Glutactin, a novel *Drosophila* basement membrane-related glycoprotein with sequence similarity to serine esterases

Pamela F.Olson, Liselotte I.Fessler,
Robert E.Nelson, Rachel E.Sterne¹,
Andrew G.Campbell² and John H.Fessler

Molecular Biology Institute and Biology Department, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024–1570, USA

¹Present address: Department of Biochemistry, University of California, Berkeley, CA.

²Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, CA

Communicated by W.Huttner

Glutactin, a new acidic sulfated glycoprotein, was isolated from *Drosophila* Kc cell culture media. Immunofluorescence microscopy located it to embryonic basement membranes, particularly to the sequentially invaginated envelope of the central nervous system, muscle apodemes and dorsal median cell processes. Its chromosome locus is 29D. The nucleic acid sequence coding for the 1023 residue long polypeptide contains one intron and was confirmed by partial amino acid sequencing. Glutactin has a signal peptide and an amino domain of >500 residues that strongly resembles acetylcholine esterases and other serine esterases, but lacks the catalytically critical serine residue. The amino and carboxyl domains of glutactin are separated by 13 contiguous threonine residues. Glutamine and glutamic acid make up 44% of glutactin's very acidic carboxyl domain. Glutactin preferentially binds Ca²⁺ in the presence of excess Mg²⁺ and four of its tyrosines are O-sulfated. Several similarities with mammalian entactin caused our previous, preliminary mention of glutactin as a putative *Drosophila* entactin, but sequence comparison now shows them to be different proteins.
Key words: basement membranes/development/*Drosophila*/serine esterases/tyrosine sulfate

Introduction

The extracellular matrix that is immediately adjacent to cells influences their differentiation and function (Hay, 1981). Here we describe a new glycoprotein component of this pericellular matrix in *Drosophila*. The protein occurs in basement membranes, which are condensations of specialized extracellular matrix that are adjacent to sheets of epithelial cells, surround muscle cells and envelop parts of the nervous system. Although basement membranes are widely distributed among metazoans, their fundamental roles in cell biology are poorly understood. We are exploring the molecular biology of basement membranes during the development of *Drosophila melanogaster* because this animal is particularly amenable to genetic and molecular investigations.

Cultured cells of *D.melanogaster* secrete several

macromolecules that locate to *Drosophila* basement membranes by immunomicroscopy. Among these materials are major components of basement membranes, such as *Drosophila* collagen IV, laminin, the proteoglycan papilin and the present protein, glutactin. At least two major basement membrane macromolecules, collagen IV and laminin, have been remarkably conserved during evolution (Fessler and Fessler, 1989).

Glutactin consists of a single polypeptide which comigrates in SDS–polyacrylamide gel electrophoresis with a glycoprotein of vertebrate basement membranes called entactin or nidogen (Hogan *et al.*, 1982; Durkin *et al.*, 1988; Mann *et al.*, 1989; Olsen *et al.*, 1989). As both proteins contain sulfated tyrosine residues and occur in equivalent morphological locations, we referred to glutactin as a putative *Drosophila* entactin in preliminary communications (Fessler *et al.*, 1984; Olson *et al.*, 1987). However, we now conclude that this *Drosophila* polypeptide is a new protein because most of its amino acid sequence is entirely different from that of vertebrate entactin (Durkin *et al.*, 1988; Mann *et al.*, 1989; Olsen *et al.*, 1989). Nearly half of the carboxyl domain of glutactin consists of glutamic acid or glutamine residues, and to denote this unusual composition we refer to this protein henceforth as *Drosophila* glutactin. Here we provide a partial characterization of glutactin as gene and protein, and describe its immunolocalization in *Drosophila* embryos. Particular attention is given to structures associated with the developing nervous system.

Results

The sulfated glycoprotein glutactin was isolated from the cell culture medium of *Drosophila* Kc cells and antibodies were raised against the purified protein. After immunomicroscopy had located glutactin to basement membranes the antibodies were used to isolate cDNA clones from an expression library. The deduced amino acid sequence was corroborated by partial amino acid sequencing of the protein, and stimulated further investigations of the properties of glutactin.

Isolation and purification of glutactin

Velocity sedimentation partially separates the glycoproteins secreted by *Drosophila* cells labeled with [³⁵S]sulfate (Figure 1). The slower sedimenting, strongly labelled protein in fractions 14–16 consists of a single polypeptide of apparent mol. wt 155 kd and is glutactin. The faster sedimenting labeled protein has been described as the sulfated proteoglycan papilin (Campbell *et al.*, 1987). The pooled fractions 14–16 contained glutactin, collagen IV and lower mol. wt proteins and were chromatographed by stepwise increases of NaCl on a Mono Q column. Glutactin and some lower mol. wt proteins were eluted with 0.5 M NaCl. The smaller proteins were removed by molecular sieve chromatography through a Superose 12 column from which glutactin elutes with $K_{AV} = 21.5$. The glutactin was further

Fig. 1. Isolation of glutactin from cell culture media. The concentrated proteins of *Drosophila* Kc cell culture media were sedimented through a sucrose gradient. (A) shows an electrophoretogram of successive, reduced fractions. The direction of sedimentation was from right to left. Fractions 14–16 were pooled, glutactin was purified and its electrophoretic analysis is shown in (B). Coomassie blue staining was used for A and B. Proteins were similarly analysed from Kc cell cultures that had been supplied with [³⁵S]O₄²⁻, and (C) shows the distribution of radioactivity through a sucrose gradient sedimented as in (A). The corresponding electrophoretic analysis of fractions 14–17 of that gradient are shown, without further purification, as a fluorogram in (D). X indicates an as yet unnamed extracellular glycoprotein.

purified on a Superose 6 column, from which it eluted as a sharp peak with $K_{Av} = 41.5$. This material gives a single electrophoretic band with traces of breakdown products which increase with storage.

Antibody characterization

Antibodies to glutactin specifically immunoprecipitated [³H]leucine-labeled glutactin. In a Western blot of the mixture of medium proteins, and of an SDS embryo extract, only glutactin reacted with the antibodies (Figure 2). The same results were obtained with either reduced or non-reduced proteins. The pre-immune sera showed no reaction on the blots. Pre-adsorption of the antisera with glutactin immobilized on a microtiter well eliminated the reaction. Antibodies which were affinity purified on a glutactin–Sepharose CL4B column gave the same reactions and the same immunofluorescence patterns as the immune serum. This showed that the antibodies react specifically with glutactin isolated from cell cultures or from embryos.

Cloning and nucleic acid sequence of glutactin

A Kc cell cDNA library made in λ_{gt11} was screened with antiserum against *Drosophila* glutactin. Restriction mapping, cross-hybridization and DNA sequencing showed that clones A7, A12 and A14 (Figure 3A,B) overlapped and contained an open reading frame, followed by a polyadenylation signal (Proudfoot and Brownlee, 1976) and a poly(A)tail. RNA probes which were synthesized from each of these clones hybridized to a band of ~3.8 kb on a Northern blot of Kc cell total RNA. Corresponding DNA probes hybridized to

Fig. 2. Immunoblot. Total proteins from Kc cell culture medium (lanes a–e) and from an SDS extract of 10–18 h embryos (lanes f, g), non-reduced or reduced with dithiothreitol (dt), were electrophoresed and blotted onto nitrocellulose. These blots were reacted with anti-glutactin (lanes b, c, f) or pre-immune serum (lanes d, e, g) and then with anti-rabbit IgG conjugated to alkaline phosphatase. The blot in lane a was stained with amido black.

a single *Drosophila* chromosome location at position 29D. To confirm that these clones were representative of the glutactin protein, antibodies were affinity selected from *Drosophila* glutactin antisera by adsorption to fusion proteins from clones A7 or A12; or wild-type λ_{gt11} and tested against several Kc cell secreted proteins. Only the antibodies which were affinity selected using either A7 or A12 fusion proteins reacted specifically with glutactin on a Western blot and with both A7 and A12 fusion proteins, while antibodies selected with the wild-type λ_{gt11} fusion protein neither reacted on a Western blot nor with any fusion proteins.

A Kc cell cDNA library made in λ_{gt10} was screened for clones which extended the overlapping cDNAs in a 5' direction. Restriction mapping, cross-hybridization, and partial sequencing showed that 15 new subclones overlapped and extended the cDNA sequence. Clone H9 contained the complete coding region, from the initiation methionine to the poly(A)tail. All of the cDNA clones analysed had overlapping restriction maps which were in complete agreement. The cDNA clones which were partially or completely sequenced are shown in Figure 3B. The 3329 bp length of the assembled cDNA sequence (Figure 4) is consistent with a 3.8 kb message. The cDNA contains a single open reading frame of 1023 predicted amino acids which is flanked by several in-frame stop codons. The sequence surrounding the initiation codon contains a good ribosome initiation site (Kozak, 1986). Beginning with the initiation methionine there is a signal peptide of 17 amino acids which contains a good cleavage site (Perlman and Halvorson, 1983). Downstream of the termination codon there is a polyadenylation signal, followed further downstream by a poly(A)tail.

Glutactin gene structure

To determine the extent of the glutactin gene, Southern blots of *EcoRI* or *BamHI* digested genomic DNA were hybridized with a probe that represented the complete glutactin cDNA. Two *EcoRI* bands of ~4.7 kb each and two *BamHI* bands of 1.6 kb and ~14 kb hybridized (not shown). As there is one site each for *EcoRI* and *BamHI* in the cDNA, the pattern

Fig. 3. *Drosophila* glutactin cDNA clones. (A) A partial restriction map of the composite cDNA. (B) Overlapping cDNA clones which were partially or completely sequenced. (C) A partial restriction map of the glutactin gene. Regions of genomic DNA which were sequenced are indicated by black bars. The schematic representation of the mRNA encoding region is shown. The protein coding region is indicated by cross-hatching. The 5' half of exon 1, indicated by dotted lines, was deduced from the predicted transcriptional start site following the TATA box in the genomic sequence and was not included in the cDNA. (D) *EcoRI* and *BamHI* restriction map of 25 kb of genomic DNA surrounding the glutactin gene. The glutactin gene is indicated by the black bar. (B = *BamHI*, E = *EcoRI*, G = *BglII*, H = *HindIII*, P = *PstI*, X = *XbaI*).

of hybridization suggests that glutactin is a single copy gene. A reduced stringency genomic Southern blot did not reveal any homologous genes. The *Drosophila* genomic cosmid library JT101 was screened with a probe which represented the complete glutactin cDNA. Three unique cosmid clones were restriction mapped with *EcoRI* and *BamHI* and shown to comprise ~60 kb of genomic sequence spanning the glutactin gene. A fine restriction map was constructed for the 18 kb of DNA encompassing the gene. The genomic restriction map (Figure 3D) is consistent with the genomic Southern hybridization results. Comparison of the cDNA and fine genomic restriction maps indicated two introns (Figure 3A and C). Selected regions of the genomic DNA, indicated in Figure 3C, were sequenced. The genomic sequence (Figure 4) contains a TATA box which is 162 nucleotides from the 5' end of the cDNA. The locations and extent of the introns were determined by comparing the genomic and cDNA sequences. The first intron is 335 nucleotides long and is within the 5'-untranslated region of the message. The second intron is 394 nucleotides long, is within the amino domain of the protein and splits a codon. Both introns contain boundary sequences which are good 5'-donor and 3'-acceptor splice sites (Mount, 1982), as well as internal sequence elements which are conserved in *Drosophila* introns (Keller and Noon, 1985). The only discrepancy between the cDNA and genomic sequences is a single nucleotide, number 4342, in the 3'-untranslated region. The cDNA and genomic libraries were derived from different fly strains. Comparison of the genomic sequence with the cDNA sequences of clones A7 and A12 reveals multiple sites of polyadenylation (Figure 4).

The regions of the glutactin sequences which were verified by amino acid sequencing of the protein are underlined in Figure 4. The amino terminus of secreted glutactin was determined by peptide sequencing and agreed for 20 amino acids with the predicted sequence following the signal peptide. Glutactin was cleaved with CNBr. The electrophoretically estimated sizes of the resulting peptides were consistent with those predicted from the sequence, with the exception of a large, highly acidic peptide that corresponds to the carboxyl end of glutactin. The apparent mol. wt of this peptide was 90 kd, while that predicted from the nucleic acid sequence is 60 kd. The aberrant electrophoretic behavior of this peptide paralleled that of glutactin, and the peptide contains most of the high negative charge of glutactin. The N-terminal amino acid sequence of this large CNBr peptide

was determined and 11 residues matched those following a Met residue, as predicted from the nucleotide sequence. The predicted COOH terminus of glutactin is -ERNSRN. In agreement with this, digestion of glutactin with carboxypeptidase A, which is blocked by R, only set free asparagine, N. Supplementation with carboxypeptidase B, and more carboxypeptidase A, gave N, R, S in the molar ratios 2.0:2.3:0.8. This is in agreement with the ratio 2:2:1 expected from the terminal sequence -ERNSRN if further digestion is blocked by E. No glutamic acid, E, was found in the digest. Amino acid analyses of purified glutactin (not shown) were consistent with the amino acid composition predicted from the cDNA sequence.

Characterization of glutactin

Glutactin is a strongly anionic protein, and this is also predicted from its derived amino acid sequence. Its isoelectric point is <3, as indicated by attempted chromatofocusing and pseudo-isoelectric focusing. The protein is unstable at very low pH, probably due to hydrolysis of Asp-Pro peptide linkages. The electrophoretic behavior of glutactin in SDS-polyacrylamide gels is aberrant. The apparent mol. wt relative to standard proteins in SDS-polyacrylamide gel electrophoresis decreased from 160 kd to 145 kd with increasing acrylamide concentration (4.5–8.0%). The apparent mol. wt of denatured reduced glutactin was 135 kd by molecular sieve chromatography. Replacement of SDS by the cationic detergent hexadecyl trimethyl ammonium chloride (Eley *et al.*, 1979) during polyacrylamide gel electrophoresis substantially increased the electrophoretic mobility of glutactin relative to the standard proteins (not shown). After CNBr cleavage of glutactin an equivalently aberrant electrophoretic behavior was also shown by a peptide that encompassed the highly acidic carboxyl region of glutactin, but not by other fragments. Probably, the electrophoretic mobility of glutactin in ionic detergent buffers is influenced by the protein's excess of negatively charged amino acid residues as well as by attached carbohydrate and sulfate. Therefore the mol. wt of the glutactin polypeptide chain cannot be obtained from electrophoretic comparisons with the usual protein standards.

This glycoprotein gives a strong periodic acid Schiff's reaction and binds to a phenylboronate agarose column. In cell cultures, radioactive glucose, mannose or glucosamine are extensively incorporated into glutactin (not shown). There are five sites for Asp-linked glycosylation in the glutactin

Experiment	1	2	3	4	5	6
Kd (μM)	90	30	30	100	80	80
Binding Sites	8	4	5	10	15	8

sequence. Removal of asparagine-linked carbohydrate chains with endoglycosidase H or F caused an apparent decrease of mol. wt of 7 kd. This did not remove [³⁵S]sulfate from biosynthetically labeled glutactin. Additional sugar chains could be linked to serine or threonine residues of glutactin. In equilibrium buoyant density gradients of 4 M guanidine hydrochloride/CsCl, glutactin is buoyant in the density group of general proteins, and not of sulfated proteoglycans (Campbell, 1986; Campbell *et al.*, 1987).

The sulfate incorporated into glutactin is exclusively present as tyrosine sulfate. When [^{35}S]sulfate-labeled, purified glutactin was hydrolysed with $\text{Ba}(\text{OH})_2$ which liberates tyrosine sulfate and free sulfate from sulfated carbohydrate side chains, all the radioactive label was found in the soluble fraction after neutralization with H_2SO_4 . ^{35}S -labeled tyrosine sulfate was identified by two-dimensional electrophoresis on cellulose thin layer plates at pH 1.9 and 3.5, where the labeled component co-electrophoresed with standard tyrosine sulfate (not shown). Also, like tyrosine sulfate, it was quantitatively hydrolysed to [^{35}S]sulfate after 5 min in 1 M HCl at 100°C . A parallel analysis of [^{35}S]sulfate-labeled mouse entactin isolated by the same procedures from a PF-HR9 cell culture also demonstrated that the sulfate was present as tyrosine sulfate exclusively. Similar analyses of mouse entactin/nidogen showed tyrosine sulfation (Paulsson *et al.*, 1985). Paper electrophoretic analysis of an alkaline hydrolysate of [^3H]tyrosine-labeled glutactin showed that 8% of the radioactivity co-localized with a standard of tyrosine sulfate, 77% was tyrosine and 15% was present as an unidentified component. We conclude that about four of the 46 tyrosine residues of glutactin were

Fig. 6. Immunostaining of whole mount embryos. *Drosophila* embryos were stained with rabbit anti-glutactin antibody and with anti-rabbit IgG-fluorescein (A–E) or anti-rabbit IgG-horseradish peroxidase (G, H). The embryo in F is stained with anti-glutactin antibody and anti-rabbit IgG-rhodamine which labeled the dorsal median cell (dmc) red, and anti-horseradish peroxidase-fluorescein which labeled the commissures (c) green, though both appear bright in this black and white print. The abbreviations are: a = apodeme; amc = antenno-maxillary complex; b = brain; c = commissures; co = chordotonal organ; dmc = dorsal median cells; g = gut; gc = garland cells; nc = nerve cord; m = muscles; the arrowheads point to dorsal median cells and their processes. The bar represents 10 μ m.

sulfated when the cells were grown in DMEM medium containing 1.5 mM sulfate. The incorporation of sulfate into glutactin by Kc cells increased linearly when the sulfate concentration of the medium was increased from 1 to

500 μ M, but increased little further at higher concentrations (not shown). The rate of glutactin biosynthesis seemed to be independent of sulfate concentration.

Three tests showed that glutactin binds Ca^{2+} ions. The

glutactin band of SDS–polyacrylamide electrophoresis gels (Figure 1) reacts strongly with a stain for calcium-binding proteins (Campbell *et al.*, 1983) (not shown). When Western blots of partially purified media proteins were reacted with [^{45}Ca]Cl $_2$ in the presence of 5 mM MgCl $_2$ only glutactin gave a strong autoradiographic band, though laminin also bound some radioactive calcium (not shown). This binding was inhibited by 5 μM LaCl $_3$. Native glutactin was equilibrated with 4–640 μM [^{45}Ca]Cl $_2$, in the presence or absence of 5 mM MgCl $_2$ by microdialysis (Figure 5). This showed that 4–15 sites bound Ca $^{2+}$ with apparent K_D values in the range 30–100 μM independently of the presence of Mg $^{2+}$ ions, Triton X-100 and buffer containing up to 150 mM salt.

Immunofluorescence staining of *Drosophila* embryos with anti-glutactin antibodies

Immunofluorescence staining of whole *Drosophila* embryos (Figure 6) shows distinct outlines of the internal organs such as the gut which are enclosed by basement membranes. The brain, nerve cord and sensory bodies are also covered by a thin, strongly immunostaining basement membrane, and most outstanding is the presence of segmentally spaced channels through the nerve cord which are lined with a basement membrane. These show up as distinct points when the nerve cord is seen in a dorsal view. The channels are seen as they traverse the nerve cord when the embryo is rotated. Careful focusing onto the dorsal surface of the nerve cord reveals the presence of two dorsal median cells which have processes extending to the lateral muscles joining at the apodemes at the segmental borders. There are 11 sets of channels with pairs of dorsal median cells at their dorsal surface of the nerve cord. By focussing at different levels one can see processes extending to the muscle junctions in the abdominal, but not in the thoracic segments. These processes of the dorsal median cells are distinct from the segmental nerves, and lie between them. This was proven by double immunostaining of slightly older embryos with anti-glutactin, highlighting the dorsal median processes, and anti-horseradish peroxidase, staining the segmental nerves (not shown). The channels are located at the exact median position between the posterior commissures of one segment and the anterior commissures of the next segment. When the nerve cord is dissected out, the channels are clearly visible, and at the dorsal surface the cells are anchored and the broken lateral processes can be seen (not shown). The channels through the nerve cord are first observed in embryos of 7–9 h of development. They are not seen in 6–7 h embryos. The channels appear to form by closure of the basement membrane. When whole embryos or the dissected nerve cord were immunostained with antibodies to *Drosophila* laminin or collagen IV these channels along the nerve cord and the dorsal median cells and their processes were also stained. Pre-immune sera, on the other hand, gave no specific staining pattern and the entire embryos showed only a very weak overall fluorescence (not shown). The extension of these processes to the lateral muscles is also described by Caudy *et al.* (1986) in analyses of such *Drosophila* embryo fillets. These dorsal median cells were shown to be of mesectodermal origin by Beer *et al.* (1987). After immunostaining for glutactin they serve as excellent segmental markers along the nerve cord for positioning other antigens, e.g. the antennapedia gene product (Wirz *et al.*,

1986). When the nerve cord segmentation patterns are altered in the mutants *fushi tarazu*, *hairy*, *evenskipped* and *run1* the number and location of the nerve cord commissures are altered. Immunostaining for glutactin shows that the channels through the nerve cord and the dorsal median cells follow this altered structure in the mutant embryos. Immunofluorescence staining of sections of *Drosophila* embryos colocates glutactin with laminin (Fessler *et al.*, 1987) at sites where basement membranes have been demonstrated by electron microscopy. The basement membranes that envelop imaginal discs are prominently stained by antibodies against glutactin or laminin (Fessler *et al.*, 1987) (not shown).

Discussion

Glutactin is a new, sulfated and strongly anionic glycoprotein of *Drosophila* basement membranes. The mol. wt calculated from the peptide sequence is 117 kd. An additional contribution of carbohydrate is at least 7 kd. The excess of 52 acidic amino acid residues and about four *O*-sulfated tyrosines contribute to glutactin's high negative charge. While all the sulfate is attached to tyrosine, there could be other acidic carbohydrate groups. Like some other strongly charged polypeptides this protein behaves aberrantly in SDS–PAGE (Gerdes *et al.*, 1989). The glutactin sequence contains seven tyrosine residues that conform to a consensus sequence for post-translational sulfation (Bauerle *et al.*, 1988) and four of these are in the acidic carboxyl domain. However, we do not know which tyrosine residues are sulfated. Glutactin preferentially binds calcium ions in the presence of competing MgCl $_2$ with an apparent binding coefficient in the range 30–100 μM , but lacks an EF hand calcium-binding domain. Glutactin is slowly degraded by proteolysis into defined smaller molecules, and is probably more sensitive to proteolysis after calcium has been removed from it.

The nucleic acid sequence of glutactin is supported by three independently established amino acid sequences of the secreted polypeptide. The predicted 17 residue long signal peptide ends in a good consensus cleavage signal and is followed by the sequence found at the amino terminus of the secreted protein. The 1023 residues of glutactin are separated into two domains by a consecutive sequence of 13 threonine plus one serine residues, starting at position 603. Several *Drosophila* proteins contain similar homooligomeric stretches of amino acid residues (Fessler and Fessler, 1989). The carboxyl domain is highly acidic, and 44% of its residues are either glutamine or glutamic acid (Figure 6). The sequence of this carboxyl domain shows ample opportunity for α -helical folding, and has a low average hydrophobicity. In contrast, the amino domain has a predicted net charge near zero and contains many more hydrophobic residues (Figure 7). There are several patterns of the many E and Q residues in the carboxyl domain. The canonical sequence EREQQ is repeated, with variations, 17 times between positions 659 and 948, and several times a few copies are in contiguous succession. From positions 684 to 723 the canonical sequence QEER(I/L)RQQRE occurs four times in contiguous succession, with only three single amino acid substitutions.

A substantial segment of the amino domain has a strongly significant similarity of 19–33 standard deviation units with several serine esterases, but lacks the key, catalytically active serine residue (Table IA,B). When horse myoglobin and

Fig. 7. Amino acid composition of the amino and carboxyl domains of *Drosophila* glutactin. The hydropathy plot of glutactin is indicated by a solid line. Average hydropathy was determined over a window of 11 residues. A zero line for the hydropathy plot has been drawn through the amino domain. The relative content of Glu and Gln residues, summed over a window 11, is shown by a dotted line. The positions of basic and acidic residues are indicated by the histogram. The amino domain of glutactin corresponds to residues 18–602 and the carboxyl domain to residues 617–1023 of the unprocessed polypeptide.

horse α -hemoglobin were compared in the same way, a homology of 14 standard deviation units was obtained and, similarly, horse α - and β -hemoglobin chains showed a homology to 29 standard deviation units. Any value >10 standard deviation units strongly indicates relatedness. These significant similarities of glutactin are with several acetylcholine and butyrylcholine esterases, rat pancreatic lysophospholipase precursor and *Drosophila* esterase-6 (carboxylic-ester hydrolase). Glutactin has strong similarity with that portion of thyroglobulins which is homologous to the catalytic region of choline esterases, but like glutactin lacks the critical serine residue (Swillens *et al.*, 1986). The best local homology alignments of glutactin with these proteins cover much of the amino domain of glutactin (Table IA). Correspondingly, most of the lipase (12–571 out of the total 612 residues of the lipase) and the esterase (25–484 out of the total 548 residues of the esterase) are involved in these alignments. To obtain a measure of the potential three-dimensional similarity in the domains that span the serine site of serine esterases we calculated the hydrophobic correlation coefficient of Sweet and Eisenberg (1983) for pairs of proteins as shown in Table IB. All the coefficients were >0.4 , the minimum value for proteins with similar three-dimensional structure. In a region of ~ 150 amino acid residues of glutactin, which encompasses an alanine that replaces the canonical serine, the correlation coefficient with human acetylcholine esterase reaches 0.7. This value is similar to the correlation coefficient of the human and *Torpedo* acetylcholine esterases in this domain. For comparison, the hydrophobic correlation coefficient between horse α - and β -globin chains, which are of this size, is 0.68. All this suggests that the three-dimensional folding of a portion of glutactin is similar to that of serine esterases, and that much of the large amino domain of glutactin resembles them. Two common substrates for esterases, α -naphthol butyrate and *p*-nitrophenyl butyrate, were not hydrolysed by glutactin. Presumably this lack of catalytic activity is due to the lack of the critical serine residue, though glutactin might still maintain a non-catalytic binding site for related substrate/ligands.

Although optimal alignment of the amino acid sequences of glutactin and entactin indicates a region of 159 residues (residues 121–279 of Figure 4) with a similarity significant to six standard deviation units, the remaining sequences are quite different and glutactin lacks $>90\%$ of the cysteines

Table I. A. Local homologies of glutactin, thyroglobulins and serine esterases^a

	Homologous regions		Significance
	Test protein	Glutactin	(SD units)
Acetylcholine esterases			
<i>Torpedo</i> (X05497)	23–551	54–588	28
Human (A00772)	1–517	58–578	24
<i>Drosophila</i> (X05893)	153–605	134–588	19
Butyrylcholine esterase			
Human (M16541)	25–545	54–578	31
Carboxyl ester hydrolase			
<i>Drosophila</i> (M15961)	25–484	85–543	33
Lysophospholipase			
Rat (M15893)	12–571	78–643	27
Thyroglobulins			
Bovine (A01533)	2195–2697	53–571	24
Rat (A22016)	397–889	58–566	21

B. Hydrophobic correlation coefficients^b

	Glutactin		Acetylcholine esterase (X05497)	
	Coeff.	No. of residues	Coeff.	No. of residues
Acetylcholine esterase				
<i>Torpedo</i> (X05497)	0.50	313 ^d	0.58	543 ^c
<i>Drosophila</i> (X05893)	0.42	452 ^c		
Human (A00772)	0.50	313 ^d		
	0.73	150 ^e		
Thyroglobulin				
Bovine (A01533)	0.45	294 ^d	0.57	314 ^d
	0.54	152 ^f		
Lysophospholipase				
Rat (M15893)	0.47	461 ^c	0.54	445 ^c
	0.55	217 ^g	0.84	308 ^h
Carboxyl ester hydrolase				
<i>Drosophila</i> (M15961)	0.45	458 ^c	0.49	374 ^c
	0.51	217 ^g	0.57	280 ^h

^aThe computer program SEQHP was used to find regions of local homology between amino acid sequences of glutactin and proteins identified by their Genbank accession numbers. The amino acid residue numbers that demarcate each alignment are listed. The statistical significance of each alignment was determined with the program SEQDP set for 100 scrambles of each sequence.

^bHydrophobic correlation coefficients are shown for selected portions of glutactin and matching regions of several proteins. For comparison the coefficients of the equivalent regions of five of these proteins with *Torpedo* acetylcholine esterase are given. For each match the number of amino acid residues of glutactin or of *Torpedo* acetylcholine esterase are shown.

^cMatches were first obtained by the SEQHP program.

^dThe glutactin region (86–398) was further aligned with some proteins and portions with the least gaps were chosen. The same portions were compared with the *Torpedo* acetylcholine esterase.

^eA subset of d, of ~ 150 residues in glutactin region 132–282, gave the highest correlation coefficient with human acetylcholine esterase.

^fThe same region of glutactin as in e was compared with thyroglobulin.

^gAlignment with glutactin region 85–302.

^hOptimal fit for a subset of c that is approximately comparable to the residue length of d.

of entactin. We conclude that glutactin, entactin and rat aspartactin (Clegg *et al.*, 1988) are all different proteins. The sedimentation coefficient, s_{w20} , of *Drosophila* glutactin (1006 amino acid residues) is 4.1S and of mouse entactin (1217 residues) is 4.5S, and only their SDS electrophoretic mobilities are closely similar. Under our conditions of isolation, glutactin showed minimal or no interaction with laminin, as judged by sedimentation analyses and equilibrium gel filtration (Horwitz *et al.*, 1985) (not shown), while entactin binds strongly to laminin, and this complex sediments faster than free laminin.

The apposition of glutactin to muscle and nerve cells makes this calcium-binding, esterase-like protein interesting, but we neither know its function, nor whether it has any true catalytic activity. The immunofluorescent detection of glutactin at a variety of sites where basement membranes are known to occur in *Drosophila* causes us to regard it as a basement membrane protein. However, this does not exclude it from being a component of other extracellular matrix structures. It is likely that glutactin may also occur in vertebrates, and we are currently investigating this possibility.

Materials and methods

Isolation and characterization of glutactin

D.melanogaster Kc cells [line 7E10(0)] were grown and radioactively labeled in suspension culture in D22 medium without serum, and as described (Campbell *et al.*, 1987; Fessler *et al.*, 1987), the culture media proteins were partially separated by velocity sedimentation on a 5–20% sucrose gradient with a 60% sucrose pad in a Beckman SW41 rotor at 4°C at 39 000 r.p.m. for 19 h. The slower sedimenting fractions containing glutactin were chromatographed on a Mono Q anion exchange column (Pharmacia LKB Biotechnology) at 4°C in 2 M urea, 30 mM Tris–HCl, pH 8, 0.1% Triton X-100, using 0.1, 0.15 and 0.2 M NaCl, and then 0.5 M NaCl for elution of glutactin. Glutactin was passed first over a Superose 12 column and then a Superose 6 column (Pharmacia) in a 50 mM Tris–HCl pH 7.5, 150 mM NaCl, 500 mM sucrose and 0.1% Triton X-100 buffer. Sedimentation coefficients of *Drosophila* glutactin and [35 S]sulfate-labeled mouse entactin, isolated free of laminin from PF-HR9 cell cultures, were determined (Fessler and Fessler, 1974) after co-sedimentation in one tube of a SW60 Beckman rotor at 4°C and 56 000 r.p.m. for 20 h. For mol. wt estimation, denatured reduced glutactin in 6 M GuCl₂, 50 mM Tris–HCl pH 7.6, 0.1% Triton X-100 and 50 mM dithiothreitol was also passed through a Superose 6 column. Amino acid sequencing was according to Hunkapiller *et al.* (1983). Tyrosine sulfate standard preparation and analyses were as described (Hüttner, 1984; Fessler *et al.*, 1986). Enzymatic digestion with endoglycosidase F or H (DuPont NEN) was according to manufacturer's directions. To determine the COOH end of glutactin, separate solutions of carboxypeptidase A (A) and carboxypeptidase B (B) (Worthington Biochemicals) in 0.2 M triethylammonium bicarbonate, 1 mM phenylmethylsulfonyl fluoride, pH 8.0 (TEAB), were autolysed at 40°C for 60 min, then passed through PD-10 columns to remove free amino acids and concentrated. Then, 18 µg A was added to 5 nmol glutactin in TEAB. Aliquots were taken immediately and after 30 min digestion at 40°C, and stored in acidified methanol. Then, 7 µg A and 50 µg B were added and aliquots were taken after 1.5 min and 30 min further incubation. The free amino acids of each sample were determined.

Calcium binding of glutactin

SDS–PAGE gels were stained for calcium binding as described (Campbell *et al.*, 1983). Other gels were electroblotted onto nitrocellulose paper, equilibrated in 60 mM KCl, 5 mM MgCl₂, 10 mM imidazole-HCl, pH 6.8, then reacted with 1.6 µM CaCl₂ and 1 µCi/ml 45 Ca²⁺ in the same buffer and washed as described (Maruyama and Nonomura, 1984). After autoradiography the blots were stained with amido black to identify protein bands. Native, purified glutactin, 1–7 µM, in 0.05% Triton X-100, 60 mM KCl, 10 mM imidazole-HCl, pH 6.8, was dialysed against this buffer containing Chelex 100 resin (Bio-Rad) to remove divalent metals. Solutions of glutactin in 150 mM NaCl, 20 mM Tris–HCl, pH 7.4, with or without 0.05% Triton X-100 were similarly prepared. CaCl₂, containing a constant amount of 45 Ca²⁺, was added to both protein and buffer solutions to

4–640 µM, and in some experiments MgCl₂ to 5 mM. 25 µl of each solution were placed in opposite, equal chambers of a rotary equilibrium microdialyser (Hoefer Scientific). After 48 h equilibration at 7°C the radioactivities of protein and buffer solutions were determined and the results were analysed with the non-linear regression analysis program Enzfitter of R.J. Leatherbarrow. Similar experiments were also made with larger volumes of solution in DiaCell dialysis capsules (Instrumed Inc.). Reagent solutions, treated with Chelex resin, were kept in plastic containers and their Ca²⁺ ion concentrations were measured by atomic absorption.

Antibody preparation

Purified non-reduced glutactin was electrophoresed and the gel band was injected into rabbits. The IgG fraction of the antiserum was isolated and affinity purified on a glutactin–Sepharose column (Fessler *et al.*, 1987). The specificity of the antibodies was demonstrated by immune precipitation or on Western blots (Fessler *et al.*, 1987) using the Kc cell culture medium proteins or an SDS extract of 10–18 h *Drosophila* embryos.

Immunofluorescence of whole *Drosophila* embryos

Staged embryos grown at 25°C were whole-mount stained essentially as described (Wirz *et al.*, 1986). In double labeling experiments the embryos were also treated with goat anti-horseradish peroxidase linked to fluorescein (Organon Teknica-Cappel). The embryos were mounted in PBS–Gelvatol–polyvinylalcohol (Monsanto). Some embryos were stained with the biotinylated anti-rabbit IgG–avidin–peroxidase ABC Vectastain (Vector Labs). After color development the embryos were dehydrated and suspended in methyl salicylate for photography. Cryostat sections of *Drosophila* embryos or imaginal discs were stained as described (Fessler *et al.*, 1987).

Characterization of the glutactin cDNA and gene

A Kc cell cDNA library constructed in λ gt11 (Blumberg *et al.*, 1988) was screened with anti-glutactin antiserum and the ProtoBlot immunoscreening system (Promega). Except where indicated, all nucleic acid manipulations were performed according to standard protocols (Maniatis *et al.*, 1982). The cDNA inserts were subcloned into either pBluescribe or pBluescript KS+ (Stratagene Cloning Systems) for restriction mapping and further characterization. A Kc cell cDNA library constructed in λ gt10 (Blumberg *et al.*, 1987) was screened twice by plaque hybridization using RNA probes complementary to either all of clone A14 or the 5' 450 bp of clone H9. [α - 32 P]UTP-labelled RNA probes were synthesized from pBluescribe or pBluescript KS+ subclones according to Stratagene protocols. Northern blots of 20 µg total Kc cell RNA were hybridized with RNA probes complementary to clones A7, A12 or A14.

Each fusion protein was adsorbed to nitrocellulose filters from confluent lysed bacterial lawns. After 10 min incubation in 3 M urea, 10 mM MgCl₂, the filters were blocked for 30 min in 10 mM Tris–HCl, pH 7.4, 100 mM MgCl₂, 0.5% Tween-20, 1% BSA, 5% fetal calf serum. After incubation with 1:50 diluted antisera at 20°C for 1 h the filters were washed twice for 10 min with blocking solution. The antibodies were eluted with three washes for 1 min in 3 ml 0.15 M glycine pH 2.0 and collected into 2 ml 2 M Tris–HCl pH 8.0. The antibodies were dialysed at 4°C against 10 mM Tris–HCl, pH 8.0, 150 mM NaCl, 0.05% Tween-20, then diluted with an equal volume of blocking solution and stored at 4°C. The affinity-selected antibodies were reacted with Western blots of glutactin and fusion proteins.

Selected restriction fragments were subcloned into either M13mp18, M13mp19 or Phagescript (Stratagene) and sequenced by the dideoxy chain termination method using [α - 35 S]dATP and either Klenow (Pharmacia) or modified bacteriophage T7 polymerase (Sequenase, U. S. Biochemical). The clones were sequenced on both strands and across all restriction sites used in subcloning using described methods (Blumberg *et al.*, 1988).

D.melanogaster genomic DNA was prepared from embryo nuclei (gift of A.J. MacKrell) (Brutlag *et al.*, 1977). Genomic Southern blotting was performed with 15 µg aliquots of genomic DNA, digested with either *Eco*RI or *Bam*HI, capillary blotted onto Hybond C+ (Amersham), and hybridized with [α - 32 P]dCTP oligo-labelled (Pharmacia) *Eco*RI restriction fragments of cDNA clones A7 and H9A. The *Drosophila* genomic cosmid library JT101 (kind gift of Dr John Tamkun and Dr Matt Scott) was screened by colony hybridization using a probe similar to that used for the genomic Southern. Selected restriction fragments of hybridizing cosmid clones were subcloned into pBluescribe for further characterization. The method of *in situ* hybridization with biotinylated DNA probes to chromosomes was as described (Langer-Safer *et al.*, 1982).

Acknowledgements

We appreciate helpful discussions with A.J. MacKrell, Dr D. Eisenberg, Dr W.B. Hüttner, Dr J. Engel, Dr R. Doelz, and the advice of Ms L. Williams

on amino acid sequence analyses. We thank Dr W. Gehring for his hospitality during an enjoyable visit to his laboratory, where we learned whole mount immunostaining techniques, and thank Dr M. Noll for mutant embryos. This research was supported by USPHS research grant AG-02128 and by a grant from the Muscular Dystrophy Association. PFO was supported by National Research Service Award GM-07185.

References

- Baeuerle, P.A., Lottspeich, F. and Hüttner, W.B. (1988) *J. Biol. Chem.*, **263**, 14925–14929.
- Beer, J., Technau, G.M. and Campos-Ortega, J.A. (1987) *Wilhelm Roux's Arch. Dev. Biol.*, **196**, 222–230.
- Blumberg, B., MacKrell, A.J., Olson, P.F., Kurkinen, M., Monson, J.M., Natze, J.E. and Fessler, J.H. (1987) *J. Biol. Chem.*, **262**, 5497–5500.
- Blumberg, B., MacKrell, A.J. and Fessler, J.H. (1988) *J. Biol. Chem.*, **263**, 18328–18337.
- Brutlag, D., Appels, R., Dennis, E.S. and Peacock, W.J. (1977) *J. Mol. Biol.*, **112**, 31–47.
- Campbell, A.G. (1986) Ph.D. Thesis, University of California, Los Angeles.
- Campbell, A.G., Fessler, L.I., Salo, T. and Fessler, J.H. (1987) *J. Biol. Chem.*, **262**, 17605–17612.
- Campbell, K.P., MacLennan, D.H. and Jorgensen, A.O. (1983) *J. Biol. Chem.*, **258**, 11267–11273.
- Caudy, M., Jan, Y.N. and Jan, L. (1986) *Soc. Neurosci. Abstr.*, **12**, 196.
- Clegg, D.O., Helder, J.C., Hann, B.C., Hall, D.E. and Reichardt, L.F. (1988) *J. Cell Biol.*, **107**, 699–705.
- Durkin, M.E., Chakravarti, S., Bartos, B.B., Liu, S.-H., Friedman, R.L. and Chung, A.E. (1988) *J. Cell Biol.*, **107**, 2749–2756.
- Eley, M.H., Burns, P.C., Kannapell, C.C. and Campbell, P.S. (1979) *Anal. Biochem.*, **92**, 411–419.
- Fessler, L.I. and Fessler, J.H. (1974) *J. Biol. Chem.*, **249**, 7637–7646.
- Fessler, J.H. and Fessler, L.I. (1989) *Annu. Rev. Cell Biol.*, **5**, 309–339.
- Fessler, J.H., Lunstrum, G., Duncan, K.G., Campbell, A.G., Sterne, R., Bächinger, H.-P. and Fessler, L.I. (1984) In Trelstad, R. (ed.), *The Role of Extracellular Matrix in Development*. Alan R. Liss, NY, pp. 207–219.
- Fessler, L.I., Brosh, S., Chapin, S. and Fessler, J.H. (1986) *J. Biol. Chem.*, **261**, 5034–5040.
- Fessler, L.I., Campbell, A.G., Duncan, K.G. and Fessler, J.H. (1987) *J. Cell Biol.*, **105**, 2383–2391.
- Gerdes, H.H., Rosa, P., Phillips, E., Bauerle, P., Frank, R., Argos, P. and Hüttner, W.B. (1989) *J. Biol. Chem.*, **264**, 12009–12015.
- Hay, E.D. (1981) *Cell Biology of the Extracellular Matrix*. Plenum Press, NY.
- Hogan, B.L.M., Taylor, A., Kurkinen, M. and Couchman, J.R. (1982) *J. Cell Biol.*, **95**, 197–204.
- Horwitz, A., Duggan, K., Greggs, R., Decker, C. and Buck, C. (1985) *J. Cell Biol.*, **101**, 2134–2144.
- Hunkapiller, M.W., Hewick, R.M., Dreyer, W.J. and Hood, L.E. (1983) *Methods Enzymol.*, **91**, 399–413.
- Hüttner, W.B. (1984) *Methods Enzymol.*, **107**, 201–223.
- Keller, E.B. and Noon, W.A. (1985) *Nucleic Acids Res.*, **13**, 4971–4981.
- Kozak, M. (1986) *Cell*, **44**, 283–292.
- Langer-Safer, P.R., Levine, M. and Ward, D.C. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4381–4385.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Mann, K., Deutzmann, R., Aumailley, M., Timpl, R., Raimondi, L., Yamada, Y., Pan, T.C., Conway, D. and Chu, M.L. (1989) *EMBO J.*, **8**, 65–72.
- Maruyama, K. and Nonomura, Y. (1984) *J. Biochem.*, **96**, 859–870.
- Mount, S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
- Olsen, D.R., Nagayoshi, T., Fazio, M., Mattei, M.G., Passage, E., Weil, D., Timpl, R., Chu, M.L. and Uitto, J. (1989) *Am. J. Hum. Genet.*, **44**, 876–885.
- Olson, P.F., Sterne, R., Fessler, L.I. and Fessler, J.H. (1987) *J. Cell. Biochem., Suppl.* **11C**, 26.
- Paulsson, M., Dziadek, M., Suchanek, C., Hüttner, W.P. and Timpl, R. (1985) *Biochem. J.*, **231**, 571–579.
- Perlman, D. and Halvorson, H.O. (1983) *J. Mol. Biol.*, **167**, 391–409.
- Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature*, **263**, 211–214.
- Sweet, R.M. and Eisenberg, D. (1983) *J. Mol. Biol.*, **171**, 479–488.
- Swillens, S., Ludgate, M., Mercken, L., Dumont, J.E. and Vassart, G. (1986) *Biochem. Biophys. Res. Commun.*, **137**, 142–148.
- Wirz, J., Fessler, L.I. and Gehring, W.J. (1986) *EMBO J.*, **5**, 3327–3334.

Received November 3, 1989; revised January 23, 1990