Insect Biochem. Molec. Biol. Vol. 24, No. 1, pp. 21–27, 1994
Printed in Great Britain.

0965-1748/94 $6.00 + 0.00

cDNA-inferred Amino-acid Sequence of a C Protein, a Heparin-binding, Basic Secretion Product of the Tubular Accessory Sex Glands of the Mealworm Beetle, *Tenebrio molitor**

GUIDO C. PAESEN,† GEORGE M. HAPP†‡

Received 7 July 1992; revised and accepted 20 April 1993

C proteins represent one of the four major protein groups secreted by the tubular accessory glands of male mealworm beetles (*Tenebrio molitor*). They are basic proteins with an apparent molecular mass of 21.9 kDa. In this paper we present the deduced amino-acid sequence of two, almost identical C proteins, termed C1 and C2. The C proteins contain a consensus sequence for a heparin-binding site, and they are efficiently purified from accessory gland homogenates by heparin-affinity chromatography. No sequence resemblance was found with other proteins in the databases, but their high avidity for heparin suggests a possible involvement of the C proteins in sperm capacitation.

C protein Accessory gland Amino acid sequence Heparin-binding Seminal fluid Sperm capacitation *Tenebrio molitor*

INTRODUCTION

A variety of functions is attributed to peptides and proteins secreted by the accessory sex glands of male insects. In many species, accessory gland polypeptides will contribute to the formation of a spermatophore, an acellular construction in which sperm and seminal fluid are packaged for transfer to the female. Other peptides and proteins become part of the seminal fluid and play a role in sperm storage, maturation or activation, or they act upon the physiology and behavior of the female after copulation (Leopold, 1976; Chen, 1991; Happ, 1992).

The reproductive apparatus of the male mealworm beetle (*Tenebrio molitor*) has two pairs of accessory sex glands, termed BAGs and TAGs. The BAGs, short for bean-shaped accessory glands, produce most, if not all, of the structural proteins that compose the spermatophore (Dailey *et al.*, 1980; Grimnes *et al.*, 1986; Paesen *et al.*, 1992a). The tubular accessory glands or TAGs, produce mainly four groups of proteins, which were termed A, B, C and D proteins, according to their mobility on SDS-gels (Happ *et al.*, 1977; Black *et al.*, 1982). These TAG proteins probably contribute to the seminal fluid; antisera against A, B and D group proteins recognize epitopes in the lumen of the spermatophore (Black *et al.*, 1982; Paesen *et al.*, 1992b). The exact function of the TAG proteins, however, remains unclear.

The best described TAG products are the highly repetitive D proteins, which appear immediately after eclosion of the adult beetle. Although their molecular masses on SDS-gels vary between 24 and 29 kDa, the D proteins form a very homogenous group with little sequence variation from one protein to another (Grimnes and Happ, 1985; Paesen *et al.*, 1992b).

The A and B proteins have only been characterized on the basis of their electrophoretical mobility and reaction to polyclonal antisera. There are indications, however, that they might be related to each other: they are close in molecular mass (17.9 and 19.0 kDa, respectively) and pI (about 5.3), and share epitopes. Both proteins also appear at the same stage of the adult development (day 2) and accumulate at a similar rate (Black *et al.*, 1982; Black and Happ, 1985). The term A/B proteins refers to the two proteins together.

C proteins are basic, contrary to the A/B and D proteins (Black *et al.*, 1982). Although heterogenous on 2D-gels, C proteins appear as one band on one-dimensional SDS–polyacrylamide gels, corresponding with a molecular mass of 21.9 kDa. The C proteins, like the A/B proteins, are barely detectable on Coomassie-stained SDS-gels before day 2 of the adult development. Thereafter, they accumulate dramatically.

Although we did not have a specific antiserum at our

*The sequence presented in this paper has been submitted to Genbank and is accessible under No. M97262.

† Department of Zoology, University of Vermont, Burlington, VT 05405–0086, U.S.A.
‡Author for correspondence.

disposal, we were able to extract C protein clones from a TAG expression library, using a polyclonal antiserum against the A/B proteins, that was shown to cross-react with C proteins (Black and Happ, 1985).

EXPERIMENTAL PROCEDURES

Animals

Mealworms (*Tenebrio molitor*) were purchased from a commercial supplier; they were kept at room temperature and reared on Purina Chick Labchow. Males and females were separated in the pupal stage.

Protein purification and characterization

TAGs were excised from 10–20 days adult mealworms, rinsed in distilled water containing protease inhibitors (1 μg/ml Leupeptin, 0.5 μg/ml Antipain, 1 μg/ml Pepstatin A, 50 μg/ml PMSF, 1 mM EDTA, pH 8.0; all Sigma products), and collected on dry ice. The sample was homogenized in a low salt Tris–saline buffer (10 mM Tris–HCl, 75 mM NaCl, pH 7.5) and centrifuged at 12,000 *g* (4°C, 5'). The supernatant was submitted to heparin affinity chromatography (Fig. 1). The C-protein containing fraction (identified by SDS–PAGE) was quickly concentrated using Centricon-100 microconcentrators (Amicon), and submitted to reversed phase HPLC (Fig. 1). The purified product was analyzed on one-dimensional (Laemmli, 1970) and two-dimensional polyacrylamide gels (O'Farrell, 1975).

N-terminal sequence analyses were performed in the protein analytical facility of the Medical Biochemistry Department at the University of Vermont, on an Applied Biosystems 475A Protein Sequencing System, employing a gas phase Edman degradation protocol with online HPLC identification of PTH derivatives of amino acids.

cDNA library screening and construction of a C-gene probe

A Lambda Zap II cDNA-library (Paesen *et al.*, 1992), constructed with mRNA from adult TAGs, was screened with a rabbit polyclonal antiserum (anti-A/B), according to Mierendorf *et al.* (1987). Although the antiserum was made against A and B proteins, it was shown to cross-react with C-proteins (Black and Happ, 1985). The pBluescript SK (−) phagemid of positive clones was excised *in vivo*, using the R408 helper phage, as described by Short *et al.* (1988). The plasmids were submitted to sequencing (see below), and one of the clones (C1), suspected to contain C protein–cDNA, was used to construct a digoxygenin-labeled probe (see below). The Lambda Zap library was then submitted to a second round of screening, this time by DNA hybridization of plaque lifts (Sambrook *et al.*, 1989), using the digoxygenin-labeled probe. An anti-digoxygenin antiserum conjugated with alkaline-phosphatase (Boehringer Mannheim) was used for probe detection. BioTrace NT-nitrocellulose filters (Gelman Sciences) were used for plaque transfer, blocking reagent was provided by Boehringer Mannheim, as were NBT and BCIP, used for vizualization of kinase activity.

Probe construction

The C1 clone was grown to saturation in liquid LB. Plasmid DNA was isolated using PlasmidQuik columns (Stratagene). The insert was excised with the restriction endonuclease *Not* I, and separated from the plasmid arms by electrophoresis through a 1% agarose gel onto a piece of NA-45 DEAE–cellulose membrane (Schleicher & Schuell, Inc.). The cDNA was recovered by submerging the membrane in a high salt buffer (1 M NaCl, 0.1 mM EDTA, 20 mM Tris, pH 8), precipitated with ethanol and redissolved in distilled water. It was then used as a template for random primer labeling with digoxygenin (Boehringer Mannheim).

Sequencing

Sequencing was carried out by the Sanger dideoxy-mediated chain termination reaction (Sanger and Coulson, 1975), using Sequenase (UCB). Double stranded (in the case of C2), as well as single stranded (C1) templates were prepared. Single stranded DNA was rescued and purified according to Short *et al.* (1988). Double stranded phagemid DNA was purified on Plasmid Quik columns (Stratagene) and alkali-denatured (Mierendorf and Pfeffer, 1987). Compressions were resolved by substituting dITP for dGTP. Sequence data were analyzed using IBI Pustell and Genetics Computer Group (GCG; Devereux *et al.*, 1984) sequence analysis software.

Northern blotting

mRNA was isolated from TAGs from adult beetles, using the MicroFastTrack isolation system (InVitrogen). 2 μg RNA samples were subjected to electrophoresis on formaldehyde containing agarose (0.8%) gels, according to Sambrook *et al.* (1989). The RNA was transferred to a nylon membrane (Hybond N+, Amersham)—using 20 × SSC as transfer buffer—and hybridized with the digoxygenin-labeled probes. Nitroblue tetrazolium salt (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) were used for vizualization of bound anti-digoxygenin alkaline phosphatase conjugate.

RESULTS

C protein purification (Fig. 1)

The C protein was very efficiently purified by heparin affinity chromatography followed by a filtration-centrifugation over Centricon-100 units. Despite their high cut-off value (100 kDa), the membranes of the filter units do retain the purified protein (21.5 kDa), suggesting aggregation or polymerization. The isolated protein, obviously a major TAG product, could easily be identified as a C-group protein, based upon its mobility on 12% SDS-gels and 2D-polyacrylamide gels. However, when the C-protein fraction was submitted

to an additional purification by reversed phase HPLC, two peaks appeared (a and b in Fig. 1C). The N-terminal sequences of the protein in fraction a (KPSXNXADKXN) and fraction b (KPSSNSADKK) are identical (apart from the unidentified amino acids), suggesting that the two peaks correspond with two C protein variants with a slightly different hydrophobicity.

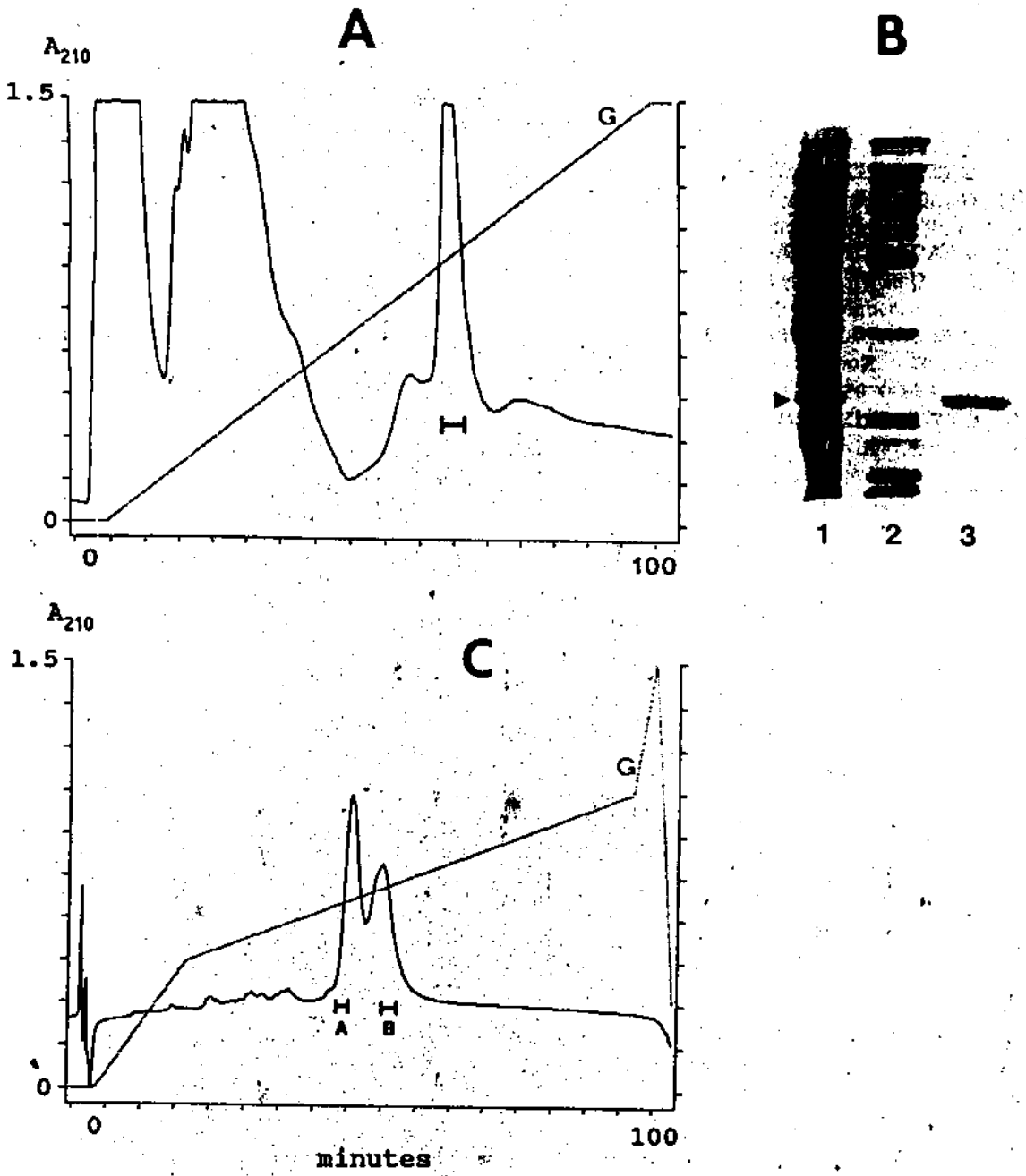

FIGURE 1. Purification of the C protein. A—Heparin-affinity chromatography. A TAG homogenate was applied to a heparin-affinity chromatography column (Econo-Pac Heparin Cartridge, Biorad). Proteins were eluted in a sodium phosphate buffer (60 mM Na_2HPO_4 was brought to pH 6.95 with a concentrated NaH_2PO_4 solution), over a linear gradient (G) from 0 to 2 M NaCl in 100 min. The flow rate was 0.75 ml/min. The absorbance was measured at a 210 nm wavelength (A_{210}). B—The fraction, indicated by the bar in Fig. 2A, was concentrated using Centricon-100 filter units (Amicon). An aliquot was resolved on a 12% SDS–polyacrylamide gel (Laemmli, 1970) (lane 3). Lane 1 contains total TAG homogenate. The triangle indicates the C protein position. Lane 2 contains molecular mass standards (Biorad) (a: 31 kDa; b: 21.5 kDa). C—C protein fractions from several heparin-chromatography runs were pooled and further purified by means of HPLC, using a 15 cm C18 Reversed Phase Column (Vydac) and a 0.1% trifluoroacetic acid (Pierce) running buffer over a 0–100% acetonitrile (J. T. Baker Chemicals) gradient. Two peaks appeared. We collected an early fraction of the first peak (bar A) and a late fraction of the second (B). Both fractions were submitted to N-terminal sequence analysis.

```
CCTCACCCATTGAATTCATCAGGAGACTAAAGTAAAATCATATTCCCAGCCAACTGTTCC

CAGAGTATTTTGCTACTATTTTGTCTGTATCTTGCACATAGAGTTATCAAAATGAGAGTG
                                                    M  R  V

GTTCTGATTGCGTTGATTTGGGCTGTTGCCAGCGATTTCAACTTCTATTCGAAAGTTGAA
V  L  I  A  L  I  W  A  V  A  S  D  F  N  F  Y  S  K  V  E

GCAAAGCCATCCTCTAATTCGGCCGATAAGAAGAACATAATTCTTCATTTTAATCAGAAC
A  K  P  S  S  N  S  A  D  K  K  N  I  I  L  H  F  N  Q  N
-1 ↑     ▲  ▲     ▲

GGTCCTGTTAATGTCCGTGGAGCTCCGGTAAATCTCGAAAGTGCGTTGAAACCTCTTTTG
20 G  P  V  N  V  R  G  A  P  V  N  L  E  S  A  L  K  P  L  L

GAACAATTTTCAGCAGTGATGATGGAGGAGATTACCAAAGTACTTCAAAAGCCAAACGGC
40 E  Q  F  S  A  V  M  M  E  E  I  T  K  V  L  Q  K  P  N¹ G

TGTTCAAAACATAAAGGTTCAAGATCATTCCGTCTGCCAATACAGAAGAAAACATTGAGC
60 C  S  K  H  K  G  S  R  S²  F  R  L  P  I  Q  K  K  T  L  S
                              ▲  △                          ▲

CTTCCGGCAGCCGATCCACCTAGTGGTGATCGAGAGAAGGGAAATACTTCTGAAGGGGCA
80 L  P  A  A  D  P  P  S  G  D  R  E  K  G  N  T  S  E  G  A
                           ▲                 ▲

ACCGTCGAAGCTCCTGTCGCAGCAATAGTAGAGGTTCAAAATGGTGACTCGAAACAGCCA
100 T  V  E  A  P  V  A  A  I  V  E³ V  Q  N  G  D  S  K  Q  P
 ▲

AATCACGAGATAGAAGTCCTAACTCAATAATCAAGACCAGCACCACAACCAATGCGGCGT
120 N  H  E  I  E  V  L  T  Q

CGACAAAATTTGCACATCACAGAAACTCCAAGAATCACAAGAATCAAAACTACCATTG
```

FIGURE 2. cDNA and deduced amino-acid sequence of C2. The repeated sequences in the region upstream of the startcodon are underlined. The arrow indicates the putative signal peptidase cleavage site. The sequence obtained by N-terminal sequencing of the purified protein is overlined. The amino acids that are underlined and marked with a number (N,$^1S^2$ and E^3), do not occur in C1; they are replaced by a serine, a proline and a glutamine, respectively. The putative N-glycosylation site is boxed. The putative heparin-binding site is double underlined. The triangles indicate possible phosphorylation sites (the open triangle corresponds with a putative phosphorylation site in C2 that is absent in C1).

cDNA sequences

We sequenced a cDNA clone, termed C1, which was isolated from the library by screening with the polyclonal anti-A/B antiserum. Additionally, a second clone (C2) was sequenced, which was obtained by screening with digoxygenin labeled C1. The nucleic acid sequence of the C2 clone (658 bases) and its inferred amino-acid sequence are shown in Fig. 2. Since the region upstream of the presumed startcodon contains a stopcodon, we believe that the startcodon in Fig. 2 corresponds with the first methionine codon in the reading frame.

The C2 cDNA (not shown) lacks the first 41 basepairs. In the region which the two clones have in common, differences are observed at 12 positions. The translation products, however, differ in only three amino-acids. Neither C1, nor C2 contains a polyadenylation signal. The startcodon in the mRNA (the first methionine-codon of the open reading frame) is preceded by at least 111 nucleotides (the number of basepairs in the upstream region of the C1 cDNA). In this rather long region, the sequence TAT(T/C)TTG is repeated three times (see Fig. 2). The meaning of this repeat, if any, is unclear. Hitherto, there are no data that suggest an eventual regulatory role in transcription or translation.

The translation products

The N-terminal sequence of the HPLC-purified protein is, in the cDNA translation product, preceded by a 24 amino-acid pro-piece. This pro-piece most probably constitutes the signal sequence, a common feature in secretory proteins. Although the most likely signal peptidase site, determined with the method of Von Heijne (1986), is located in between amino acids-12 and-11 (ALIWAVASDFNF.), the second highest score was obtained for a signal peptidase site at the expected position, i.e. immediately preceding the N-terminus of the secreted protein (indicated by the arrow in the

to spermatozoa and modulate capacitation (Miller *et al.*, 1990). It is believed that the addition of the seminal proteins to sperm cells provides the latter with extra heparin-binding sites, which in turn enables a substantial increase in the binding of heparin-like glycosaminoglycans (that are present in the female reproductive tract) to the sperm. The addition of glycosaminoglycans to the sperm surface enhances capacitation and induces acrosome reactions, at least *in vitro* (Gebauer *et al.*, 1978; Hurst *et al.*, 1988; Chandonnet *et al.*, 1990; Miller and Ax, 1990; Nass *et al.*, 1990).

In invertebrate phyla, a similar mechanism has been described in the fertilization of sea urchin eggs, where the binding of a (proteoglycan-like) fucose sulfate glycoconjugate to a sperm surface glycoprotein induces the acrosome reaction (Rossignol *et al.*, 1984; DeAngelis and Glabe, 1990). Sperm activation in insects has been best studied in the Lepidoptera by Shepherd (1974) and the Osanai group (Aigaki *et al.*, 1987; Osanai and Kasuga, 1990) where a proteolytic cascade is mediated by accessory gland secretions. We have seen no sequence data for these proteins. Sulfate-containing glycosaminoglycans have been detected in the ovarioles of the Colorado potato beetle, *Leptinotarsa decemlineata*, particularly in secretion vesicles of the follicle cells and in the intercellular space between the oocyte and the follicle (Stynen *et al.*, 1986).

Proteases have been described in insect accessory glands of locusts (Cheeseman *et al.*, 1990) and silkmoths (*Bombyx mori*) (Aigaki *et al.*, 1987). In *Bombyx*, the proteases from the glandula prostatica have been implicated in a cascade of sperm activation steps. Vertebrate protease inhibitors are typically heparin binders (Jackson *et al.*, 1991), and the C-protein might conceivably function in that role or might bind to sperm, but prelimiary experiments were inconclusive. A protease inhibitor has been identified in the secretion of the male accessory glands of *Drosophila funebris* (Schmidt *et al.*, 1989). However, the 63 amino-acid inhibitory peptide of *Drosophila* has no obvious sequence resemblance with the *Tenebrio* C protein.

REFERENCES

Aigaki T., Kasugo H. and Osanai M. (1987) A specific endopeptidase, BAEE esterase, in the glandula prostatica of the male reproductive system of the silkworm, *Bombyx mori*. *Insect Biochem.* 17, 323–328.

Barbas J. A., Chaix J-C., Steinmetz M. and Goridis C. (1988) Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse. *EMBO J.* 7, 625–632.

Black P. N., Landers M. H. and Happ G. M. (1982) Cytodifferentiation in the accessory glands of *Tenebrio molitor*. VIII. Crossed immunoelectrophoretic analysis of terminal differentiation in the postecdysial tubular accessory glands. *Devs. Biol.* 94, 106–115.

Black P. N. and Happ G. M. (1985) Isolation, partial characterization, and localization of the A and B proteins from the tubular accessory glands of male *Tenebrio molitor*. *Insect Biochem.* 15, 639–650.

Cardin A. D. and Weintraub H. J. R. (1989) Molecular modeling of protein–glucosaminoglycan interactions. *Arteriosclerosis* 9, 21–32.

Chandonnet L., Roberts K. D., Chapdelaine A. and Manjunath P. (1990) Identification of heparin-binding proteins in bovine seminal plasma. *Molec. Reprod. Dev.* 26, 313–318.

Cheeseman M. T., Gillott C. and Ahmed I. (1990) Structural spermatophore proteins and a trypsin-like enzyme from the reproductive accessory glands of the male grasshopper, *Melanoplus sanguinipes*. *J. exp. Zool.* 255, 193–204.

Chen P.-S. (1991) Biochemistry and molecular regulation of the male accessory gland secretions in *Drosophila (Diptera*. *Ann. Soc. ent. Fr.* 27, 231–244.

Dailey P. J., Gadzama N. M. and Happ G. M. (1980) Cytodifferentiation in the accessory glands of *Tenebrio molitor*. VI. A congruent map of cells and their secretions in the layered elastic products of the male bean-shaped gland. *J. Morphol.* 178, 139–154.

DeAngelis P. L. and Glabe C. G. (1990) Specific recognition of sulfate esters by bindin, a sperm adhesion protein from sea urchins. *Biochem. biophys. Acta* 103, 100–105.

Devereux J., Haeberli P. and Smithies O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* 12, 387–395.

Gebauer H., Lindler H. R. and Amsterdam A. (1978) Synthesis of heparin-like glycosaminoglycans in rat ovarian slices. *Biol. Reprod.* 18, 350–358.

Glass D. B., El-Maghrabi M. R. and Pilkis S. J. (1986) Synthetic peptides corresponding to the site phosphorylated in 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase as substrates of cyclic nucleotide-dependent protein kinases. *J. biol. Chem.* 261, 2987–2993.

Grimnes K. A. and Happ G. M. (1985) Partial characterization of the D group proteins of the tubular accessory glands of *Tenebrio molitor*. *Insect Biochem.* 15, 181–188.

Grimnes K. A., Bricker C. S. and Happ G. M. (1986) Ordered flow of secretion from accessory glands to specific layers of the spermatophore of mealworm beetles: demonstration with a monoclonal antibody. *J. exp. Zool.* 240, 275–286.

Happ G. M. (1992) Maturation of the male reproductive system and its endocrine regulation. *A. Rev. Entomol.* 37, 303–320.

Happ G. M., Yuncker C. and Huffmire S. A. (1977) Cytodifferentiation in the accessory glands of *Tenebrio molitor*. II. Patterns of leucine incorporation in the tubular glands of post-ecdysial adult males. *J. exp. Zool.* 200, 223–236.

Hurst R. E., Bynum R. L., Einfeldt S. E. and Roy B. J. (1988) The identification of a heparin-binding protein on the surface of bovine sperm. *Biochem. biophys. Res. Commun.* 153, 289–293.

Jackson R. L., Busch S. J. and Cardin A. D. (1991) Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiol. Rev.* 71, 481–539.

Kishimoto A., Nishiyama K., Nakanishi H., Uratsuji Y, Nomura H., Takeyama Y. and Nishizuka Y. (1985) Studies on the phosphorylation of myelin basic protein by protein kinase C and adenosine 3′:5′-monophosphate-dependent protein kinase. *J. biol. Chem.* 260, 12492–12499.

Kuenzel E. A., Mulligan J. A., Sommercorn J. and Krebs E. G. (1987) Substrate specificity determinants for casein kinase II as deduced from studies with synthetic peptides. *J. biol. Chem.* 262, 9136–9140.

Laemmli V. K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* 277, 680–685.

Leopold R. A. (1976) The role of male accessory glands in insect reproduction. *A. Rev. Ent.* 21, 199–221.

Mierendorf R. C. and Pfeffer D. (1987) Direct sequencing of denatured plasmid DNA. *Meth. Enzymol.* 152, 556–562.

Mierendorf R. C., Percy C. and Young R. A. (1987) Gene isolation by screening lambda gt11 libraries with antibodies. *Meth. Enzymol.* 152, 458–469.

Miller D. J. and Ax R. L. (1990) Carbohydrates and fertilization in animals. *Molec. Reprod. Dev.* 26, 184–198.

Miller D. J., Winer M. A. and Ax R. L. (1990) Heparin-binding proteins from seminal plasma bind to bovine spermatozoa and modulate capacitation by heparin. *Biol. Reprod.* 42, 899–915.

Nass S. J., Miller D. J., Winer M. A. and Ax R. L. (1990) Male accessory sex glands produce heparin-binding proteins that bind the

not predict a lot of periodic secondary structure.

The calculated molecular mass of the C2 translation product (without signal sequence), was 13,775 Da. The basic pI (calculated to be 8.82) is due to the relatively high content of lysine (12 residues/mol; together with serine, the most abundant residue in the molecule).

The C2 translation product contains one putative N-glycosylation signal (Oikawa *et al.*, 1987), and eight putative phosphorylation sites (Weller, 1979; Kishimoto *et al.*, 1985; Glass *et al.*, 1986; Kuenzel *et al.*, 1987).

The amino-acid sequence from position 61 to 68 (SKHKGSRS) corresponds with a consensus sequence for heparin binding (XBBBXXBX, where X stands for a hydropathic residue and B for a basic residue; Cardin and Weintraub, 1989).

Northern blotting

The Northern blot in Fig. 3 shows that the amount of mRNA coding for C proteins is relatively low in young adult males (1 day after emergence), but increases later on in the development. This is completely in accordance with the dramatic increase in C-protein production from day 2 of adult development (Black *et al.*, 1982). The abundance of the C protein mRNA in the adult males is reflected in the ease with which C-protein cDNAs could be extracted from the library, which was constructed from TAG mRNA of 5–8 day old adults. Using the dioxygenin-labeled probe, only one round of screening was required to obtain tens of positive clones from a single 100 mm agar plate containing about 5000 plaque forming units.

DISCUSSION

The two C proteins presented in this paper are very similar to one another. The three positions in which they differ from each other, still contain similar amino acids, suggesting that there is little difference in the overall characteristics of the two proteins. The question remains whether the minute differences between C1 and C2 suffice to explain the appearance of two peaks after reversed phase chromatography. Maybe C1 and C2 account for only one peak, while another, yet unidentified C protein gives rise to the second peak. However, we do not expect to find C proteins that are considerably different from the ones published here. SDS–polyacrylamide-gel electrophoresis of TAG homogenates does not support the existence of a lot of divergence in the C group.

There is a considerable discrepancy between the calculated molecular mass of the cDNA translation products (13.8 kDa) and the apparent molecular mass of the purified C protein on SDS–polyacrylamide gels (21.9 kDa). The binding of concanavalin A (data not column. For comparison, typical heparin binders, such as p30, a rat cell adhesion protein (Rauvala *et al.*, 1988) and human plasma fibronectin (Zlatopolsky *et al.*, 1992) are eluted with no more than 1 M NaCl.

At this point, we do not know what is the exact function of the C protein. The GenBank or NBRF databases do not contain sequences that are significantly homologous to C1 or C2. One of the higher scores obtained in the database search was with a region in a mouse neural cell adhesion molecule (Barbas *et al.*, 1988). The best-fit program of GCG (gap weight 3.0, length 0.1) predicts 47% similarity and 20.5% identity between the complete C protein sequence and a fragment of the mouse neuronal adhesion molecule. However, this similarity is dependent upon choice of fragments and gap weight conditions, and we conclude only that the similarity indicates analogous binding functions. This indication, together with the fact that C proteins show a high avidity for heparin, suggests a possible role in sperm coating and capacitation. In mammals, it has been demonstrated that heparin-binding proteins from seminal plasma (among others a 24 kDa, basic protein) bind

FIGURE 3. Northern blotting. The content of mRNA for C-protein in TAGs of 1, 3, and 6 day adult mealworm beetles is shown. Comparison with standards indicates that the single major band is from mRNA between 530 and 780 bases in length.

cauda epididymal spermatozoa and are testosterone dependent. *Molec. Reprod. Dev.* **25**, 237–246.

O'Farrell, P. H. (1975). High-resolution two-dimensional electrophoresis of proteins. *J. biol. Chem.* **250**, 4007–4021.

Oikawa S., Nakazato, H. and Kosaki G. (1987) Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA sequence. *Biochem. biophys. Res. Commun.* **142**, 511–518.

Osanai M., Kasuga H. and Aigaki T. (1989) Induction of motility of apyrene spermatozoa and dissociation of eupyrene sperm bundles of the silkmoth, *Bombyx mori*, by iniatorin and trypsin. *Inv. Reprod. Dev.* **15**, 97–103.

Paesen G. C., Schwartz M. B., Peferoen M., Weyda F. and Happ G. M. (1992a) Amino-acid sequence of Sp23, a structural protein of the spermatophore of the mealworm beetle, *Tenebrio molitor*. *J. biol. Chem.* **267**, 18852–18857.

Paesen G. C., Weyda F. and Happ G. M. (1992b) Deduced sequences show multiple repeats in two D proteins from the tubular accessory glands of *Tenebrio molitor*. *Insect Biochem. Molec. Biol.* **22**, 387–398.

Rauvala H., Merenmies J., Pihlaskari M., Korkolainen M., Huhtahla M-L. and Panula P. (1988) The adhesive and neurite-promoting molecule p30: analysis of the amino-terminal sequence and production of anti-peptide antibodies that detect p30 at the surface of neuroblastoma cells and of brain neurons. *J. Cell Biol.* **107**, 2293–2305.

Rossignol D. P., Earles B. J., Decker G. L. and Lennarz W. J. (1984) Characterization of the sperm receptor on the surface of eggs of *Strongiyocentrotus purpuratus*. *Dev. Biol.* **104**, 308–321.

Sambrook J., Fritsch E. F. and Maniatis T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory.

Sanger F. and Coulson A. R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Molec. Biol.* **94**, 441–448.

Schmidt T., Stumm-Zollinger E., Chen P-S., Boehlen P. and Stone S. (1989) A male accessory gland peptide with protease inhibitory activity in *Drosophila funebris*. *J. biol. Chem.* **264**, 9745–9749.

Shepherd J. G. (1974) Activation of saturniid sperm by a secretion of the male reproductive tract. *J. Insect Physiol.* **20**, 2107–2122.

Short J. M., Fernandez J. M., Sorge J. A. and Huse W. D. (1988) Lambda ZAP: a bacteriophage lambda expression vector with *in vivo* excision properties. *Nucl. Acids Res.* **16**, 7583–7599.

Stynen D., Tyteca L. and De Loof A. (1986) Localization of glycosaminoglycans in the ovarioles of the Colorado potato beetle, *Leptinotarsa decemlineata* (Say). *Ann. Soc. r. zool. Belg.* **116**, 131–136.

Von Heijne G. (1986) A new method for predicting signal cleavage sites. *Nucl. Acids Res.* **14**, 4683–4690.

Weller M. (1979) *Protein Phosphorylation: The Nature, Function, and Metabolism of Proteins which Contain Covalently Bound Phosphorus*. Pion, London.

Zlatopolsky A. D., Chubukina A. N. and Berman A. E. (1992) Heparin-binding fibronectin fragments containing cell-binding domains and devoid of hep2 and gelatin-binding domains promote human embryo fibroblast proliferation. *Biochem. biophys. Res. Commun.* **183**, 383–389.

Acknowledgements—The assistance of Tim Valley in determination of amino acid composition and in amino-terminal sequence analysis is acknowledged. We thank the National Institutes of Health (AI-15662) for financial support.