this occurred independently several times during insect evolution (Figs 3, 6) (Telfer & Kunkel, 1991; Burmester et al., 1998).

Other functions of hemocyanins have not been thoroughly investigated so far, therefore it is uncertain whether hormone binding, for example, is an exclusive feature of the hexamerins. However, in view of the widespread capabilities to bind to small organic molecules (Fig. 6), it is more likely that a general affinity to these compounds is a rather ancient feature. This view is supported by the observation that chelicerate hemocyanin may bind to ecdysone (Decker & Foll, 1993). A similar line of argumentation can be put forward for a function in cuticle formation, because the presence of hemocyanins has been demonstrated in this structure (Paul et al., 1994).

CONCLUSIONS AND PROSPECTS

While for a long time hexamerins have been considered mainly as storage proteins, there is growing evidence that these proteins also serve other specific functions. Of course these findings do not reject either an exclusive or an important role of these proteins in amino acid storage, but the significant differences in the pattern of the expression and utilisation of hexamerins during development cannot be explained unless there are stage-specific requirements. Moreover, the pattern of conservation of distinct hexamerin classes in the Lepidoptera and Diptera (maybe in other insect orders as well) for more than 200 million years is striking. It seems unlikely that such diversity has been maintained unless there are at least some specific functions that cannot be taken over by another hexamerin or haemolymph protein. However, the picture is far from being complete, and several other functions of these most prominent insect proteins may be discovered in the future.

ACKNOWLEDGEMENTS. This paper is an expanded version of a talk presented at the Vth European Congress of Entomology, České Budějovice, August 1998. I am grateful to the organisers of this meeting for giving me the opportunity to present my work. I thank D. Wheeler and H. Beneš for making sequence data available before publication. I am grateful to H. Beneš, K. Scheller, J.-A. Lepesant, H. Decker, J. Markl and H. Massey for many helpful discussions. H. Beneš read the entire manuscript, improved the language and made many useful suggestions. My work cited in this review received support from the European Molecular Biology Organisation and the Deutsche Forschungsgemeinschaft (Bu956/3; Schle95/7).

REFERENCES


Burmester T. & Scheller K. 1995a: Ecdysterone-mediated uptake of arylphorin by larval fat bodies of Calliphora vicina;


Jones G., Hiremath S.T., Hellmann G.M. & Rood E. 1988: Juvenile hormone regulation of mRNA levels for a highly}


Received November 3, 1998; accepted January 28, 1999