Molecular Characterization and Developmental Expression of a Retinoid- and Fatty Acid-binding Glycoprotein from Drosophila 

A PUTATIVE LIPOPHORIN* 

(Received for publication, April 22, 1996, and in revised form, June 12, 1996) 

R. Krishnan Kutty†§, Geetha Kutty†, Ravi Kambadur†, Todd Duncan†, Eugene V. Koonin**, Ignacio R. Rodriguez†, Ward F. Odenwald†, and Barbara Wigger† 

From the Laboratory of Retinal Cell and Molecular Biology, NEI; the Neurogenetics Unit, Laboratory of Neurochemistry, NINDS; and the *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, 20892 

A detailed understanding of the mechanism of lipid transport in insects has been hampered by the inability to identify the proapolipophorin gene that encodes apolipoporphins I and II, the principal protein components of lipophorin, the lipid transport vehicle. Here we provide the first molecular description of the Drosophila gene encoding a retinoid- and fatty acid-binding glycoprotein (RFABG) and present evidence that it is a member of the proapolipophorin gene family. The gene, localized to the chromosome 4 (102 F region), encodes a 3351-amino acid protein that could serve as the precursor for the 70-kDa and >200-kDa polypeptides associated with RFABG. The N-terminal sequence of the 70-kDa polypeptide and that predicted for the >200-kDa polypeptide showed high sequence similarity to blowfly apolipophorin II and apolipophorin I, respectively. The RFABG precursor contains a signal peptide and exhibits a multidomain mosaic protein structure, which is typical of extracellular proteins. It has structural domains similar to lipid-binding proteins, namely vitellogenins and apolipoprotein B. The protein also contains a domain similar to the D domain of von Willebrand factor and mucin. The gene is expressed in the Drosophila embryo during development in cells that make up the amnioserosa and fat bodies. Immunolocalizations using specific antibodies against RFABG reveal that the protein is initially dispersed through the embryonic amnioserosa sac and later concentrated at skeletal muscle-epidermis apodeme contact junctions during larval development. This novel gene may play an important role in the transport of lipids, including retinoids and fatty acids, in insects.

Lipophorin, the lipoprotein present in the hemolymph, is thought to be the transport vehicle for various types of lipids in insects (1-5). Two large apoproteins known as apolipophorin I (230-250 kDa) and apolipophorin II (70-85 kDa) are the major protein components of lipophorin (1-6). They are also implicated in the transport of hydrophobic ligands like juvenile hormones, pheromone hydrocarbons, and carotenoids (2, 7–9). The two apolipophorins are derived from a common precursor known as proapolipophorin (3, 10). The unusually large size of the proapolipophorin mRNA has hampered various efforts to clone and sequence it from several sources.

A novel retinoid- and fatty acid-binding glycoprotein (RFABG)1 was isolated in our laboratory from Drosophila melanogaster heads (11). It was capable of binding both retinoids and fatty acids with high affinities. RFABG was also found to contain endogenous fatty acids. Two major protein bands at >200 kDa and ~70 kDa, resembling apolipophorin I and apolipophorin II, respectively, were seen on SDS-PAGE of the purified RFABG. The ability to bind hydrophobic ligands, the presence of endogenous fatty acids, the subunit composition, and glycoprotein nature indicated that RFABG could indeed be related to lipophorin.

In order to further characterize RFABG, specific antibodies were raised against it in rabbits and used to screen a Drosophila head cDNA library. The cloned gene encodes a large protein (3351 amino acids) that contains a 25-amino acid signal peptide. The predicted sequence contains regions showing high similarity to the N-terminal sequences of blowfly apolipophorins I and II (7). It also showed domains similar to lipid-binding proteins, namely vitellogenin and apolipoprotein B, as well as a domain similar to the D domain of von Willebrand factor (vWF). The gene is expressed in the amnioserosa, fat body, and apodemes of the developing Drosophila embryo. 

EXPERIMENTAL PROCEDURES 

Preparation of Proteins for in Situ Digestion and Microsequencing 

Soluble proteins from the heads of Drosophila melanogaster were subjected to ConA-Sepharose affinity chromatography as described previously (11). The eluate containing RFABG was subjected to SDS-PAGE (approximately 50 μg of protein/lane) in 10-20% Tricine gels using a Novex electrophoresis system (Novex, San Diego, CA) (12). Proteins from the gel were then electroblotted for 3 h onto nitrocellulose membrane using a Novex transfer apparatus (13). Nitrocellulose membranes were then immersed for 60 s in a solution of 0.1% (v/v) Ponceau S dye in 1% acetic acid. Excess stain was removed from the blots by gentle agitation for 1–2 min in 1% acetic acid. Stained protein bands corresponding to >200 and ~70 kDa were excised and transferred to Eppendorf tubes (1.5 ml), washed with HPLC grade water for 1–2 min, and then stored wet at ~20 °C. A minor ~140-kDa band that appeared in certain preparations of RFABG was also excised and processed the same way. Sixteen to 18 such nitrocellulose pieces containing the same protein band were pooled in a single Eppendorf tube and stored at −70 °C until use. 

Microsequencing of the excised bands was performed as described previously (11). The procedure involved in situ digestion of the proteins and microsequencing of the released aminoterminal amino acids. 

References 

1 The abbreviation used is: RFABG, retinoid- and fatty acid-binding glycoprotein; PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction; VWF, von Willebrand factor; MAP, multiple antigenic peptide; PAGE, polyacrylamide gel electrophoresis; ConA, concanavalin A; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethy]glycine; kb, kilobase pair(s); bp, base pair(s); aa, amino acid(s).
tube and sent to the Harvard Microchemistry Facility and Cambridge Prochem, Cambridge, MA for microsequencing. In situ tryptic digestion and peptide sequence analysis was performed as described by Aebersold et al. (14).

**FIG. 1.** The predicted protein sequence of the Drosophila RFABG. The peptides determined by microsequencing are shown in bold type, those from the ~70-kDa, >200-kDa, and the ~140-kDa bands associated with the purified RFABG are underlined with single line, dotted line, and double line, respectively. The putative N-glycosylation sites are indicated by asterisks. The secondary structure prediction was shown as: H, α-helix; E, β-strand. The protein secondary structure prediction was by the neural network PHD method (21, 22).
rabbit antibodies against a peptide sequence obtained as described above from the 70-kDa protein band. The peptide, YIPGNYYDYSFD-SILTIGASSDVPNDSD, was synthesized as a four-branch MAP by Applied Biosystems, Foster City, CA. Rabbits were immunized subcutaneously with 1 mg of MAP in complete Freund’s adjuvant. Two booster injections of 1 mg of MAP in incomplete Freund’s adjuvant were given at 2-week intervals. Rabbits were bled 2 weeks after the third injection.

Antibodies against the 70-kDa Protein Band—The fraction (30–40 mg of protein) capable of binding retinol and palmitic acid obtained by Superose-6 chromatography of the ConA-Sepharose eluate was subjected to SDS-PAGE and transblotting as described above under microsequencing. The 70-kDa stained protein band on the nitrocellulose membrane was excised, and six such strips were dissolved in 0.5 ml of Me2SO. Rabbits were immunized subcutaneously with the dissolved band in incomplete Freund’s adjuvant. Two booster injections of six strips dissolved in Me2SO in incomplete Freund’s adjuvant were given at 2-week intervals, and the rabbits were bled 1 week after the third injection.
Drosophila Retinoid- and Fatty Acid-binding Glycoprotein

A

701 GGRAKK SIVDDV

B

RFABG/DROME  26 ENACLGCSPXSDNGLKYYPCN
APOLIPOPHORIN II  1 ESSTRGSPESDNGLYSPETN

C

RFABG/DROME  707 SIVVSVXSISSLKYYGTV
APOLIPOPHORIN I  1 SIAASSXSVQKXYYGK

Fig. 2. Possible mode of generation of ~70- and ~200-kDa subunits of Drosophila RFABG and their similarity to blowfly lipophorins I and II. A, region of RFABG sequence showing consensus sequence (RRXR) for proteolytic processing; * indicates the protease cleavage site (RR), region of RFABG sequence and the N-terminal sequence of blowfly apolipophorin II. Conserved amino acid residues are shown in the middle line. C, comparison of RFABG sequence and the N-terminal sequence of blowfly apolipophorin I.

Antibody against Native RFABG—The fraction obtained by ConA-Sepharose affinity chromatography of soluble proteins from Drosophila heads was subjected to Superose-6 size exclusion chromatography. The fraction identified as being able to bind exogenous radiolabeled retinol and palmitic acid (11) was rechromatographed on Superose-6, and the resultant protein peak binding retinol and palmitic acid was used for antibody production. Rabbits were immunized subcutaneously with 150 μg of this protein fraction in complete Freund's adjuvant. Two booster injections of 150 μg of this protein fraction in incomplete Freund's adjuvant were given at 1-month intervals. Rabbits were bled 2 weeks after the third injection.

Western Blot Analysis

Drosophila extracts were subjected to SDS-polyacrylamide gel electrophoresis using 10–20% Tricine gels (Novex, Encinatas, CA) and the protein bands were transferred to a nitrocellulose membrane (Schleicher & Schuell) (12, 13). The membrane was blocked with a 5% solution of Carnation dried milk in Tris-buffered saline and then incubated at 0–4°C overnight with antigen diluted 200-fold. The blot was washed with Tris-buffered saline and incubated with an alkaline phosphate-conjugated goat F(ab’)2 fragment against rabbit IgG (Cappel, Durham, NC). The blot was washed again, and the immunoreactive bands were visualized with 4-chloro-1-naphthol and H2O2 as substrates.

Poly(A)+ RNA preps from Drosophila were subjected to agarose gel electrophoresis in the presence of formaldehyde. The RNA from the gel was capillary blotted onto an Immobilon N membrane (Millipore Corp., Bedford, MA), UV cross-linked and hybridized as described by Sambrook et al. (16) with a 1.7-kb EcoRI/HindIII cDNA fragment. The probe was labeled with 32P by random priming (17). The hybridization was carried out at 65°C for 16 h with the labeled probe (2 x 106 cpm/ml) using Hybrisol II ( Oncor, Gaithersburg, MD) containing 0.5 mg/ml denatured salmon sperm DNA. The blots were then washed twice in 2 x SSC and 0.1% SDS for 30 min at room temperature, once in 0.1 x SSC and 0.1% SDS for 10 min at 65°C, and exposed to Kodak XAR film.

cDNA Cloning

An adult Drosophila head library in a λExilo vector (Novagen, Madison, WI) was plated using BL26(DE3)pLyse strain. The plates were incubated with nylon filters impregnated with isopropyl-β-D-thiogalactopyranoside. The filters were subjected to microwave screening (In Vitrogen Corp., San Diego, CA) using the antibody preparation against native RFABG. Positive clones were isolated and purified by three rounds of rescreening. The positive phage clones were converted to plasmid subclones by Cre-loxP autosubcloning.

cDNA Sequencing

The DNA preparation from the P1 clone was reverse-transcribed with an oligod(T) primer. The first strand cDNA preparation was used as a template for PCR (19), which was performed with a GeneAmp PCR system 9600 using the XL PCR Kit (Perkin Elmer). Several sets of primers designed to span various regions of the genomic sequence were employed. Hot start PCR conditions were used under the conditions described by the manufacturer. Samples were subjected to thermocycling: 1 min at 94°C; 16 cycles of 30 s at 94°C and 10 min at 62°C; 20 cycles of 30 s at 94°C and 10 min at 62°C with an increment of 1°C cycle; 72°C for 10 min. The samples were then analyzed by electrophoresis on a 1% agarose gel, and amplification products were visualized by ethidium bromide staining. The bands were excised and extracted from the gel with a Qiagen Gel Extraction Kit (Qiagen, Chatsworth, CA) or a Wizard PCR Preps DNA Purification System (Pierce, Madison, WI) for direct DNA sequencing.

DNA Sequencing

Sequencing of cDNA clones and PCR fragments were done on an Applied Biosystem 370A DNA sequencer using Taq dyeoxy termination sequencing kit (Applied Biosystems). P1 clone and PCR fragments were sequenced manually using Amplitop sequencing kit (Perkin Elmer) and Sequagel-6 (National Diagnostics, Atlanta, GA).

Computer Analysis of the Sequence

Protein sequences were from the non-redundant (NR) data base supported at the National Center for Biotechnology Information (National Institutes of Health, Bethesda, MD). The analysis of protein physicochemical properties was performed using programs available through the ExPasy Worldwide Web server at the University of Geneva (http://expsy.hugich.egw.html). Specifically, the protein molecular weight and pl were computed with the ProtParam program. The signal peptide prediction was performed with the Signalp program (20). The protein secondary structure prediction was by the neural network PHD method (21, 22). Prediction of transmembrane helices in proteins was by the PHDhtm program (23). Protein sequences were compared to the NR data base with the BLAST program, version 1.4 (24). Low complexity regions that tend to produce spurious “hits” in data base searches were masked in the query sequences with the SEG program (25). Multiple sequence alignment programs were constructed with MACAW program (26).

Chromosome Localization

The DNA preparation from the P1 clone was random prime labeled (27) with biotin-dCTP. Polytene chromosomes were hybridized with this probe and signals were detected with peroxidase (28). Chromosomes were stained with Giemsa.

In Situ Hybridization and Immunohistochemistry

The techniques employed for in situ hybridization and immunochemistry of Drosophila embryos were as described before (29). Whole-mount embryos in situ hybridization was performed according to Tautz and Pfeifle (30) using a 1.7-kb EcoRI/HindIII cDNA fragment. The probe was labeled with digoxigenin by random priming using the Genius kit (Boehringer Mannheim). For immunostaining, the embryos were dechorionated and fixed as described earlier (29). They were incubated with primary antibody (antiserum raised against the ~70-kDa protein band) diluted in phosphate-buffered saline containing 0.1% Tween 20 and 4% horse serum at room temperature for 1–2 h. The embryos were then treated with secondary antibody conjugated to biotin (Vector Labs), followed by streptavidin-horseradish peroxidase. The color was developed using diamobenzidine as the horseradish peroxidase substrate described by Kania et al. (31).
RESULTS

Purification and Characterization of Retinoid- and Fatty Acid-binding Glycoprotein (RFABG)—The purification and characterization of RFABG has been described before from this laboratory (11). Briefly, the glycoprotein was purified from the soluble protein fraction of Drosophila heads using ConA-Sepharose affinity chromatography followed by Superose-6 chromatography. As expected, the purified protein fraction showed two protein bands with molecular masses of 70 kDa and 200 kDa when subjected to SDS-PAGE. Fluorescence titration analysis showed that RFABG binds retinol with a $K_d$ of $2.9 \times 10^{-7}$ and 16-$[9$-anthroyloxy]$palmitic$ acid$ with$ a$ $K_d$ of $3.5 \times 10^{-7}$. Antibodies were raised against RFABG in order to further characterize it by cDNA cloning.

Cloning of the Gene for the Retinoid- and Fatty Acid-binding Glycoprotein—A Drosophila head cDNA library was screened using antibodies raised against native RFABG. Although several clones were obtained, all of them had the same 2.9-kb insert. Screening of a Drosophila genomic library (18) constructed using a bacteriophage P1 vector was then employed to screen by PCR a Drosophila genomic library (18) constructed using a bacteriophage P1 vector. A clone isolated was sequenced by primer walking. The genomic sequence thus obtained was used to design several sets of primers to generate amplification products overlapping the entire sequence. They were used for RT-PCR using poly(A) RNA preparations isolated from Drosophila. The amplification products obtained were then sequenced. The comparison of genomic sequence with the combined sequence obtained from RT-PCR products enabled us to identify the intron and exon regions.

The cDNA sequence thus obtained showed a large open reading frame with two potential initiation codons at positions 136 and 145; the upstream codon was in the more favorable context and was likely to be utilized predominantly if not exclusively (32). The stop codon TAA was found at position 10189 bp.

Sequence Analysis of the Retinoid- and Fatty acid-binding Glycoprotein— FIG. 3. Structural features of Drosophila RFABG. A, a scheme for the domain organization of RFABG; α indicates the predicted α-helical domains, the arrow the cleaved signal peptide, and the bold arrow the possible proteolytic cleavage site for the generation of the ~70-kDa and the ~200-kDa polypeptides. B, conservation of the D domain in RFABG, von Willebrand factor, and mucin. The consensus shows amino acid residues that are conserved in RFABG and at least six of the eight aligned D domain sequences are indicated by asterisks. The distances from protein termini and the distances between the aligned segments are indicated by numbers. The predicted secondary structure is shown as: h, α-helix; b, β-strand; l, loop; ?, no prediction for the given position.

Drosophila Retinoid- and Fatty Acid-binding Glycoprotein 20645
Glycoprotein—The large cDNA sequence obtained for RFABG encodes a putative protein that consists of 3351 amino acid residues, with a predicted M, of 372,690 and theoretical pI of 8.07 (Fig. 1). The putative protein is predicted to contain a cleaved signal peptide consisting of 25 amino acid residues. The protein is relatively hydrophilic, and even though several stretches of hydrophobic amino acid residues are present, none of them is strongly predicted to form a transmembrane helix. Therefore, it is most likely that RFABG is either an extracellular or a soluble intracellular protein, rather than an integral membrane protein. However, the presence of a cleaved signal peptide in the predicted sequence of RFABG is in agreement with the extracellular nature of this protein. The protein contains 35 cysteine residues, some of which are likely to form disulfide bonds. Seven potential N-glycosylation sites were predicted.

The protein sequence deduced from the cDNA sequence was verified by analyzing peptides obtained by in situ digestion and microsequencing of the >200- and ~70-kDa bands present in the ConA-Sepharose-purified RFABG. A minor band at ~140 kDa that appears in certain preparations of RFABG was also analyzed by microsequencing. The sequences of 21 peptides thus obtained were found to account for nearly 10% (324 amino acid residues) of the deduced sequence (Fig. 1). Five peptides obtained from the ~70-kDa band spanned the N-terminal region (aa 26–615). The regions spanned by seven peptides obtained from the >200-kDa band (aa 1063–2132) and nine peptides obtained from the ~140-kDa band (aa 1788–2936) did overlap and were close to the C-terminal region. It appears that the ~140-kDa band could be an artifact of purification possibly generated from the proteolysis of the >200-kDa band. The fact that the former band is not detected when a crude extract of Drosophila is analyzed by Western immunoblotting using antibody raised against native RFABG is in support of this conclusion. The N-terminal sequence of the ~70-kDa polypeptide was determined and was found to span the aa 26–50 region of the sequence shown in Fig. 1. Thus, the first 25 amino acids marked as a signal peptide in Fig. 1 are indeed absent in the purified protein. The N terminus of the >200 kDa polypeptide was blocked; therefore, the sequence could not be determined.

Fig. 2 shows that the predicted sequence of RFABG contains a single consensus cleavage sequence RXRR for dibasic processing endoproteinases of the subtilisin family similar to that found at subunit cleavage sites of vitellogenins from mosquito and boll weevil (33–35). It appears that such a protease can generate a 681-amino acid polypeptide (~70 kDa) spanning the N-terminal region, and a 2645-amino acid peptide (>200 kDa) spanning the C-terminal region can be generated from RFABG lacking the signal peptide. As shown in Fig. 2, a comparison of the RFABG sequence with those in the non-redundant sequence data base (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD) showed 70% identity between the N-terminal sequence of the ~70-kDa subunit of RFABG and that reported by Trowell et al. (7) for blowfly apolipoporphin II. Also, the deduced N-terminal sequence of the ~200 kDa subunit of RFABG showed 54% identity to that of blowfly apolipophorin I.

A comparison of the RFABG protein sequence with the non-redundant sequence data base also showed statistically highly significant similarities between two distinct domains of RFABG and two groups of extracellular proteins (Fig. 3A). An extended region located, approximately, between the residues 150 and 900 of RFABG is similar to lipid-binding proteins, namely lipovitellins (36) and apolipoprotein B (37). The probability of the RFABG sequence matching the lipovitellin (or vitellogenin) and apolipoprotein B sequences by chance was below 10⁻⁶.

A region approximately 150 residues long, located between positions 2790 and 2940 in the RFABG sequence, is related to the D domain of human prepro-vWF, with the probability of matching by chance below 10⁻⁵; a somewhat lower similarity to the related D domain of intestinal mucin was also observed (Fig. 3B). Both vWF and mucin contain four copies of the D domain (38, 39), but only one copy was found in RFABG, with no additional, divergent copies detectable. The highest similarity was observed between the single D domain of RFABG and the N-terminal D1 domain of vWF, which belongs to the "prepro" sequence. Of the 6 cysteine residues that are conserved in vWF and mucin, only 2 are found also in RFABG; thus the D domain of RFABG may contain only one disulfide bond, in contrast to the larger number of such bonds in its homologs (Fig. 3B).

Chromosome Localization and Expression of the Retinoid and Fatty Acid-binding Glycoprotein Gene—The RFABG gene was localized to the Drosophila polytene chromosome 4 to the 102 F region by in situ hybridization of polytene chromosomes (data not shown). A functional analysis of this gene was not possible since mutants were not reported for this particular locus. Northern blot analysis showed that the gene is expressed as a ~10-kb mRNA (Fig. 4).

The expression of RFABG was investigated using several antibody preparations raised against it. Purified preparations of the RFABG when subjected to SDS-PAGE showed major bands at the >200-kDa and ~70-kDa regions (11). Antibodies raised against a peptide sequence derived from the ~70-kDa band of RFABG, as expected, reacted with a ~70-kDa band when extracts from Drosophila head or body and the purified protein fraction were subjected to Western blot analysis (Fig. 5A). Similar results were also obtained with antibody prepa-
tions raised against the entire ~70-kDa band (Fig. 5B). As expected, the antibody preparation raised against purified native RFABG showed a high molecular immunoreactive band at >200 kDa in addition to the ~70-kDa band for both head and body extracts as well as for the purified protein (Fig. 5C). A band that migrates slightly faster than the 200-kDa band was also observed for the purified protein (Fig. 5C, lane 4), but not for crude extracts from Drosophila head or body (Fig. 5C, lanes 2 and 3). A possible cause for the generation of this band is proteolytic degradation occurring during the purification of the protein.

The expression of the RFABG in the developing Drosophila embryo is shown in Fig. 6. During stage 12, cells that make up the amnioserosa accumulate high levels of RFABG message (Fig. 6, A and B). Protein immunostaining (Fig. 6, C and D) at this time in development reveal that RFABG is dispersed throughout the yolk sac. By late stage 14, the lateral fat body cells express RFABG (Fig. 6, E). Starting at stage 14, the RFABG protein localizes to the apodemes (Fig. 6, F and G; arrows in G indicate the lateral muscle fibers that underlie the RFABG-immunopositive apodemes).

**DISCUSSION**

The gene encoding the retinoid- and fatty acid-binding glycoprotein (RFABG) has been cloned and localized to Drosophila chromosome 4 (102 F region). It is expressed as a very large
mRNA (>10 kb). The protein sequence deduced from the cDNA sequence shows that RFABG is produced as a precursor that contains 3351 amino acid residues (Fig. 1). The first 25 amino acid residues of RFABG are predicted to constitute a signal peptide. This conclusion is supported by the data from the N-terminal analysis of the purified RFABG preparation. Microsequencing analysis of the peptides obtained from the purified RFABG indicates that the precursor protein may be proteolytically processed to yield >200- and ~70-kDa polypeptides. Indeed, a mechanism involving dibasic processing endopeptidase cleavage site proposed for the proteolytic processing of mosquito and boll weevil vitellogenins could generate these subunits (Fig. 2; Refs. 34 and 35). The N-terminal sequence of ~70-kDa subunit showed high similarity to that reported for blowfly apolipopophorin II (7). Also, the predicted N-terminal sequence of the >200 kDa subunit showed high similarity to that obtained for blowfly apolipopophorin I (7). Van der Horst et al. (3, 10) have reported that apolipopophorins I and II in locust are derived from a common precursor called proapolipophorin. Thus, it appears that the RFABG gene could belong to the hitherto uncharacterized proapolipophorine family. However, the sequence similarity between RFABG and the N-terminal of apolipopophorin I or cockroach lipophorin was not striking (3, 8). Interestingly, comparison of limited sequence information available for blowfly, locust, and cockroach lipophorins I and II showed high similarity, indicating that the proapolipophorine gene may not be highly conserved during evolution. However, it should be noted that more sequence information is needed from these insects for a more meaningful comparison. The partial sequence (801 aa) of apolipopophorin I (3) showed 25% identity with 49% similarity to RFABG sequence. Like the RFABG sequence, the locust apolipopophorine I (3) also contained regions similar to apoRFABG sequence. Like the RFABG sequence, the locust apolipophorin I (3) showed 25% identity with 49% similarity to apoRFABG sequence. Like the RFABG sequence, the locust apolipophorin I (3) also contained regions similar to apoRFABG sequence.

RFABG is a multidomain, mosaic protein, which is typical of extracellular proteins (40, 41). The lipovitellin-related domain of RFABG is predicted to consist of two β-sheets flanking an approximately 300-residue α-helical domain. This is reminiscent of the structure of the lipovitellin lipid-binding cavity (42, 43) and suggests that in RFABG, this domain is directly involved in retinol and fatty acid binding. The role of the WVF-related D domain in RFABG is less clear. Interestingly, vitellogenin contains a C-terminal domain that is distantly related to the D domain of von Willebrand factor (36). Even though the similarity between this domain and the D domain-related region in RFABG was much lower than between the latter and WVF, and not statistically significant (data not shown), these observations indicate that RFABG and vitellogenins may have a principally similar domain organization. The D1 domain of WVF is involved in multimerization of this protein (44, 45), and by inference multimerization mediated by the D domain may be postulated for RFABG. In fact, lipophorin has been implicated in hemolymph clotting reactions in certain insects (46, 47), and the multimerization of RFABG might be important for this process.

The physiological significance of the retinoid binding activity of this novel glycoprotein remains to be elucidated. Drosophila RFABG exhibits properties similar to that of bovine interphotoreceptor retinoid-binding protein (11). Both proteins are glycosylated, contain covalently and noncovalently bound endogenous fatty acids, and show similar binding affinity toward retinoids and fatty acids. However, the predicted sequence of the RFABG shows no detectable similarity to the bovine interphotoreceptor retinoid-binding protein sequence (48). The ability of the Drosophila RFABG to bind lipids such as retinoids and fatty acids is reflected in its structural domain that is similar to that of lipovitellin, vitellogenin, and apolipoprotein B and appears to contain a lipid-binding cavity analogous to that of lipovitellin. A major portion of the region resembling the lipid-binding domain of lipovitellin and apolipoprotein B resides in the ~70-kDa band of RFABG. The newly characterized protein may play an important role in transport of lipids in Drosophila. The ability to bind retinoids may stem from the ability of RFABG to transport carotenoids assimilated from dietary sources to the visual system where it can be processed to visual pigments. In fact lipophorin is thought to transport carotenoids (2).

RFABG appears to play an important role in Drosophila development. It is expressed in amnioserosa, fat body, and amnioserosa. It is interesting to note that the lipophorins or their precursor, proapolipophorin, are shown to be produced in the fat body in other insects (1–4, 10). The striped appearance of staining is interestingly similar to tiggrin, a protein found in amnioserosa, the extracellular matrix at muscle insertions (49).

In summary we have identified a gene in Drosophila that encodes a retinoid- and fatty acid-binding glycoprotein. Molecular analysis shows that this protein may belong to the apolipophorin family. It exhibits the properties of an extracellular protein with structural features characteristic of lipid binding proteins such as vitellogenin and apolipoprotein B. The expression of this gene is developmentally regulated. The cloning of this gene from Drosophila opens up opportunities for genetic intervention to study the mechanism of transport of lipids and hydrophobic molecules such as retinoids, hydrocarbons (pheromones), and juvenile hormones.

Acknowledgment—We thank Dr. Gerald J. Chader for support and interest in this study. We gratefully acknowledge Robert J. ones and Dr. Daniel Hartl for their help with chromosome localization. We also thank Dr. Fintan Steele, Dr. Susan Haynes, Dr. David Smoller, and Ted Fletcher for their help and advice. We are deeply indebted to Dr. Thomas Kaufmann for supplying us with the thousands of frozen flies necessary to obtain enough protein for antibody production and microsequencing.

REFERENCES

47. Gotz, M. (1983) Naturwissenschaften 70, 45–46
Molecular Characterization and Developmental Expression of a Retinoid- and Fatty Acid-binding Glycoprotein from Drosophila: A PUTATIVE LIPOPHORIN
R. Krishnan Kutty, Geetha Kutty, Ravi Kambadur, Todd Duncan, Eugene V. Koonin, Ignacio R. Rodriguez, Ward F. Odenwald and Barbara Wiggert

doi: 10.1074/jbc.271.34.20641

Access the most updated version of this article at http://www.jbc.org/content/271/34/20641

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 47 references, 12 of which can be accessed free at http://www.jbc.org/content/271/34/20641.full.html#ref-list-1