1 Very high-density lipoprotein and vitellin as carriers of novel
2 biliverdins IX α with a farnesyl side-chain presumably derived from
3 heme A in *Spodoptera littoralis*.

4
5
6 Hartmut Kayser^{a*}, Manfred Nimtz^b, Philippe Ringler^c and Shirley A. Müller^c
7
8

9 ^aInstitute of General Zoology and Endocrinology, Ulm University, Helmholtzstrasse 8/1,
10 89081 Ulm, Germany

11 ^bHelmholtz Centre for Infection Disease, Cellular Proteome Research, Inhoffenstrasse 7,
12 38124 Braunschweig, Germany

13 ^cCenter for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel,
14 Mattenstrasse 26, 4058 Basel, Switzerland
15
16
17

18 *To whom correspondence should be addressed.

19 E-mail: hartmut.kayser@uni-ulm.de

20 Tel.: +49 731 50 15110

21 Fax: + 49 731 50 32609
22
23

24 **Running title**

25 Protein complexes with farnesyl biliverdins.
26
27

28 **Abbreviations**

29 AS, ammonium sulphate; CD, circular dichroism; CV-bilin, bilin from *Cerura vinula*; EM,
30 electron microscopy; ESI-MS, electrospray ionization mass spectrometry; MALDI-TOF,
31 matrix-assisted laser desorption/ionization time-of-flight; STEM, scanning transmission EM;
32 TEM, transmission EM; VHDL, very high-density lipoprotein.

Abstract

Bilins in complex with specific proteins play key roles in many forms of life. Biliproteins have also been isolated from insects; however, structural details are rare and possible functions largely unknown. Recently, we identified a high-molecular weight biliprotein from a moth, *Cerura vinula*, as an arylphorin-type hexameric storage protein linked to a novel farnesyl biliverdin IX α ; its unusual structure suggests formation by cleavage of mitochondrial heme A. In the present study of another moth, *Spodoptera littoralis*, we isolated two different biliproteins. These proteins were identified as a very high-density lipoprotein (VHDL) and as vitellin, respectively, by mass spectrometric sequencing. Both proteins are associated with three different farnesyl biliverdins IX α : the one bilin isolated from *C. vinula* and two new structurally closely related bilins, supposed to be intermediates of heme A degradation. The different bilin composition of the two biliproteins suggests that the presumed oxidations at the farnesyl side-chain take place mainly during egg development. The egg bilins are supposedly transferred from hemolymph VHDL to vitellin in the female. Both biliproteins show strong induced circular dichroism activity compatible with a predominance of the *M*-conformation of the bilins. This conformation is opposite to that of the arylphorin-type biliprotein from *C. vinula*. Electron microscopy of the VHDL-type biliprotein from *S. littoralis* provided a preliminary view of its structure as a homodimer and confirmed the biochemically determined molecular mass of ~350 kDa. Further, images of *S. littoralis* hexamerins revealed a 2x3 construction identical to that known from the hexamerin from *C. vinula*.

Keywords

Bilin structure; Biliprotein; Insect storage protein; Mass spectrometry; Protein imaging; *Spodoptera littoralis*.

1. Introduction

Tetrapyrroles are essential cofactors in many diverse processes in all forms of life, acting as cyclic porphyrins or as open-chain bilins (Frankenberg and Lagarias, 2003). These bilins are derived from the IX α -isomer of biliverdin by modification of the chromophore and covalent binding to specific proteins. In vertebrates, bilins, commonly of the IX α -type, are produced by the degradation of heme from haemoglobin and various other heme proteins, and are finally destined for excretion. Insect bilins display more structural diversity, since either the α - or the γ -isomer of biliverdin is produced depending on the taxonomy of the species; the γ -isomer is almost exclusively present in Lepidoptera (Kayser, 1985). The known insect bilins are non-covalently associated with protein. Although a number of insect biliproteins have been purified, most from Lepidoptera, the proteins and the bilins in particular have only been firmly identified in a few cases, most of the pigments being tentatively described as 'biliverdin' or 'biliverdin-like', and their roles are largely subject to speculations (Kayser, 1985, 2005).

Recently, we isolated a high-molecular weight biliprotein from larval hemolymph of a moth, *Cerura vinula*, and identified the apoprotein as an arylphorin-type hexamerin specifically associated with a novel bilin (Kayser et al., 2009). It represents the first biliprotein among the arylphorin class of insect storage proteins. Arylphorins are produced in high quantities together with the methionine-rich storage proteins, another major class of the hexamerins, as reserves for developmental transformation (Burmester, 1999). The structure of this bilin, denominated CV-bilin, was fully elucidated using the most recent spectrometric methodology, and shown to be biliverdin IX α with a modified farnesyl side-chain and a hydroxymethyl group at one of the pyrrole rings (Kayser et al., 2014). The substituent pattern of this unique bilin suggests that it may result from α -cleavage of mitochondrial heme A (or its precursor, heme O), which is the only known tetrapyrrole carrying a farnesyl group (Brown et al., 2002). So, heme A cleavage, a conserved key step of heme degradation, is expected to give rise to a farnesyl bilin. By contrast to heme B, however, the degradation pathway of heme A is unknown to date (Hederstedt, 2012).

The presence of an α -isomer of biliverdin in *C. vinula*, a member of the Notodontid family, was unexpected because Lepidoptera are known to produce the γ -isomer, as mentioned above. This prompted us to study another species, *Spodoptera littoralis*, which is a member of a different moth family, the Noctuidae. Here, we isolated not only a bilin suggested to be identical with CV-bilin from all molecular data obtained, but also two other structurally closely related farnesyl biliverdins that are considered to be intermediates of the proposed transformation of heme A to CV-bilin. In *S. littoralis*, however, the three bilins are not associated with an arylphorin as in *C. vinula*, but with a very high-density lipoprotein (VHDL) in larval hemolymph and with vitellin in eggs. The bilins may be transferred from VHDL to vitellin, as bilin synthesis in ovaries has not yet been shown. The bilin composition changes with the *Cerura* bilin being predominant in the egg biliprotein, the two other bilins are mainly associated with the hemolymph VHDL. In addition, we studied the structure of the VHDL and of hexamerins by electron microscopy (EM). The images present a preliminary

view on an insect VHDL and confirmed the conserved architecture of the hexamerin complexes.

2. Materials and methods

2.1. Chemicals and insects

Unless stated otherwise, chemicals were obtained from Merck or Sigma-Aldrich. Larvae and eggs of the cotton leaf worm, *S. littoralis* (Noctuidae), were from an established culture maintained on an artificial diet.

2.2. Protein purification

Hemolymph was collected in cold 0.05 M Tris/HCl, pH 7.4, containing 0.05 M NaCl, and processed as in previous work (Kayser et al., 2009). Briefly, the blue biliprotein was precipitated from centrifugation supernatants by the addition of solid ammonium sulphate (AS) in the saturation range of 45% to 70% at 5 °C. Usually, narrower ranges of saturation were applied to improve purification (see 3.1.). The biliprotein was purified by gradient chromatography on DEAE-Sephacel (2.5 x 34 cm). Blue fractions were characterized by their UV-visible spectra and pooled according to purity ($A_{662\text{nm}}/A_{278\text{nm}}$). Frozen eggs, harvested 6 to 12 h after deposition, and newly hatched larvae, were homogenized first in cold 0.05 M Tris/HCl, pH 7.4, containing 0.05 M NaCl, and then in 0.05 M Tris/HCl, pH 7.4, containing 0.5 M NaCl. Egg extracts were fractionated by the addition of AS in the range of 45% to 67% saturation. Samples were purified on a short column of DEAE-Sephacel (1.4 x 13 cm or 2 x 5 cm), which was step-eluted with 0.125 M and 0.25 M NaCl in 0.05 M Tris/HCl, pH 7.4.

2.3. Size-exclusion chromatography

Size-exclusion chromatography was performed on a TSK HW 55-S column (2.5 x 34 cm) (Merck) with 0.05 M Tris/HCl, pH 7.4, at a rate of 0.8 mL/min at room temperature. The high-molecular weight kit (GE Healthcare Life Sciences) of proteins was used for calibration.

2.4. Polyacrylamide gel electrophoresis

SDS-PAGE was performed with pre-cast NuPAGE Novex Bis-Tris 10% and 4-12% gels using either Mes or Mops buffer, as described (Kayser et al., 2009).

2.5. Density gradient centrifugation

Density gradients were generated with 46% KBr in 0.05 M Tris/HCl, pH 7.4, in Opti Seal tubes spun at 100,000 rpm (543,000 g) using a Beckman Optima TLX ultracentrifuge with rotor TLA 100.4. The running time was 5 h at a set temperature of 20 °C. After termination, the tubes were punctured at the bottom and fractions were collected for gravimetric determination of density and for UV-visible spectroscopy. Alternatively, bands were collected from the top using a Pasteur pipette.

2.6. Protein identification

Proteins separated by SDS-PAGE, were identified by in-gel tryptic digestion of stained gel slices and sequencing of cleavage peptides by mass spectrometry (MS), as in previous work

(Kayser et al., 2009). The (glyco-)peptides were analysed by electrospray ionization mass spectrometry (ESI-MS/MS) (Nimtz et al., 2004). An Orbitrap XL mass spectrometer (ThermoScientific, Bremen, Germany) equipped with a nanospray ion source was used. For collision-induced dissociation experiments, parent ions were selected in the ion trap and fragmented in the HCD cell, using nitrogen as collision gas. The resulting daughter ions were then separated in the Orbitrap analyser at a resolution of ca. 60,000. These spectra of tryptic peptides were used for manual *de novo* sequencing. If necessary for sequence assignments, MS3 experiments were performed. As Leu and Ile cannot be distinguished with this technique L always means L or I. Database searches against insect proteins were performed using the fasta 3 algorithm. For matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) peptide mass fingerprinting, protein bands were cut out from SDS-gels, reduced and carboxamidomethylated, and then subjected to in-gel tryptic digestion. The resulting peptides were extracted, desalted using ZipTip devices (Millipore) and analysed by MALDI-TOF-MS with alpha-cyano-4-hydroxycinnamic acid as matrix, using a Bruker Ultraflex TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). The MASCOT search engine (<http://www.matrixscience.com>) was employed to identify proteins.

2.7. Bilin isolation and characterisation

As in previous work (Kayser et al., 2014), bilins were liberated by precipitation of the carrier protein with methanol, and isolated from the supernatant by solid-phase adsorption to a Sep-Pak C18 cartridge and elution with methanol, or by extraction into chloroform. Methylation, acetylation, micro degradation and reverse-phase HPLC of bilins were performed also as described (Kayser et al., 2014).

2.8. Spectroscopy

UV-visible spectra were routinely recorded on a Perkin Elmer Lambda 15 spectrophotometer. The molar bilin-to-protein ratio was estimated from the extinction coefficients of VHDL from *Heliothis zea* (see 3.1.) (205.5/mM at 278 nm (<http://www.expasy.org/protparam>)) and the Soret absorption of biliverdin IX α (50/mM at 368 nm (Heirwegh et al., 1991)). Circular dichroism (CD) spectra and UV-visible spectra were obtained using an Applied Photophysics Chirascan photometer. High resolution and high accuracy mass spectra of bilins were acquired on a Bruker APEXIII Fourier Transform mass spectrometer equipped with an ESI-source operated in both the positive and negative ion modes. Thirty-two spectra were accumulated and internally calibrated using the signals from the Agilent ES tune mix solution. The measured mass accuracies were between 0.6 and 1.4 ppm at a resolution of about 30,000, and only one molecular formula was returned by the system within the given constraints. High resolution ESI-MS was performed with a Thermo Scientific LTQ Orbitrap mass spectrometer with an accuracy of +/- 3 ppm.

2.9. Effect of α -cyclodextrin

Solutions of biliproteins in 0.05 M Tris/HCl, pH 7.4, containing 0.05 M NaCl, were incubated with α -cyclodextrin at 0.04 M final concentration for 30-45 min at 5 °C (Jouni et al., 2000), spun at 14,000 *g* for 10 min, and the supernatant was saved for UV-visible and CD spectroscopy. Controls without α -cyclodextrin were treated in parallel. The protein-bound

state of bilin was examined by micro-filtration using Microcon YM50 centrifugal filter devices (Amicon-Millipore).

2.10. Electron microscopy

Protein fractions from DEAE-chromatography were stored in 0.05 M Tris/HCl, pH 7.4, at 4 °C until use. They were diluted with the same buffer immediately before grid preparation to give final protein concentrations of 5-10 µg/mL. Mass measurement of unstained samples by scanning transmission electron microscopy (STEM) (Engel, 1978) was performed as described (Kayser et al., 2009). Sample aliquots (4 µL) were adsorbed for 1 min to the glow-discharged grids, which were then washed on 6 drops of quartz bi-distilled water, blotting between each step, immediately plunge-frozen in liquid nitrogen and freeze-dried at -80 °C overnight in the Vacuum Generators STEM HB-5 microscope. Dark-field images were recorded at an accelerating voltage of 80 kV and a nominal magnification of 200,000x. Recording doses varied from 300 to 660 electrons/nm². Beam-induced mass-loss was determined as described (Kayser et al., 2009; Müller et al., 1992). Tobacco mosaic virus (kindly supplied by R. Diaz-Avalos, Janelia Research Campus, Virginia, USA) was used for absolute mass calibration. For negative stain transmission electron microscopy (TEM), samples were diluted and adsorbed and washed as outlined above, negatively stained on two 10-µL droplets of 2% uranyl acetate, pH 4.5, blotting between each step, and air dried. Samples were imaged at a nominal magnification of 130,000x, using a Philips CM10 TEM operating at 80 kV. Micrographs were recorded on a 2,000 by 2,000 pixels charge-coupled device camera (Veleta, Olympus soft imaging solutions GmbH, Münster, Germany) mounted in the 35 mm port of the TEM, yielding a final pixel size of 3.8 nm at the specimen level.

2.11. Image processing

STEM images of unstained samples were analysed using MASDET (Krzyżanek et al., 2009). Mass values were corrected for beam-induced mass-loss, scaled according to the mass standard, binned in a histogram and described by Gauss curves. The overall experimental uncertainty of the results was estimated from the standard error, SE, of the data and the 5% uncertainty in the calibration of the microscope, E_{cal} , being $\sqrt{(SE^2 + E_{cal}^2)}$. TEM images were analysed using the EMAN1 software package (Ludtke et al., 1999). Over 1,000 particle projections were manually selected for each negatively-stained sample using the boxer subroutine, aligned and classified; the maximum number of projections per class was restricted to 50. Fifteen representative class averages are shown for each sample.

3. Results

3.1. Biliprotein from larval hemolymph, a very high-density lipoprotein

Purification of the blue hemolymph protein by gradient chromatography on DEAE-Sephacel provided two pale/yellowish flow-through fractions, both with a single absorption maximum at 277 nm. A blue protein eluted at mid-gradient as a single peak. All blue fractions showed almost identical UV-visible spectra with maxima at ~662 nm, 374 nm and 278 nm (Fig. 1A, upper panel); their purity, defined by the ratio of A_{662nm}/A_{278nm} , was typically <0.15. Narrower

1 ranges of AS saturation revealed that most (up to 70 %, based on $A_{662\text{nm}}$) of the blue protein
2 precipitated between 55 % and 65 % AS. Importantly, 65-70 % AS fractions provided almost
3 pure blue protein; however, the yields were usually low (5 to 10 % of total).

4
5 On size-exclusion chromatography, the blue protein eluted between catalase (232 kDa) and
6 ferritin (440 kDa), suggesting a molecular mass in the range of 350 kDa. The purity ratio,
7 $A_{662\text{nm}}/A_{278\text{nm}}$, was 0.18; the ratio of the bilin absorption maxima, $A_{662\text{nm}}/A_{374\text{nm}}$, was 1.5.

8
9 On SDS-PAGE, samples purified either by ion exchange and/or by size-exclusion
10 chromatography consistently showed two prominent bands corresponding to masses of about
11 150 kDa and 85 kDa, respectively (Fig. 2A,B). Obviously, these proteins aggregate to
12 complexes, which are not discriminated on the TSK column. Analysis of fractions across the
13 blue-protein elution peaks from DEAE- and TSK-chromatography, respectively, confirmed
14 that the 150-kDa and 85-kDa proteins almost co-eluted on both chromatographic matrices
15 (Fig. 2A,B). Close inspections of gels and spectral analyses indicated that the 150-kDa protein
16 elutes slightly ahead of the 85-kDa protein on the anion exchange matrix (Fig. 2A), and it is
17 slightly retarded on the size-exclusion matrix (Fig. 2B). Therefore, separation of these
18 proteins could not be achieved by either method.

19
20 As mentioned, almost quantitative precipitation of the 85-kDa protein, together with 150-kDa
21 protein, was obtained at 65% AS saturation, while complete precipitation of the 150-kDa
22 protein required up to 70 % saturation. Hence, fractions obtained at >65% AS saturation
23 usually only contained the 150-kDa protein, as shown by SDS-PAGE (Fig. 2C). These
24 fractions were blue with spectral ratios of >0.22 for $A_{662\text{nm}}/A_{278\text{nm}}$ and >1.6 for $A_{662\text{nm}}/A_{374\text{nm}}$
25 indicating that the bilin was exclusively associated with the 150-kDa protein and not with an
26 85-kDa protein. Conversely, samples that only contained 85-kDa proteins were colourless.

27
28 Sequence analysis of tryptic peptides of the 150-kDa protein revealed very high similarity
29 with VHDL from *Helicoverpa zea* (UniProt: A8CMX8) (Supplementary Fig. S1): most expect
30 values were between E-12 and E-6. FASTA search with all peptides resulted in VHDL from
31 *H. zea* as the only insect protein. The elution behaviour of the native 150-kDa protein from
32 *Spodoptera* on the DEAE matrix suggested a weakly acidic pI value; this would be in
33 agreement with the (theoretical) pI value of 5.57 for VHDL from *H. zea*
34 (<http://www.expasy.org/protparam/>). Two peptides were glycosylated with a proximally
35 fucosylated truncated oligomannosidic structure (GlcNAc₂Man₂Fuc) and a high-mannose
36 glycan (GlcNAc₂Man₉), respectively. The site of the fucosylated glycan is at the N-terminus
37 of the respective peptide (marked in Supplementary Fig. S1); the site of the high-mannose
38 glycan could not be identified. Mass spectrometry of the native biliprotein provided major
39 two (monomeric) masses of 178 kDa and 180 kDa; this mass difference suggested a
40 difference in glycosylation.

41
42 Four proteins with masses of about 85 kDa were identified as members of the hexamerin class
43 of insect storage proteins by MS sequencing or MALDI peptide mass fingerprinting
44 (Supplementary Fig. S2). Two proteins from the run-through of the DEAE-column showed
45 high similarity to the methionine-rich storage protein (UniProt: Q9U5K5) (expect value: E-
46 13) and the moderately methionine-rich storage protein (UniProt: Q9U5K6) (expect value: E-

3), respectively, both from *S. litura*. A third protein from the run-through was similar to the arylphorin subunit (*UniProt: Q9U5K4*) from *S. litura* (expect value: E-17). The other prominent 85-kDa protein present in blue gradient fractions, showed high sequence similarity to a hexamerin (*UniProt: A6YRR6*) from *S. exigua* (expect value: E-11); at least one site was occupied by a high-mannose type glycan, as in VHDL.

As the 150-kDa biliprotein was a lipoprotein (a VHDL) by contrast to the 85-kDa hexamerins, the two protein classes were examined by KBr density gradient centrifugation. Two well-separated bands were obtained with samples from each the DEAE and the TSK column: an upper blue band with a density of 1.22 g/mL, compatible with a VHDL, and a lower yellowish band with a density of 1.26 g/mL (Fig. 3A). The isolated yellowish run-through fractions containing 85-kDa proteins from *S. littoralis* behaved like the *C. vinula* biliprotein, a hexamerin with a density of 1.26 g/mL (Fig. 3B,C). Samples containing pure 150-kDa protein provided a single blue band with a density of 1.22 g/mL (Fig. 3D). A mixture of the biliproteins from *S. littoralis* and *C. vinula* separated into two blue bands according to their densities (Fig. 3E).

These studies conclusively prove that the bilin is associated exclusively with VHDL in the larval hemolymph of *S. littoralis*.

3.2. Biliprotein from eggs, a vitellin

Extracts from eggs were green, not blue like samples from hemolymph. On DEAE-chromatography, a green protein was completely eluted at low salt (<0.25 M NaCl). It precipitated when diluted with water or dialysed against NaCl-free buffer, allowing easy concentration. The UV-visible spectra showed a relatively low and broad maximum at 650 to 670 nm and a double peak between 370 nm and 390 nm (Fig. 1B, upper panel).

On size-exclusion chromatography, the egg biliprotein behaved similarly to the hemolymph biliprotein and to the hexamerins. This result is thought to be due to the low resolution of the TSK-matrix in the respective mass range. SDS-PAGE of the egg biliprotein from the DEAE- and TSK columns, respectively, showed identical protein pattern with one prominent band at ~150 kDa and two minor bands at ~180 kDa and ~37 kDa, respectively (Fig. 4). MS sequencing of cleavage peptides identified the proteins of ~150 kDa and ~180 kDa as vitellogenin-related, based on high similarity to vitellogenin from *S. litura* (*UniProt: A7XS62*) (expect values: E-11 and E-10, respectively) and many other vitellogenins. As no N-terminal peptides were obtained from both proteins they may represent processed forms of vitellogenin (Supplementary Fig. S3). The mass difference between the two proteins may be due to differences in glycosylation and/or other modifications. The ~37-kDa protein was identified as the cleaved N-terminal sequence of vitellogenin. Another minor protein with a mass >200 kDa according to SDS-PAGE (not shown), was identified as the full-length vitellogenin precursor (expect value: E-20), which is processed to the exported form, vitellin (Supplementary Fig. S3). VHDL and hexamerins, both prominently present in larval hemolymph, were not detected in eggs by SDS-PAGE. Conclusively, the bilin in eggs was

exclusively associated with vitellin, in accordance with the behaviour of this biliprotein on the DEAE matrix.

Extracts from eggs, three days after deposition, were light reddish-brown with a green hue indicating the presence of biliprotein. This protein, precipitated at 45-67 % AS saturation, showed a flat broad absorption maximum in the range of 650 nm, comparable to the spectrum of extracts from fresh eggs (6-12 h after deposition). The spectrum of the isolated bilin was identical to that from hemolymph (see 3.3.). On the KBr gradient (not shown), the density of the green protein was ≥ 1.23 g/mL, slightly higher than that of the blue hemolymph protein. Extracts from four-day old eggs, i.e., less than one day before larval hatching, were dark reddish-brown and the dissolved AS sediments were devoid of biliprotein, according to visible spectra. The same negative result was obtained from extracts from freshly hatched larvae.

3.3. 'The bilin', a mixture of three farnesyl biliverdins

The UV-visible spectra of *Spodoptera* bilin (total sample, see below) from larval hemolymph were comparable to those from eggs. In both cases the absorption maxima and peak ratios were almost identical to those of *Cerura* bilin, and the maxima were blue-shifted by 30 -40 nm, compared to those of biliverdin IX α (Table 1).

Methylation and acetylation of *Spodoptera* bilin yielded less polar products on TLC, indicating the presence of free carboxylic acid and hydroxyl groups, respectively (data not shown). Chromate degradation of native bilin (performed with hemolymph samples) resulted in a polar zone near the site of application on TLC. Chromate degradation of bilin dimethyl ester provided a pyrrole dialdehyde methyl ester, which co-chromatographed with that of the dimethyl ester of biliverdin IX α ; a second, more polar aldehyde product was formed at a comparable amount, according to the staining intensity. The latter dialdehyde product indicated that one of the inner pyrrole units carried a polar group, as is unique to *Cerura* bilin. Less polar products, such as methyl vinyl maleimide (obtained with the biliverdin isomers β , γ and δ) were not detected. Taken together, *Spodoptera* bilin was shown to belong to the IX α series of biliverdin isomers, and was degraded to products identical to those of *Cerura* bilin.

Analysis by reverse-phase HPLC, however, revealed that the bilin from hemolymph was not a single compound, but a mixture of three bilins with retention times of 12 min (bilin I), 16.5 min (bilin II) and 21 min (bilin III). Their relative amounts varied among samples, with bilin III dominating; in some samples bilin III accounted for >90 % of the total bilins (Fig. 5A). The molecular masses of the bilins were obtained by MS of the HPLC fractions: 852 for bilin I corresponding to C₄₈H₆₀N₄O₁₀ (found m/z 853.439, calculated 853.438 for [M+H]⁺), 836 for bilin II corresponding to C₄₈H₆₀N₄O₉ (found m/z 837.443, calculated 837.443 for [M+H]⁺), and 820 for bilin III corresponding to C₄₈H₆₀N₄O₈ (found m/z 821.447, calculated 821.448 for [M+H]⁺). In another sample, analysed as Na⁺ adducts, the two major fractions were identified as bilin II (found m/z 859.424, calculated m/z 859.425 for [M+Na]⁺) and bilin III (found m/z 843.429, calculated m/z 843.430 for [M+Na]⁺).

High-resolution MS of whole samples (not separated by HPLC) of hemolymph bilin confirmed the above results (found m/z for $[M+H]^+$): 853.43845 for bilin I, 837.44351 for bilin II, and 821.44859 for bilin III (Fig. 6). Their relative amounts estimated from the ion intensities agreed with bilin III being the major form in hemolymph. In one sample, the composition was 9 % bilin I, 27 % bilin II and 64 % bilin III; another sample contained 21 % bilin I, 27 % bilin II and 52 % bilin III. The results from further samples were consistent with the presence of three different bilins in hemolymph VHDL in the relative ratio of bilin I < bilin II < bilin III. Loss of water was observed for all bilins; the addition of water was seen for bilin I at m/z 871.449, and for bilin II at m/z 855.453, not for bilin III (Fig. 6).

The bilins associated with the vitellin-type biliprotein from eggs were the same as those identified as $[M+H]^+$ from the hemolymph VHDL-type protein: bilin I ($C_{48}H_{60}N_4O_{10}$; found m/z 853.437); bilin II ($C_{48}H_{60}N_4O_9$; found m/z 837.442) and bilin III ($C_{48}H_{60}N_4O_8$; found m/z 821.447). Compared to hemolymph, however, egg samples contained more bilin I than bilin II and bilin III (Fig. 5B). The composition was typically 68 % bilin I, 25 % bilin II and 7 % bilin III, as estimated from mass spectra. The proportions of the three bilins varied among samples, possibly reflecting heterogeneity with respect to the developmental age of the eggs. The typical ratio among the three bilins from eggs may be described as bilin I > bilin II > bilin III. On HPLC, bilin I from eggs and from hemolymph exhibited the same retention time as CV-bilin, in agreement with their identity, as shown by MS (Fig. 5C). The presence of lipid in extracts from the egg biliprotein made it necessary to apply less mild conditions when processing the bilin samples; this resulted in the formation of some artefacts, which were not further studied (Fig. 5B). Formation of artefacts was also observed with CV-bilin, relating to oxidation of the terminal hydroxyl group at the farnesyl chain and to the formation of esters and/or ethers (Kayser et al., 2014).

The molar bilin-to-protein ratio was calculated for the *Spodoptera* hemolymph biliprotein based on the absorption data of the holoprotein of highest purity (i.e., lowest A662nm/A278nm ratio). An estimate of ~0.5 bilin per VHDL monomer was obtained from several calculations, indicating a stoichiometry of ~1 bilin per VHDL dimer, the native biliprotein complex. Contamination by bilin-free VHDL would have resulted in a ratio of <1 per dimer. Corresponding calculations could not be performed for the egg biliprotein due to insufficient spectral quality in the Soret region (Fig. 1B).

3.4. Bilin conformation varies with the type of binding protein

The circular dichroism (CD) spectrum of the VHDL-type biliprotein from hemolymph showed a strong negative band at ~650 nm and a strong positive band in the Soret region at ~370 nm, both corresponding to their visible absorbance maxima (Fig. 1A). The green vitellin-type biliprotein from eggs exhibited less optical activity, but showed the same type of CD spectrum; the maxima were shifted to ~670 nm and ~380 nm, respectively (Fig. 1B, lower panel). Free bilin gave a similar CD spectrum; however, its CD activity was weaker than that of the two protein-bound forms (Fig. 1C, lower panel). All of these spectra indicate a predominance of the left-handed *M*-conformation of the bilin in the equilibrium of the two enantiomers, if a cyclic-helical conformation is assumed.

The colour of a solution of hemolymph VHDL biliprotein changed on the addition of α -cyclodextrin from intense blue(-green) to light olive-green, due to a blue-shift of about 35 nm and a 66 % decrease of the long-wave absorption peak. As the Soret absorption was little affected, the peak ratio decreased from 1.20 to 0.46 (Fig. 1A, upper panel). Corresponding changes were seen in the CD spectrum (Fig. 1A, lower panel). The colour and visible absorption of a solution of the egg vitellogenin biliprotein changed very little in the presence of α -cyclodextrin; however, the effect on the CD spectrum was comparable to that seen with the VHDL biliprotein (Fig. 1B). Although in the presence of α -cyclodextrin the UV-visible and CD spectra of the biliproteins were comparable to those of free bilin (Fig. 1C), the bilin clearly remained associated with the carrier protein as it did not pass 50-kDa filters. In contrast to the biliproteins from *S. littoralis*, spectral changes were not seen when the arylphorin-based biliprotein from *C. vinula* was exposed to α -cyclodextrin (Fig. 7). This negative result may be due to the absence of lipid in arylphorins, as in hexamerins in general.

3.5. Electron microscopy provides a first view of VHDL structure

Several protein samples obtained on the purification of the hemolymph biliprotein were examined. Negative-stain TEM of samples from the (colourless) run-through fractions of the DEAE-column revealed a homogeneous population of protein complexes (Fig. 8A, top panel). Six domains became clearly evident when their circular top-view projections were classified by single particle analysis and averaged. These form the corners of two triangles, resulting in D3 symmetry and a maximum diameter of ~14 nm (Fig. 8A, centre panel). A central ring-like domain can also be distinguished. Correspondingly, generated rectangular side-view averages indicate that the complex is ~11 nm high. Analysis of unstained complexes by STEM gave a mass of 525 ± 46 kDa ($n = 1497$; overall uncertainty = ± 26 kDa) (Fig. 8A, bottom panel).

The blue fraction from the DEAE gradient was less homogeneous, as expected from SDS-PAGE analysis (Fig. 2A,B). Two distinct populations were distinguished by negative-stain TEM and on mass analysis by STEM (Fig. 8B). While many of the complexes were clearly the hexamers found in the run through fractions (measured mass, 490 ± 43 kDa ($n = 368$; uncertainty = ± 25 kDa)), others appeared to be dimers with a mass of 352 ± 48 kDa ($n = 254$; overall uncertainty = ± 18 kDa).

The blue fractions obtained by AS precipitation at 65-70% saturation contained almost exclusively this second particle population (Fig. 8C, top panel). Many of the class averages are side-views of ~15 nm-wide and ~14 nm-long dimeric complexes that have a stain-filled cavity at one end; other class averages might correspond to monomers (Fig. 8C, centre panel). STEM mass analysis gave a histogram with a minor peak at 185 ± 43 kDa ($n = 126$; overall uncertainty = ± 10 kDa) arising from a small population of monomers, and a major peak at 366 ± 44 kDa ($n = 623$; uncertainty = ± 18 kDa), clearly representing dimers (Fig. 8C, bottom panel). Accordingly, most of the 150-kDa protein documented by SDS-PAGE (Fig. 2A-C) was present as a dimer. The occurrence of some monomers indicates that the complex is maintained by relatively weak bonds, in contrast to the hexamerins.

4. Discussion

4.1. The farnesylated bilins are presumably derived from heme A

The present study of *S. littoralis* confirms and extends preceding results on the farnesylated bilin, CV-bilin, isolated from *C. vinula* (Kayser et al., 2014): Firstly, in addition to CV-bilin (described here as bilin I) two other closely related bilins, termed bilins II and bilin III, which may be intermediates of the proposed transformation of heme A to bilin I. Farnesyl biliverdins IX α represent possible degradation products of heme A (or its precursor, heme O) that may also occur in other organism. Secondly, by contrast to *C. vinula*, we show that in *S. littoralis* the bilins are associated with two different proteins depending on the developmental stage. Thirdly, there is a developmental shift from bilin III to bilin I, possibly pointing to a role of bilin I in embryonic development, as already suggested for CV-bilin (Kayser et al., 2014).

No difference was found between bilin I from *S. littoralis* and CV-bilin in the results from UV-visible spectrometry, CD spectroscopy, high-resolution MS, oxidative degradation and reverse-phase HPLC. NMR spectrometry could not be applied in the present work because of the limited amounts of bilins from *Spodoptera*; hence, a NMR confirmation of the identity of the two bilins was not possible. The two new bilins, II and III, from *Spodoptera* differed from bilin I by having one, respectively two, oxygen atoms less. This is in agreement with their different polarities on reverse-phase HPLC. The observed developmental shift from bilin III to bilin II and bilin I led us propose that these two oxygen atoms are those of the epoxide and the terminal hydroxyl group of the farnesyl chain, as present in CV-bilin. Assuming that bilin I is identical to CV-bilin, then bilin II should only have one of these oxygen functions, and bilin III should have an unmodified farnesyl chain as in heme A (Brown et al., 2002). Overall, the presence of all formally expected intermediates of the proposed transformation of heme A to bilin I supports the hypothesis that heme A (or heme O) is the only source of these novel farnesylated biliverdins. As depicted in Figure 9, the proposed pathway would involve first α -cleavage of heme A and reduction of the formyl group at C-13 (not necessary in case of heme O) resulting in bilin III, but leaves open the sequence of the further modifications at the hydroxyethylfarnesyl side-chain: epoxidation of one of the double bonds and hydroxylation of one of the terminal methyl groups of the farnesyl chain of bilin III. Such modifications of a farnesyl chain would not be unique as corresponding structural changes are known from the biosynthesis of several insect juvenile hormones from farnesol, as shown *in vivo* and *in vitro* (Richard et al., 1989; Mauchamp et al., 1999; Kotaki et al., 2009). The proposed epoxy and hydroxy intermediates (bilin IIa and bilin IIb; Fig. 9) could not be discriminated by the methods used. MS provided evidence that one molecule of water was added to bilin I and bilin II, but not to bilin III (data not shown); as this reaction most likely takes place at the epoxide, the proposed bilin II intermediate may mainly be the epoxybilin IIa. As all three bilins are present in the two biliproteins, discussed below, the proposed transformations from bilin III to bilin I may take place within the bilin-protein complex. Larval biliproteins, like other hemolymph proteins and vitellins, are synthesized in the fat body and/or epidermis

(Kayser, 2005). We are not aware of studies that showed the synthesis of bilins in insect ovaries or developing oocytes.

By contrast to *S. littoralis*, where the biliproteins are associated with three structurally related bilins, the *C. vinula* biliprotein is linked to a single bilin, CV-bilin (Kayser et al., 2014), that is probably identical to bilin I from *S. littoralis*. However, reverse-phase HPLC of bilin samples from *C. vinula* showed one, occasionally two, very small peaks of unknown identity, in addition to CV-bilin (Fig. 5C). As their retention times were identical to those of bilin II and III, these trace compounds may be these two putative precursors, which are more prominent in *Spodoptera*. Obviously, the developmental time scale of the proposed transformation of heme A to CV-bilin and bilin I, respectively, is quite different in the two species. On the other hand, a common feature of *C. vinula* and *S. littoralis* is the presence of farnesyl bilin(s) in eggs. This fact supports the idea that this unusual type of bilin or a derivative thereof, whether protein-bound or not, may play a role in embryonic development; fully developed *Spodoptera* larvae are devoid of bilin, as we show here. There are only few studies on how bilins may function in cellular regulation, discussed in preceding work (Kayser et al., 2014). While several other insects are also known to contain bilin(s) in eggs, this is not true for all. A well-studied example is the egg-specific biliprotein from the bug *Riptortus clavatus*; other biliproteins, collectively called cyanoproteins, are present in the hemolymph of males and females (Chinzei et al., 1990). The “biliverdin”-ligand has not been identified in this bug, however.

4.2. Very-high density lipoprotein and vitellin as development-specific bilin carriers

While in *C. vinula* the single bilin is linked to an arylphorin, a member of the large hexamerin class of insect storage proteins (Kayser et al., 2009), the bilins in larval hemolymph of *S. littoralis* are associated with VHDL, a glycoprotein that is not related to the hexamerins. Very similar VHDL-type biliproteins have been purified from *Heliothis (Helicoverpa) zea* (Haunerland and Bowers, 1986; Sum and Haunerland, 2007) and other moths (Jones et al., 1988; Yoshiga and Tojo, 1995) and described as homodimers of 165-kDa monomers that are glycosylated and contain lipid, resulting in a density of 1.26 g/mL. Their chromophores have been described as biliverdin-like. Size-exclusion chromatography and mass measurement by STEM showed that the *S. littoralis* VHDL-type biliprotein also occurs as a dimer (see 4.3.); its lower density of 1.22 g/mL may indicate a higher lipid load.

The only known primary sequence of an insect VHDL is that from *H. zea* (Sum and Haunerland, 2007). It shows sequence similarity to vitellogenins suggesting that VHDL may be derived from vitellogenins by several losses of internal and C-terminal sequences during evolution; therefore, VHDL is significantly smaller than vitellogenins. Our results on *S. littoralis* suggest that the bilins, possibly only a fraction thereof, are transferred from VHDL in larval hemolymph to vitellogenin, or vitellin, in the course of egg development.

Vitellogenins are synthesized and processed in the fat body into a large and a small protein (Chen et al., 1997; Tufail and Takeda, 2008). In the present study, the ~37-kDa protein was identified as an N-terminal fragment and the two larger proteins of ~150 kDa and ~180 kDa, respectively, as C-terminal sequences. The two latter proteins may differ in glycosylation

and/or phosphorylation, which are typical modifications of vitellogenins taking place prior to their secretion into hemolymph as vitellins (Tufail and Takeda, 2008). In honey bee, a 40-kDa protein was first identified as the N-terminal cleavage peptide of vitellogenin (Havukainen et al., 2011).

The present isolation of a vitellin-based biliprotein from the eggs of *S. littoralis* compares well with the corresponding protein purified from adult female hemolymph of *S. litura* (Maruta et al., 2002). The mass of this biliprotein was estimated as ~490 kDa consisting of two ~188 kDa monomers. Hexameric storage proteins of ~500 kDa composed of ~85 kDa subunits have also been reported from *S. litura* (Tojo and Yoshiga, 1994) comparing well with the two methionine-rich proteins and the arylphorin/hexamerin from *S. littoralis*.

VHDL from *H. zea* and vitellins in general contain up to 10 % lipid with diacylglycerol and phospholipids as the major components (Haunerland and Bowers, 1986; Jones et al., 1988; Tufail and Takeda, 2008; Maruta et al., 2002). α -Cyclodextrin has been reported to efficiently extract lipids, diacylglycerol with high preference, from insect lipoproteins (Jouni et al., 2000). Exploiting this property, we show here for the *S. littoralis* VHDL that the presence of lipids affects the bound bilins in some way, as documented by changes in the UV-visible and CD spectra. The different extents to which the spectra of the biliproteins with VHDL and vitellin are changed by α -cyclodextrin may be due to the quantity and composition of lipids bound, as well as to structural features of the bilin binding site, which is presumably different in these two classes of lipoproteins. The way how lipids affect the spectral properties of bound bilins remains open, it could be indirect through a conformational change of the protein or due a close contact between the bilin and the protein. Interestingly, α -cyclodextrin had no effect on the UV- and CD-spectra of the lipid-free arylphorin-type biliprotein from *C. vinula*.

There is a remarkable difference between the VHDL-type and vitellin-type biliproteins from *S. littoralis*, described here, and the recently reported arylphorin-type biliprotein from *C. vinula* (Kayser et al., 2009). While the CD-spectra of the two biliproteins from *S. littoralis* exhibited a strong negative band in the 650-nm region, which indicates a predominance of the left-handed *M*-conformation of the bilin in the equilibrium of the enantiomers, the spectrum of the *C. vinula* biliprotein with an arylphorin as carrier suggested a right-handed *P*-conformation of the bound bilin. In solution, CV-bilin has been shown to adopt an unusual 10*E* configuration (Kayser et al., 2014); its configuration in the protein-bound state is unknown, however. The same holds for the bilins from *S. littoralis* in complex with the two different proteins. Therefore, the above assignments of *P*- and *M*-conformation, respectively, may not relate to a cyclic-helical conformation of these bilins. The difference in the CD spectra of the biliproteins from *Spodoptera* and *Cerura* may (mainly) be due to the presence and absence, respectively, of lipid in the vicinity of the bilin binding site.

4.3. Electron microscopy provides a first molecular view of VHDL

The EM analysis of the DEAE run-through fractions from *S. littoralis*, in which four different proteins were identified, nevertheless revealed a structurally very homogenous sample. The size and shape of all complexes were almost identical to those of the recently studied

1 arylphorin-type hexamerin from *C. vinula* (Kayser et al., 2009). Like the latter, the *S. littoralis*
2 complexes are organized as six domains composed of two trimers (Fig. 8A). Moreover, the
3 measured mass of ~525 kDa for the complex and of ~85 kDa for a subunit support the
4 assignment of the *Spodoptera* proteins as members of the hexamerin class of insect storage
5 proteins. Two are arylphorin-type hexamerins, and two belong to the methionine-rich class of
6 hexamerins. The latter type of hexamerins is expected to be constructed in the same way as
7 the evolutionarily closely related arylphorins and, hence, unlikely to be discriminated by EM.
8 Moreover, studies of several arylphorins have even revealed a common quaternary structure
9 with arthropod hemocyanins, which is explained by their common ancestry (Kayser et al.,
10 2009; Markl et al., 1992; Kim et al., 2003).

11
12 The mass of ~360 kDa determined for the pure VHDL-type blue biliprotein from larval
13 hemolymph by STEM (Fig. 8C), is in good agreement with the 350-kDa estimate obtained by
14 gel permeation chromatography. As SDS-PAGE resulted in a single band at a mass of 150
15 kDa, the native complex is thought to be a homodimer. The STEM sample also contained a
16 small population of 185-kDa particles, which is the mass expected for monomers (Fig. 8C).
17 The suggestion that VHDL is a homodimer is supported by negative-stain TEM, which
18 showed the larger complexes to have two similar domains (Fig. 8C). The present EM study is
19 the first performed on a VHDL from a non-vertebrate species and on a novel type of
20 biliprotein. The vitellin-based biliprotein from the eggs of *S. littoralis* was not studied by EM.

21 22 **Acknowledgments**

23 We are grateful to Helma Wennemers (ETH Zürich, Switzerland) and Stefan Binder (Ulm
24 University, Germany) for giving access to the CD spectrometer and the Beckman Optima
25 ultracentrifuge, respectively, and also to Christian Guenat (Novartis, Basel, Switzerland) for
26 high-resolution mass spectra, to Françoise Erne-Brand for assistance with the STEM
27 microscopy and to Andreas Engel and Henning Stahlberg for supporting the STEM/TEM
28 project (University Basel, Switzerland). This work was made possible by the generous
29 continuous supply of insects by Syngenta Crop Protection (Basel, Switzerland).

30 31 **Conflict of interest**

32 The authors declare no conflict of interest.

33 34 **Author contribution**

35 H.K. planned the project, performed all biochemical experiments, analysed data and wrote the
36 manuscript, except for the EM part. M.N. carried out mass spectrometry, protein analyses and
37 data bank searches. P.R. and S.A.M. collected and analysed the EM data and provided the
38 related text.

39 40 **Funding**

41 The STEM and TEM microscopy was funded by the Maurice E. Müller Foundation, by Swiss
42 National Foundation Grant 3100A0-108299 to Andreas Engel and by the Swiss systems
43 biology initiative SystemsX.ch (grant CINA to Andreas Engel and Henning Stahlberg).

References

- Brown, K.R., Allan, B.M., Do, P., Hegg, E.L., 2002. Identification of novel hemes generated by heme a synthase: evidence for two successive monooxygenase reactions. *Biochem.* 41, 10906-10913.
- Burmester, T., 1999. Evolution and function of the insect hexamerins. *Eur. J. Entomol.* 96, 312-225.
- Chen, J.S., Sappington, T.W., Raikhel, A.S., 1997. Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals ancient common ancestry. *J. Mol. Evol.* 44, 440-451.
- Chinzei, Y., Haruna, T., Miura, K., Numata, H., Nakayama, S., 1990. Purification and characterization of biliverdin-associated cyanoprotein from eggs and hemolymph of the bean bug, *Riptortus clavatus* (Heteroptera: Alydidae). *Insect Biochem.* 20, 545-555.
- Engel, A., 1978. Molecular weight determination by scanning transmission electron microscopy. *Ultramicroscopy* 3, 273-281.
- Frankenberg, N., Lagarias, J.C., 2003. Biosynthesis and biological functions of bilins, in: Kadish, K.M., Smith, K.M., Guillard, R. (Eds.), *The Porphyrin Handbook*, Academic Press, Amsterdam, pp. 211-236.
- Haunerland, N.H., Bowers, W.S., 1986. A larval specific lipoprotein. Purification and characterization of a blue chromoprotein from *Heliothis zea*. *Biochem. Biophys. Res. Commun.* 134, 580-586.
- Havukainen, H., Halskau, Ø., Skjaerven, L., Smedal, B., Amdam, G.V., 2011. Deconstructing honeybee vitellogenin: novel 40 kDa fragment assigned to its N terminus. *J. Exp. Biol.* 214, 582-592.
- Hederstedt, L., 2012. Heme A biosynthesis. *Biochim. Biophys. Acta* 1817, 920-927.
- Heirwegh, K.P.M., Blanckaert, N., Van Hees, G., 1991. Synthesis, chromatographic purification, and analysis of isomers of biliverdin IX and bilirubin IX. *Anal. Biochem.* 195, 273-278.
- Jones, G., Ryan, R.O., Haunerland, N.H., Law, J.H., 1988. Purification and characterization of a very high density chromolipoprotein from the hemolymph of *Trichoplusia ni* (Hübner). *Arch. Insect Biochem.* 7, 1-11.
- Jouni, Z.E., Zamora, J., Snyder, M., Montford, W.R., Weichsel, A., Wells, M.A., 2000. α -Cyclodextrin extracts diacylglycerol from insect high density lipoproteins. *J. Lipid Res.* 41, 933-939.

- 1
- 2 Kayser, H., 1985. Pigments, in: Kerkut, G.A., Gilbert, L.I. (Eds.), Comprehensive Insect
- 3 Biochemistry, Physiology and Pharmacology, vol. 10, Pergamon Press, Oxford, pp. 367-415.
- 4
- 5 Kayser, H., 2005. Lipocalins and structurally related ligand-binding proteins, in: Gilbert, L.I.,
- 6 Iatrou, K., Gill, S. (Eds.), Comprehensive Molecular Insect Science, vol. 4, Elsevier, Oxford,
- 7 pp. 267-306.
- 8
- 9 Kayser, H., Mann, K., Machaidze, G., Nimtz, M., Ringler, Ph., Müller, S.A., Aebi, U., 2009.
- 10 Isolation, characterisation and molecular imaging of a high-molecular-weight insect
- 11 biliprotein, a member of the hexameric arylphorin protein family. J. Mol. Biol. 389, 74-89.
- 12
- 13 Kayser, H., Wray, V., Nimtz, M., 2014. Structure of a novel farnesylated bilin from an insect
- 14 – formation by α -cleavage of heme A of mitochondrial cytochrome c oxidases? FEBS J. 281,
- 15 2366-2376.
- 16
- 17 Kim, S.R., Yoon, H.J., Park, N.S., Lee, S.M., Moon, J.Y., Seo, S.J., Jin, B.R., Sohn, H.D.,
- 18 2003. Molecular cloning, expression, and characterization of a cDNA encoding the
- 19 arylphorin-like hexameric storage protein from the mulberry longicorn beetle, *Apriona*
- 20 *germari*. Arch. Insect Biochem. Physiol. 53, 49-65.
- 21
- 22 Kotaki, T., Shinada, T., Kaihara, K., Ohfune, Y., Numata, H., 2009. Structure determination
- 23 of a new juvenile hormone from a heteropteran insect. Org. Lett. 11, 5234-5237.
- 24
- 25 Krzyžánek, V., Müller, S.A., Engel, A., Reichelt, R., 2009. MASDET - A fast and user-
- 26 friendly multiplatform software for mass determination by dark-field electron microscopy. J.
- 27 Struct. Biol. 165, 78-87.
- 28
- 29 Ludtke, S.J., Baldwin, P.R., Chiu, W., 1999. EMAN: semiautomated software for high-
- 30 resolution single-particle reconstructions. J. Struct. Biol. 128, 82-97.
- 31
- 32 Markl, J., Burmester, T., Decker, H., Savel-Niemann, A., Harris, J.R., Süling, M., Naumann,
- 33 U., Scheller, K., 1992. Quaternary and subunit structure of *Calliphora* arylphorin as deduced
- 34 from electron microscopy, electrophoresis, and sequence similarities with arthropod
- 35 hemocyanin. J. Comp. Physiol. B 162, 665-680.
- 36
- 37 Maruta, K., Yoshiga, T., Katagiri, C., Ochiai, M., Tojo, S., 2002. Purification and
- 38 characterization of biliverdin-binding vitellogenin from the hemolymph of the common
- 39 cutworm, *Spodoptera litura*. Arch. Insect Biochem. 50, 97-106.
- 40
- 41 Mauchamp, B., Darrouzet, E., Malosse, C., Couilland, F., 1999. 4'-OH-JH-III: an additional
- 42 hydroxylated juvenile hormone produced by locust corpora allata in vitro. Insect Biochem.
- 43 Molec. Biol. 29, 475-480.
- 44

- Müller, S.A., Goldie, K.N., Bürki, R., Häring, R., Engel, A., 1992. Factors influencing the precision of quantitative scanning transmission electron microscopy. *Ultramicroscopy* 46, 317-334.
- Nimtz, M., Conradt, H.S., Mann, K., 2004. LacdiNAc (GalNAc β 1-4GlcNAc) is a major motif in N-glycan structures of the chicken eggshell protein ovocleidin-116. *Biochim. Biophys. Acta* 1675, 71–80.
- Richard, D.S., Applebaum, S.W., Sliter, T.J., Baker, F.C., Schooley, D.A., Reuter, C.C., Henrich, V.C., Gilbert, L.I., 1989. Juvenile hormone bisepoxide biosynthesis *in vitro* by the ring gland of *Drosophila melanogaster*: a putative juvenile hormone in the higher Diptera. *Proc. Natl. Acad. Sci. USA* 86, 1421-1425.
- Sum, H., Haunerland, N.H., 2007. VHDL, a larval storage protein from the corn earworm, *Helicoverpa zea*, is a member of the vitellogenin gene family. *Insect Biochem. Mol. Biol.* 37, 1086-1093.
- Tojo, S., Yoshiga, Z., 1994. Purification and characterization of three storage proteins in the common cutworm, *Spodoptera litura*. *Insect Biochem. Mol. Biol.* 24, 729-738.
- Tufail, M., Takeda, M., 2008. Molecular characteristics of insect vitellogenins. *J. Insect Physiol.* 54, 1447-1458.
- Yoshiga, T., Tojo, S., 1995. Purification and characterization of four biliverdin-binding proteins from larval haemolymph of the common cutworm, *Spodoptera litura*. *Insect Biochem. Mol. Biol.* 25, 575-581.

Supporting Information

Proteins from *Spodoptera littoralis* identified by mass spectrometric sequencing of tryptic peptides and homology search:

Figure S1: Very high-density lipoprotein (VHDL) (hemolymph biliprotein)

Figure S2: Hexamerin proteins (not associated with bilins)

Methionine-rich storage protein

Moderately methionine-rich storage protein

Arylphorin

Hexamerin

Figure S3: Vitellogenin (egg biliprotein)

Table 1. Absorption maxima and peak ratios of the bilins from *S. littoralis* (SL) and *C. vinula* (CV), and of biliverdin IX α .

Bilin	UV (nm)	VIS (nm)	UV/VIS ratio
<i>Solvent : Methanol</i>			
SL hemolymph	368	627	3.0
SL eggs	369	629	2.9
CV-bilin*	368	624	2.8
Biliverdin IX α *	377	666	3.7
<i>Solvent : Chloroform</i>			
SL hemolymph	375	621	2.1
SL eggs	372	624	2.0
CV-Bilin*	375	620	2.1
Biliverdin IX α *	378	660	3.3

* Data from Kayser et al. (2014)

Figure legends

Figure 1. UV-visible spectra (upper panels) and CD spectra (lower panels) of biliproteins and bilin from *S. littoralis* in the absence (continuous line) and presence (dashed line) of α -cyclodextrin. **A**, biliprotein from hemolymph; **B**, biliprotein from eggs; **C**, free bilin from egg biliprotein.

Figure 2. SDS-PAGE analysis of samples of hemolymph biliprotein (150 kDa) from *S. littoralis* after AS precipitation and chromatography. **A**, DEAE fractions 1-9 of 60-65% AS precipitate: no separation of 85-kDa and 150-kDa proteins. **B**, TSK fractions 1-5 of 45-60% AS precipitate: no separation of 85-kDa and 150-kDa proteins. **C**, DEAE fractions of 55-65% and 65-70% AS precipitates: 85-kDa and 150-kDa proteins are separated at >65% AS saturation. **M**, molecular mass markers.

Figure 3. KBr density gradient centrifugation of biliprotein and hexamerin preparations from *S. littoralis* (*SL*) and *C. vinula* (*CV*). **A**, 45-60% AS fraction from *SL*; **B**, hexamerin fraction from *SL*; **C**, purified biliprotein from *CV*; **D**, purified biliprotein from *SL*; **E**, mixture of **C** and **D**. The positions of VHDL and hexameris/arylphorins and their measured densities are given on the left and right, respectively.

Figure 4. SDS-PAGE analysis of samples of egg biliprotein obtained by AS precipitation (45-70% saturation) and gradient elution on DEAE-Sephacel: 'DEAE fractions 1-5'. Aliquots of DEAE-fractions were further subjected to TSK size-exclusion chromatography: 'DEAE > TSK'. **M**, molecular mass markers (see Figure 2B).

Figure 5. Analysis by reverse-phase-HPLC of total bilins isolated from purified biliproteins. **A**, *S. littoralis* hemolymph (VHDL); **B**, *S. littoralis* eggs (vitellogenin); **C**, *C. vinula* hemolymph (arylphorin). The peaks represent bilin I (r.t. 12 min), bilin II (r.t. 16.5 min) and bilin III (r.t. 21 min), identified by MS of a different sample. Asterisks indicate preparation artefacts of bilin I (see 3.3 for details).

Figure 6. $[H^+]$ -Mass spectrum of the total bilin fraction from purified VHDL-type biliprotein. The spectrum shows the predominance of bilin III (molecular mass of 820) over bilin II (molecular mass of 836) and bilin I (molecular mass of 852). The spectrum also shows the loss of water from bilin III at m/z 803, and the addition of water to bilin I at m/z 871.

Figure 7. UV-visible absorption spectra (upper panel) and CD spectra (lower panel) of the biliprotein from *C. vinula* in the absence and presence of α -cyclodextrin (CycDex).

Figure 8. Electron microscopy of protein samples from *S. littoralis* hemolymph. Top panels: Overview TEM image of the negatively stained complexes. Centre panels: Representative class averages. Bottom panels: Histogram displaying the masses measured by STEM and the fitted Gaussian curves. **A**. Fractions 1 and 2 from the run-through of the DEAE-column: The 27 class averages obtained from the uniform complexes show elongated side-views and roughly circular top-views that display 6 domains organised in two triangles and a central ring-shaped structure; the measured masses fall into a single peak. **B**. Blue fraction from DEAE gradient: The sample was heterogeneous: some of the 21 class averages obtained are the same as those depicted in A, others have two similar domains and are comparable to the projections depicted in C;

the measured masses fall into two peaks. **C.** Blue fraction obtained with 65-70% AS saturation: The size and shape of the complexes varied; many of the 25 class averages obtained display two similar domains implying that they arise from dimers; others are either different views of dimers or arise from monomers. The measured masses fall into two peaks and are compatible with the presence of a small population of monomers and a large population of dimers. The mass $\pm$ the standard deviation is indicated beside each peak; the overall uncertainties of the measurements were primarily determined by the accuracy of the microscope's calibration, being $\sim \pm 5\%$ of the mass values. Scale bars: 100 nm. Width of the class-average boxes: 30 nm.

Figure 9. Possible bilin biosynthesis from heme A in *S. littoralis*. The same pattern of substituents of heme A and bilin I from *S. littoralis* (proposed to be identical to CV-bilin from *C. vinula*) suggests that these bilins are derived from heme A by α -cleavage and transformation to bilin I *via* the intermediates bilin III and bilin II. The asterisks mark the transformation sites. The isomeric bilins IIa and IIb were not discriminated by reverse-phase HPLC and MS.

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

Fig. 7

Fig. 8

Fig. 9

Supplementary material to the manuscript:

Very high-density lipoprotein and vitellin as carriers of novel biliverdins IX α with a farnesyl side-chain presumably derived from heme A in *Spodoptera littoralis*.

Hartmut Kayser^{a*}, Manfred Nimtz^b, Philippe Ringler^c and Shirley A. Müller^c

^aInstitute of General Zoology and Endocrinology, Ulm University, Helmholtzstrasse 8/1, 89081 Ulm, Germany

^bHelmholtz Centre for Infection Disease, Cellular Proteome Research, Inhoffenstrasse 7, 38124 Braunschweig, Germany

^cCenter for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel, Mattenstrasse 26, 4058 Basel, Switzerland

*To whom correspondence should be addressed.

E-mail: hartmut.kayser@uni-ulm.de

Tel.: +49 731 50 15110

Fax: + 49 731 50 32609

Sequences of proteins from *Spodoptera littoralis* identified by tryptic peptides and homology search:

Figure S1: Very high-density lipoprotein

Figure S2: Methionine-rich storage protein
Moderately methionine-rich storage protein
Arylphorin
Hexamerin

Figure S3: Vitellogenin

Figure S1. Very high-density lipoprotein identified in the course of purification of the biliprotein from larval hemolymph of *S. littoralis*.

The ammonium sulphate precipitate of the biliprotein was subjected to DEAE gradient chromatography and fractions were analysed by SDS-PAGE. Cleavage peptides of the ~150-kDa protein were obtained by in-gel tryptic digestion and sequenced by ESI-MS/MS. The obtained peptide sequences are shown in red; they matched the sequences of the very high-density lipoprotein (VHDL) from *Helicoverpa zea* marked yellow, identified by homology search. Potential N-glycosylation sites of VHDL are shaded in grey. The identified N-glycosylation site on one of the *S. littoralis* peptides is shaded in light blue.

Very high-density lipoprotein from *Helicoverpa zea*

(accession code: A8CMX8; nominal mass: 173069; calculated pI value: 5.57)

1	MKLLILITAI	AAVASSPLTQ	NKSWPWQVGK	EYLYDVNTYT	VASFDGSNSN
51	GNAFKAQFAV	RVVAPGYLLA	KLTVSVYAQV	QDEKVTFSVL	PSDLKYQPVQ
101	ISDQPFEIFV	KGGRVKAFSV	PKTLSVAHEN	LLKGLISALQ	VDLSPYGHVD
151	NFPNSYDKES	FQGIFKKVEA	DLSGECETLY	TVSPVSAEWR	REFPNFASEE
201	APIEVIKSKN	YGSCCKRSAL	YSGVPEGSIW	NGIAYENDEQ	QLIKQTSEAR
251	IVASKQGTYI	KSEVLNSVFF	SPLLYGKQKA	QVTSYVKFYL	ASVQDSGAEW
301	KKPEGGRSID	SLIFTMSESM	FLPKVSDQSI	ANSQKLLQEM	APLLQSPSQL
				LLQEM	TPLLQDPNLL
351	PKADFLSKFN	ILIRHLATFN	SEQLTVLSNS	VEIARNSKNA	AKNAMWIIYR
	PK				
401	DALAQSGTVA	AFQQIKSWVL	AQKINGEEAA	ELISGIGAAL	RYPTKQVLTE
451	FFEFAINPEV	QQLKLVNTSA	LLAATRFSRN	VEDSLYVVES	IIPWLSKELK
501	QAAQNGDTVK	AQVYIRSLGN	LANPEILKVV	APYLDGSIIV	SKYLRVQIVA
551	SLKTLANLKD	ENVRAALYSI	LKNTAEPEYV	RVVAILNIFL	NEPTAEMMQI
601	MAHMTNDDPS	TQVRVAVLANG	INFAGGLKSP	RFSVLAKTAK	AVQPFVAKKE
651	FGYRISTDSL	IDEYTSDDDL	TFFRELAYVG	SEDNFFPVYH	RTALRTRGTN
	LSTDSL	LDEYTSDDDL	AYFRELSFLG	SENHLLPVYQ	R
701	LNEESQVTLT	VSGVQQIIEH	IADLIYEPNK	AKVDLKYSAK	KIAEKLNLKR
751	QKPDPIEGAL	FLENLNQQKL	ITFSEADLIA	FITSTVENIE	KLFNGVDVQY
	...DPLEGSL	FLENLNQQK			
801	TKILNNKLTY	IVFPLASGMP	FYYEYNEPLM	YSFNGQAKIK	VDRNSKSLAG
851	SLQKNIDFTY	ARNVDGSGVF	LDTLNDVYAF	AGVINKIQFY	IPANLNTVIT
901	PGQVKLSFKL	PEQDASFVHL	SVWPYTTLKK	VDSIQAISEL	PSTKIIKRAE
951	KVVSNDIKFG	YLAGVSLLFE	GYSYSSDFKS	TDLYDIDIVS	SIGTLFYPKD
1001	IALTEFELKY	NAKESTNKDV	TVTLFSDTLY	NQKQSGELAP	AALFKDVSAN
1051	SQARREEIVK	RAATGIESAQ	VELYDFSASF	EGKKKVEYVL	TAAIADSFVD
1101	NKFQGVFFLS	SWQEQINAVF	KVVVKPKIAPL	NFKEALKNQI	KVTYDLDLKV
1151	GNQKNIHLQG	YSERSETYTE	QLQNDPSAQK	CLEETNENNT	FQRDCYKLIV
1201	KAHAPDFFKA	SVTYEDKELT	PVALKSITEY	YDIFKSRNSY	TEEDVLRIT
					I
1251	ENGKLEIEAQ	FFYEENYANY	EFASQYGVLR	LNNVEGVAYY	PFGFAVYAPI
	LNGTLELVSE	AFYYDNYVNY	EFASK		
1301	TAVERSYNY	YGEQYTPYCV	VDSSKIQTFS	GRSYDYGLSG	SWHVVMVDES
1351	NDYGKWTDLV	ILARRPSEKQ	EEVYVSIVTE	DGKYLELNK	PAGIDVKTN
1401	KKVSEGAITS	YWNDEAEAPL	LQYYTLAEGV	LVFNKNGAI	RIVYDQQLA
1451	IFTDEHRKTT	RGVCGQDSTQ	TRDDYITPYG	IVDSPQLYGA	SFALDGENSD
1501	PKTEALKKEA	QQKAYQPFVK	YTNILSSDSE	WSKAKNHS	

Figure S2. Hexamerin proteins identified in the course of purification of the biliprotein from larval hemolymph of *Spodoptera littoralis*.

The ammonium sulphate precipitate of the biliprotein was subjected to DEAE gradient chromatography and fractions were analysed by SDS-PAGE. Cleavage peptides of proteins of ~85 kDa were obtained by in-gel tryptic digestion and sequenced by ESI-MS/MS. The obtained peptide sequences are shown in red matching the sequences of the proteins below, identified by homology search. The expect values were E-13 for protein 1; E-3 for protein 2; E-17 for protein 3, and E-11 for protein 4.

1. Methionine-rich storage protein from *Spodoptera litura*

(accession code: Q9U5K5; nominal mass: 89,419; calculated pI value: 9.12)

```

1   MRVLLVLA AV ALAAANVVKD DTNVFIGKDN MVNFDIKMKE LCIMKLLNHI
51  LQPTIYDDVR EVAREWVIED NVDKYLKAEV VRDFINTFRM GMLPRGEV FV
101 HTNELHIIQA VKVFKILYFA KDFDIFMR TA CWLRE RINGG MFVYALTACV
151 FHRADCRGIT RPAPYEIYPY FFVD SHVINK AFMMKMTKAA NDPVLMNYYG
201 IKVTDKNLVV IDWRKGVRRT LSQNDRI SYF TEDIDLNTYL YYLHMSYPYW
251 MNHDMYAVNK ERRGEIMSYA TMQLLARLRL ERLSHEMCDI KPIMWNEPLK
301 TGYWPKIRLH TGDEMPVRSN NMILVTKENL KWKRM LDDVE RKL RDGILNG
351 KIERRDGTVI SLKKPEDIED LARMVLGGMG LVSDDAKFMH MMHLMKRLLS
401 YNVYNFDKYT YVPTALDMYT TCLRDPVFWR LMKRVTDV FV LFKKMLPKYT
451 RDDDFDFPGVK IDRFTTDKLV TFMDEYDLDI TNALYLD DAE MKK KRPDMLM
501 VARMRRLNHH PFKVTVDVTT DKTVD AVVRI FIGPKYDCLG RLMSVNDKRL
551 DMVEIDTFKY KLETGKNTIV RNSVEMHGVI EQRPWTRRIL NNM RD TVGMI
601 SKTV DVESWW YKTRVGYPHR LLLPLGRLGG FPLQMYV IIS PVK TNNLLLP T
651 VDMSLMKERR TCRWSVCFDT MPLGFPPFDRK IDMTKFF THN MK FTDVMVYR
701 KDLGLSNNIK DVDMSDMVMK KDDYTYLDTD MLTRWSYKDV MMSKDDMMR
751 M

```

2. Moderately methionine-rich storage protein from *Spodoptera litura*

(accession code: Q9U5K6; nominal mass: 90,743; calculated pI value: 8.96)

```

1   MRAVLLLVVC LAGLAMARRD TDDYNTNNFV TMDIKQRQLV ILKLLNHVME
51  PLMYKDLEDF GK NYKFEDNF DRYTKTDVVK DFLKMLKVGF LPRGEIFTLH
101 IDRQLKEVVN MFHMLYYAKD FDTFVRTACW MRLYLNEG MF VYALTVAVRH
151 REDCKGIILP PPYEIYPYF VRGDVIQKAY LLKMKKG LLD LKLCDFYGIK
201 KTDKDIYIID ENVYDKRVYL NHDDKLR YFT EDIDLNTYYY YFHVDY PFWM
251 KDRFMDKTKT RRFELTVYMY QQILARYYLE RLSNGLGRIK DLSWYKPIKK
301 GYWPWLKLHN GIEFPVRFNN YVIARDYNLD VIRLCDEYEK IIREAIKGY
351 IEINGLRLEL TKTDDMEVLG KLIYGNIEKV DVDKTYVDSY RYLLIVMKAA
401 IGLNTFYSDK YFVVPTILDT YQTALRDPVF YMLQKRIIDL VCLYKLRLPS
451 YTKEDLYFPG VKIDNVNVDK LVTYFDDYLM DMTNAVYLTE DELKKTRSDM
501 KFFVRKRRLN HQPFKVTFDI LSDKSVDCVV RVFLGPKKDH MGR LIDININ
551 RLNFVELDTF LYKLTTGKNT IVRNSYDMHN LVRDRIMTRD LWKKVESITD
601 MRDFFIKDLR NYHTGFPTRL LLPKGFVGGM HMMLYVIVTP LKLV DNVDLN
651 VLDVNRKDFV YDFRSTVLLD KMPLGFPPDR RIDLGNFFTP NMKFVDVTIF
701 HKKQTCDMKT RWNRYVLKNY NLVDRTPIES DTYFVD TDLN MKVDRNRNLF
751 DI

```

3. Arylphorin subunit from *Spodoptera litura*

(accession code: Q9U5K4; nominal mass: 84,059; calculated pI value: 6.70)

```
1   MKTVLLLAGL VALVMGSAVP HHHDIILLKTV EPKFVEYQKK ILHLFENSEQ
51  LDYHSDYYKV GKDYDIEANI QNYSNKQAVE EFLLLYRTGT LPKYYKFSIF
101 YERMRDEAIA LFHVLYYAKD FETFYKTAAW AKVYMNEEQF FYAYYIAIVQ
151 RADTDGIVLP APYEVYPQFF FNSEIMHKLY RTKMQNNVFD EKLAQYGMV
201 KDNNNYVFYA NYSNSLSYPN KEQKLSYFTE DIGLNSYYFY FHSQMPFWWK
251 SEKLNILKDR LGEVFFYYQ QLLARYYLER LPHGLGEIPE FSWYSKFKSG
301 YYPQLPAHYI NYVQRSNDYN LHNEKNYEYI RFLDTYEKTF FQFLQKAEFK
351 SPEKEMNYVG NYWHMNSDLY AEQSNKDLHQ YSYEIIARHV LGASPKPFDK
401 YTFMPSALDF YQTSRLDPAF YQLYQRIIDY LIDYKQYVKP YNHNDLHFGV
451 VKINDVEVSE LVTYFEYFDF NVTNSAYYNK EELKAYPVNY IVRQPRINHK
501 PFNVKINVKS DVASDAVFKI FIGPKYHANG YPVNIEDDWM KFYELDWFVQ
551 KLVPGENKIE RKSSEFVLFK DDSVSINEIY KWLEQGGKVPY DMSVQPDNMP
601 RRLMLPRGSL GGPFQFLFVY VYPYNSVKKG ENVFENYILD DKPFGYPFDR
651 PVREAYFRQP NVFSKDVIY HKEPFFLYEL TIRSQFLQKK NKKPLNSF
```

4. Hexamerin from *Spodoptera exigua*

(accession code: A6YRR6; nominal mass: 82,178; calculated pI value: 6.48)

```
1   MARLALCVLA LLVAGALSSP DRSNVSPKVA DPVFLQRQVD LMILFFHLHE
51  PNHPESWKQI SKSWSLEKNI EHFSNVTAVT TYIEMLEHKW VLPRAVPFSL
101 LESGHKFEAV TLYNVLYSAK DYDSFYKTAV YLRERVNEHL FAYVLGVAIV
151 NRPDTQGIYI PR LHEVFPSY YNNGEILT TA QRINTHGHRL VESYPSTYKW
      LHEVFPSY YFNGELLT TA QR
201 DHNVVIRWNA TIWPYSNDHS TPVAYFTNDF GLNTYYNYNH LAQPSWLHGE
251 VLPLHKGRRG EWFVHVHKQI VARYMERLS NGLGEVPELG HNIVKEGYNP
301 GLLCHNGIAF PVRPDYYHLD RPVLVHALQK IDDYERRVRD AIEQGYVLNH
      D ALEQGYVLNH
351 LGERIDISSP EAIEILGSKI EANDSPNLK YYNDFISLWK KVLGNSIVHE
      R YYNDFLGLWK
401 HQYHHHVPL VVPSVPEHYQ TALRDPAFYM IWKRVLGLFQ LWQEKLAPYK
451 REQLAVPQVA IEKVDVDKLV SFFEYNYLNI SSFLHMNEEE AKELYDQVSV
501 LVQHPRLNHK KYQVRVHVKS EVAKTVLVKF FLAPKYDSHG YEIPLHVN SH
551 NFLQIDEFVH DLPAGESVIA RDSVDTSKVW DTANNVYYAF EKSLQGDQKF
      ...DEFVY DLPAGESVLA R NV DSANNVYNAF EK
601 NMEQLENMES LVQHLTLPKG RVGGMPFVLM VYISDYRAPK VPREPTNVSK
651 VSLGLTIAAR QLTDDPLGFP VNRPLYPWQV EGVKNIYFQD VLIHHKHTPE
701 IVIPHRE
```

The obtained tryptic peptides from *S. littoralis* are shown in red; they have very high similarity to the sequences, marked in yellow, of this hexamerin from *S. exigua*. Residues marked in green have probably been interchanged.

Figure S3. Vitellogenin identified in the course of purification of the biliprotein from eggs of *S. littoralis*.

The ammonium sulphate precipitate of the biliprotein was subjected to DEAE gradient chromatography and fractions were analysed by SDS-PAGE. Cleavage peptides of proteins of >200 kDa, ~180 kDa, ~150 kDa and ~37 kDa were obtained by in-gel tryptic digestion and sequenced by MALDI-TOF-MS. The obtained peptide sequences are shown in red matching the sequence of vitellogenin from *Spodoptera litura*, identified by homology search. The expect value was E-20.

Vitellogenin from *S. litura*

(accession code: A7XS62; nominal mass: 198,752; calculated pI value: 8.75)

Vitellogenin precursor (>200 kDa)

1	MKLLVLAIFI	AVVSSGNLSE	PQQNQWPWQT	GKLYRYDVNT	HILARHQEGA
51	SSGTAFKGEF	IVRVKSPGRL	QAKLEKPQHA	QIHEQLPNDM	AMPKDLK YET
101	VQNLDQIFEI	SVEGGRVLSV	NLPTTLLLSH	ENLLKGLIST	LQVDLSTHSS
151	TNNREDYLDL	EREQGLFKKM	ETDVTGNCET	MYTVVPVAAE	WRREL PQFAS
201	EEDPMEITKS	KNYGHCHHRV	AYHFGVPEGA	EWTGTAHSNE	EKQFISHSAV
251	SRLLVGKQGP	IYKAETTSTV	SVHPLIYGKQ	KAQVHSR VQF	TLLSVEQDNA
301	PEWHSPQSTR	QINTLLYSVT	TKQMAILDKT	SNLLQSSESH	EHLNNAETQR
351	ENTLNEDVSR	SSSSDSL SAY	VNEDVPNMNE	PAYAALYMSV	QSRGDKKQNA
401	MNVQKLLQDM	AQQLQNPNNM	PKADFLSKFN	ILVR IIASMS	TEQLAQISR
451	IEVGKSSNNN	VKADMWMIFR	DAVVQAGTPP	AFAQIKTWIK	NK KLQNEEAA
501	QVVSSLARTI	RYPskeIMTQ	SFDLAMSPEV	QQQKGLNTSA	LIAATRLIHM
551	AQVNNETAHN	YYPTRMYGRL	TNQHDTFVLE	IILPRLAEKM	NQAIEQQEWS
601	KAQVYIK AIG	NLGHPEILQV	FSPYLEGRIQ	VPRFIR VQMV	IQLRSLAKQH
651	DKHVKAFLFS	ILK NTAEPIE	VRVAATLNIF	LAHPTVAMMQ	DMAQMTNDDP
701	SVHVSALKS	GIVSAANLKD	PRFWHLSKTA	QAVK EQLTQE	NFGWRSSVKQ
751	FVDNYVKEDE	QEYFRETSYI	SSDNHGMPKF	QAHTWRISKIS	GWALENTISA
801	SLSDAKAILN	FIK QIMYEPL	KSNANHKHTA	QK ISEMLNIR	SEPQDPLQGA
851	FFYTILGQER	FFSFDENDLI	TLVQDIMEHM	KQVENGMEDH	YTK VFNSNQV
901	SVMFPISSGM	PFIYKYKEPV	AIHVQAKAKG	KVVRDPNTRK	EMSLMDKEL
951	LITAARNIDG	NVGFMDTLSN	KLASAGVVKK	YQLNVPVKLN	VQISSGEAKL
1001	KVEPLRPDQD	HTIAHYSVWP	YTTNQMKDTL	VPYSQDAATK	IVERPRKVSS
1051	TDVK FGQQVG	ALFQLQGYSH	SNDFRNTNVL	QVVSNNMANLL	ALRDIGLTHY
1101	NLKYL GKQSQ	NK KLTLTAAY	DEL FNQKQNG	ELKEARN MQD	VTPNSEARRA
1151	EMVKRVSSGI	NSAR AQVL DV	SATFEGSQKQ	EYVFTA AVAN	SPVDRKMQMV
1201	WFAGR NSAQQ	GNEQINVVVK	VKKPEISPLN	FLEALKKDMK	MTYEADIKFG
1251	QNGNIHIQGT	TERTKTATEQ	LKNNPWAKLV	QQQIENGNOY	QAAAHRLIR
1301	AHAPDNMKAI	VTYKNLSPMN	LNWTSQAFHI	LKQWNR NIEI	NPTKKVGDGK
1351	LQVEVQGSYL	DNTLRVEMIS	PVGLVSVDNV	PLPRFTPEVV	SLYTPFSPFE
1401	RLGNYAGNGQ	YQPFCTIDDN	KVKTFGNRSI	NYELSR SWHL	VMQEESENENR
1451	GQWNEMVILA	RRPSQQEQEL	YISYITETGK	DLEIEIKPSQ	SKRPNVQVKT
1501	NSKKISEGDL	TVYWDDVQDE	PLLEYFTEAD	DVLMLNIRDG	RLRAMYDGQR
1551	LVLTTQDHRK	TNRGICQONS	GEARDDYQTP	AGLVDQPELY	GASWALSDES
1601	SDPKAEELKK	KAQEKAYQPI	PKYTAILRSD	EQWKMAVQER	EQRLSSONLY
1651	MTRSYQRKGR	QCQLQKQIQY	YNTDR EICIS	TTPLPACPSN	CRGVAFNVES
1701	AHVVCRSNND	EQFKTYRQQI	QQGQNPQLPQ	VSHRLRKVNF	RVPTSCKA

Vitellogenin, C-Terminal fragment(s) (~150 kDa and ~180 kDa, respectively)

Almost identical sets of cleavage peptides were obtained from the protein bands of ~150 kDa and ~180 kDa; one of these sets is shown.

1	MKLLVLAIFI	AVVSSGNLSE	PQONQWPWQT	GKLYRYDVNT	HILARHQEGA
51	SSGTAFKGEF	IVRVKSPGRL	QAKLEKPQHA	QIHEQLPNDM	AMPKDLKYET
101	VQNLDQIFEI	SVEGGRVLSV	NLPTTLLLSH	ENLLKGLIST	LQVDLSTHSS
151	TNNREDYLDL	EREQGLFKKM	ETDVTGNCET	MYTVVPVAAE	WRREL PQFAS
201	EEDPMEITKS	KNYGHCHHRV	AYHFGVPEGA	EWGTGTAHSNE	EKQFISHSAV
251	SRLLVGKQGP	IYKAETTSTV	SVHPLIYGKQ	KAQVHSRVQF	TLLSVEQDNA
301	PEWHSPQSTR	QINTLLYSVT	TKQMAILDKT	SNLLQSSESH	EHLNNAETQR
351	ENTLNEDVSR	SSSSDSL SAY	VNEDVPNMNE	PAYAALYMSV	QSRGDKKQNA
401	MNVQKLLQDM	AQQLQNPNNM	PKADFLSKFN	ILVRIIASMS	TEQLAQISRG
451	IEVGKSSNNN	VKADMWMIFR	DAVVQAGTPP	AFAQIKTWIK	NKKLQNEEAA
501	QVVSSLARTI	RYPskeIMTQ	SFDLAMSPEV	QQQKGLNTSA	LIAATRLIHM
551	AQVNNETAHN	YYPTRMYGRL	TNQHDTFVLE	IILPRLAEKM	NQAIEQQEWS
601	KAQVYIKAI	GLHPEILQV	FSPYLEGRIQ	VPRFIRVQMV	IQLRSLAKQH
651	DKHVKAFLFS	ILKNTAEPIE	VRVAATLNIF	LAHPTVAMMQ	DMAQMTNDDP
701	SVHVSALKS	GIVSAANLKD	PRFWHL SKTA	QAVKEQLTQE	NFGWRSSVKQ
751	FVDNYVKEDE	QEYFRETSYI	SSDNHGMPKF	QAHTWR SKIS	GWALENTISA
801	SLSDAKAILN	FIKQIMYEPL	KSNANHKHTA	QKISEMLNIR	SEPQDPLQGA
851	FFYTILGQER	FFSFDENDLI	TLVQDIMEHM	KQVENG MEDH	YTKVFNSNQV
901	SVMFPISSGM	PFIYKYKEPV	AIHVQAKAKG	KVVRDPNTRK	EMSL LMDKEL
951	LITAARNIDG	NVGFM DTLN	KLASAGVVKK	YQLNVPVKLN	VQISSGEAKL
1001	KVEPLRPDQD	HTIAHYSWVP	YTTNQMKDTL	VPYSQDAATK	IVERPRKVSS
1051	TDVKFGQQVG	ALFQLQGYSH	SNDFRNTNVL	QVVS NMANLL	ALR DIGLTHY
1101	NLKYL GKQSQ	NKKLTLTAAY	DEL FNQKQNG	ELKEARNMQD	VTPNSEARRA
1151	EMVKRVSSGI	NSARAQVL DV	SATFEGSQKQ	EYVFTAAVAN	SPVDRKMOMV
1201	WFAGRNSAQQ	GNEQINVVVK	VKKPEISPLN	FLEALKKDMK	MTYEADIKFG
1251	QNGNIHIQGT	TERTKTATEQ	LKNNPWAKLV	QQQIENGNOY	QAAAHRLIR
1301	AHAPDNMKAI	VTYKNLSPMN	LNWTSQAFHI	LKQWNRNIEI	NPTKKVGDGK
1351	LQVEVQGSYL	DNTLRVEMIS	PVGLVSVDNV	PLPRFTPEVV	SLYTPFSPFE
1401	RLGNYAGNQ	YQPFCTIDDN	KVKTFGNRSI	NYELSR SWHL	VMQEESENENR
1451	GQWNEMVILA	RRPSQQEQEL	YISYITETGK	DLEIEIKPSQ	SKRPNVQVKT
1501	NSKKISEGDL	TVYWDDVQDE	PLLEYFTEAD	DVLM LNIRDG	RLRAMYDQQR
1551	LVLTTQDHRK	TNRGICGQNS	GEARDDYQTP	AGLVDQPELY	GASWALSDES
1601	SDPKAEELKK	KAQEKAYQPI	PKYTAILRSD	EQWKMAVQER	EQRLSSONLY
1651	MTRSYQRKGR	QCQLQKQIQY	YNTDR EICIS	TTPLPACPSN	CRGVAFNVES
1701	AHVVCRSNND	EQFKTYRQQI	QQGQNPQLPQ	VSHRLRKVNF	RVPTSCKA

Vitellogenin, N-Terminal fragment (~37 kDa)

1	MKLLVLAIFI	AVVSSGNLSE	PQONQWPWQT	GKLYRYDVNT	HILARHQEGA
51	SSGTAFKGEF	IVRVKSPGRL	QAKLEKPQHA	QIHEQLPNDM	AMPKDLKYET
101	VQNLDQIFEI	SVEGGRVLSV	NLPTTLLLSH	ENLLKGLIST	LQVDLSTHSS
151	TNNREDYLDL	EREQGLFKKM	ETDVTGNCET	MYTVVPVAAE	WRREL PQFAS
201	EEDPMEITKS	KNYGHCHHRV	AYHFGVPEGA	EWGTGTAHSNE	EKQFISHSAV
251	SRLLVGKQGP	IYKAETTSTV	SVHPLIYGKQ	KAQVHSRVQF	TLLSVEQDNA
301	PEWHSPQSTR	QINTLLYSVT	TKQMAILDKT	SNLLQSSESH	EHLNNAETQR
351	ENTLNEDVSR	SSSSDSL SAY	VNEDVPNMNE	PAYAALYMSV	QSRGDKKQNA
401	MNVQKLLQDM	AQQLQNPNNM	PKADFLSKFN	ILVRIIASMS	TEQLAQISRG
451	IEVGKSSNNN	VKADMWMIFR	DAVVQAGTPP	AFAQIKTWIK	NKKLQNEEAA
501	QVVSSLARTI	RYPskeIMTQ	SFDLAMSPEV	QQQKGLNTSA	LIAATRLIHM
551	AQVNNETAHN	YYPTRMYGRL	TNQHDTFVLE	IILPRLAEKM	NQAIEQQEWS
601	KAQVYIKAI	GLHPEILQV	FSPYLEGRIQ	VPRFIRVQMV	IQLRSLAKQH
651	DKHVKAFLFS	ILKNTAEPIE	VRVAATLNIF	LAHPTVAMMQ	DMAQMTNDDP
701	SVHVSALKS	GIVSAANLKD	PRFWHL SKTA	QAVKEQLTQE	NFGWRSSVKQ
751	FVDNYVKEDE	QEYFRETSYI	SSDNHGMPKF	QAHTWR SKIS	GWALENTISA

801 SLSDAKAILN FIKQIMYEPL KSNANHKHTA QKISEMLNIR SEPQDPLQGA
851 FFYTILGQER FFSFDENDLI TLVQDIMEHM KQVENGMEDH YTKVFNSNQV
901 SVMFPISSGM PFIYKYKEPV AIHVQAKAKG KVVDPNTRK EMSLLMDKEL
951 LITAARNIDG NVGFMDTSLN KLASAGVVKK YQLNVPVKLN VQISSGEAKL
1001 KVEPLRPDQD HTIAHYSVWP YTTNQMKDTL VPYSQDAATK IVERPRKVSS
1051 TDVKFGQQVG ALFQLQGYSH SNDFRNTNVL QVVSNNMANLL ALRDIGLTHY
1101 NLKYLKGQSQ NKKLTTLTAAY DELFNQKQNG ELKEARNMQD VTPNSEARRA
1151 EMVKRVSSGI NSARAQVLDV SATFEGSQKQ EYVFTAAN SPVDRKMQMV
1201 WFAGRNSAQQ GNEQINVVVK VKKPEISPLN FLEALKKDMK MTYEADIKFG
1251 QNGNIHIQGT TERTKTATEQ LKNNPWAKLV QQQIENGNOY QAAAHRMLIR
1301 AHAPDNMKAI VTYKNLSPMN LNWTSAFHI LKQWNRNIEI NPTKKVGDGK
1351 LQVEVQGSYL DNTLRVEMIS PVGLVSVDNV PLPRFTPEVV SLYTPFSPFE
1401 RLGNYAGNGQ YQPFCTIDN KVKTFGNRSI NYELSRSWHL VMQESNENR
1451 GQWNEMVILA RRPSQQEQEL YISYITETGK DLEIEIKPSQ SKRPNVQVKT
1501 NSKKISEGDL TVYWDDVQDE PLLEYFTEAD DVLMLNIRDG RLRAMYDGQR
1551 LVLTTQDHRK TNRGICGQNS GEARDDYQTP AGLVDQPELY GASWALSDES
1601 SDPKAEELKK KAEKAYQPI PKYTAILRSD EQWKMAVQER EQRLSSQONLY
1651 MTRSYQRKGR QCQLQKQIQY YNTDREICIS TTPLPACPSN CRGVAFNVES
1701 AHVVCRSNNQ EQFKTYRQQI QQQQNPQLPQ VSHRLRKVNF RVPTSCKA