

Identification of a *Drosophila* Vitamin K-dependent γ -Glutamyl Carboxylase*

Received for publication, March 1, 2000, and in revised form, March 23, 2000
Published, JBC Papers in Press, March 29, 2000, DOI 10.1074/jbc.M001790200

Tao Li, Chao-Tsung Yang, Dayun Jin, and Darrel W. Stafford‡

From the Department of Biology, Center for Thrombosis and Homeostasis, University of North Carolina, Chapel Hill, North Carolina 27599-3280

Using reduced vitamin K, oxygen, and carbon dioxide, γ -glutamyl carboxylase post-translationally modifies certain glutamates by adding carbon dioxide to the γ position of those amino acids. In vertebrates, the modification of glutamate residues of target proteins is facilitated by an interaction between a propeptide present on target proteins and the γ -glutamyl carboxylase. Previously, the gastropod *Conus* was the only known invertebrate with a demonstrated vitamin K-dependent carboxylase. We report here the discovery of a γ -glutamyl carboxylase in *Drosophila*. This *Drosophila* enzyme is remarkably similar in amino acid sequence to the known mammalian carboxylases; it has 33% sequence identity and 45% sequence similarity to human γ -glutamyl carboxylase. The *Drosophila* carboxylase is vitamin K-dependent, and it has a K_m toward a model pentapeptide substrate, FLEEL, of about 4 mM. However, unlike the human γ -glutamyl carboxylase, it is not stimulated by human blood coagulation factor IX propeptides. We found the mRNA for *Drosophila* γ -glutamyl carboxylase in virtually every embryonic and adult stage that we investigated, with the highest concentration evident in the adult head.

The γ -glutamyl carboxylase catalyzes the post-translational modification of specific glutamates to γ -carboxyglutamate (Gla)¹ in a number of proteins. To accomplish this reaction, the γ -glutamyl carboxylase requires, in addition to its peptide substrate, reduced vitamin K, oxygen, and carbon dioxide as co-substrates. In vertebrates, it appears that the recognition between the nascent protein substrate and enzyme is dependent upon a “docking” reaction between the propeptide of the substrate and the enzyme. In the mollusk *Conus*, a propeptide also appears to dock the substrate to the carboxylase, but the substrate is quite different in sequence than that of the vertebrate propeptides (1).

Gla was first identified in prothrombin as a modified amino acid necessary for the activity of this vitamin K-dependent protein. Historically, Gla has been associated with coagulation factors. However, in recent years, other Gla proteins with varied functions and locations have been identified: two proteins,

bone Gla protein, or osteocalcin, and matrix Gla protein, have been isolated from bone (2); the Gla protein Gas 6 is known to be a ligand for the receptor tyrosine kinase Axl (3); and two proline-rich Gla proteins, PRGP1 and PRGP2, whose functions are not known, are found in the spinal cord and thyroid gland, respectively (3).

Drosophila melanogaster (the fruit fly) has been used as a model organism because of its relatively small genome, short growth cycle, and long history of study. The *Drosophila* genome, with approximately 165 million bases, is about 20-fold smaller than that of mammals. Furthermore, the sequence of the *Drosophila* genome has been completed and will soon be available through GenBankTM. One of the most remarkable and valuable characteristics of *Drosophila* as a model organism is the fact that many of its genes have functional counterparts in humans. For example, the homeobox genes of *Drosophila* have been used to find their mammalian homologues. In the case of the *Drosophila* *Deformed* gene, the human homologue, *HOX4B*, can functionally replace its *Drosophila* counterpart (4).

Using the human γ -glutamyl carboxylase (hGC) amino acid sequence as query, we found a region on *Drosophila* chromosome 3L with strong similarity to human γ -glutamyl carboxylase. We report here the isolation of a cDNA corresponding to this gene. When expressed in *Spodoptera* insect cells, it produces vitamin K-dependent carboxylase activity. Mammalian substrates are not efficiently carboxylated by *Drosophila* γ -glutamyl carboxylase (dGC), however, which suggests differences in recognition specificities between *Drosophila* and mammalian substrates.

EXPERIMENTAL PROCEDURES

GenBankTM Search for Vitamin K-dependent γ -Glutamyl Carboxylase Homologues—The amino acid sequence of hGC was used to search GenBankTM.

RT-PCR—Forward primer 1140-F (CTTCATCACCAAGGGCTATA) and reverse primer 1842-R (CAGCATCGTTTTGTTGGTGT) were selected to avoid introns and to give a fragment of about 700 bp (ordered from Life Technologies, Inc.). 100 mg of adult *Drosophila* in 1 ml of TRIZOL reagent (Life Technologies, Inc., catalogue number 15598) was ground in a mortar, and total RNA was isolated according to the manufacturer's protocol. Total RNA was further treated by RQ1 RNase-free DNase (Promega, catalogue number TB518). First strand cDNA was synthesized with SuperScript II RNase H reverse transcriptase (Life Technologies, Inc., catalogue number 18064-014) with oligo(dT) 12–18 as primer (Life Technologies, Inc., catalogue number 18418–012). PCR reaction conditions were: predenaturing the template at 94 °C for 2 min; 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min for 30 cycles; followed by 72 °C for 5 min.

*RT-PCR Using a *Drosophila* Expression Panel*—*Drosophila* Rapid-ScanTM gene expression panel (catalogue number DSCC-101) was purchased from OriGene Technologies, Inc. (Rockville, MD). The product contains first-strand cDNAs prepared from different *Drosophila* tissues and developmental stages. The 12 cDNAs have been normalized against the transcript for RP49 (a constitutively expressed ribosomal protein

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280. Tel.: 919-962-0597; Fax: 919-962-9266; E-mail: dws@email.unc.edu.

¹ The abbreviations used are: Gla, γ -carboxyglutamic acid; hGC, human γ -glutamyl carboxylase; dGC, *Drosophila* γ -glutamyl carboxylase; PCR, polymerase chain reaction; RT, reverse transcriptase; bp, base pair(s).

gene), serially diluted over a 4-log range, arrayed onto a 48-well PCR plate, and dried. PCR conditions were: predenaturing the template at 94 °C for 3 min; 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 2 min for 35 cycles; followed by 72 °C for 5 min.

Screening of the *Drosophila* Adult Head cDNA Library—The *Drosophila* cDNA libraries were a generous gift from the Ruben Lab (BDGP EST Project, University of California, Berkeley, CA). The adult head

library was screened with the 700-bp *Drosophila* DNA fragment as probe (5). The positive clones were further characterized by sequence analysis.

Expression of γ -Glutamyl Carboxylase in High Five Cells—The full-length cDNA encoding the dGC was cloned into the pVL1392 vector; the recombinant clone with correct orientation and BacVector 3000 were co-transfected into SF9 cells. Plaques of recombinant virus were isolated via PCR using the same primer pair which was used for RT-PCR. The selected viruses were used to infect additional cells, which were checked for γ -glutamyl carboxylase activities. Plaques identified to contain recombinant virus were grown and then titered according to the manufacturer's instructions. Expression of the dGC was done by infection of $\sim 2 \times 10^6$ /ml High Five Cells with the recombinant virus at a multiplicity of infection of ~ 1 . After 48 h cells were collected by centrifugation and stored at -80 °C.

Preparation and Assay of Microsomes from High Five Cells—A total of 1.6×10^6 cells from 1.0 liter of culture expressing the recombinant *Drosophila* carboxylase was washed twice with Buffer A (20 mM phosphate, pH 7.4, 150 mM NaCl, 1 \times protease inhibitor mixture (6), and 15% glycerol) and resuspended in 100 ml of Buffer A. The sample was homogenized with 15 strokes using a Dounce homogenizer and then sonicated with four 5 s pulses using an Ultrasonic Heat Systems sonicator. Cellular debris was removed by centrifugation at $4000 \times g$ for 15 min, and the supernatant was centrifuged at $105,000 \times g$ for 1 h. The microsomal pellet was resuspended in 20 mM phosphate, pH 7.4, 500 mM NaCl, 1 \times protease inhibitor mixture, and 15% glycerol and stored

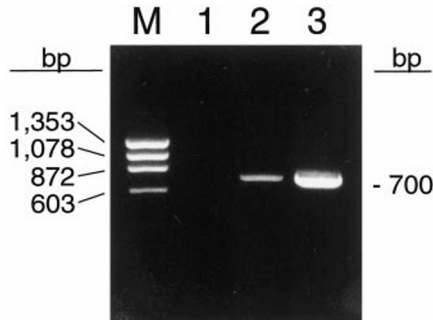


FIG. 1. PCR of total mRNA isolated from adult *Drosophila*. Lane M, PhiX174/*Hae*III DNA maker; lane 1, total RNA as PCR template (negative control); lane 2, first strand cDNA as template; lane 3, *Drosophila* genomic DNA as template (positive control).

```

cattcagattcagatcgaagatataatataatgcaatttgcgttttgcagcggcggcaactcccagttgggatgtccgcggcccgagtg
      predicted promoter
gcagttatcacaggttaccacacagcttccacacggacaggactagagacgaaactagaggaaacgaagctgggcaaatccogttgcaataa
taactcaotgacggagtataatggatgtaacatttctgaattgactctattttttacacogattccgaacgaatatttccgtggactcca
cgccagcgatccgacagaaacccgtagatattacgcaagagtcagcaaacgaaatgtaaaacacaaatattatattatgatacaaaoc
aatatgcaotgcaatgatttctgcaaaattttcaataaatttcaaaaacttaatacaaaacacacattogaggggattgatctttgggt
catctatgggcagatctgtaaggataaogtttactacttgcacaaaggacgtgttttttgcattctcaagattcaagattcaagcaga
      mRNA start site
atggccaaactccagcgaagctgccacgaaaccagtgccacagaggaogctactccgcacaagatacacotaggctogtgggaaga 90
M A N S K R K L P T K P S A T E D A T P H K D T P R S S E E
ttccocacagatgtaaatccaaagagtgccaaagagtocttctccaaacogattttctgagattctctcggatccgatctgagc 180
F P T D V N H K S A K K S F L P Q N A F F R D C L G H D L S
aaottcaotgacgtgacaaatttccagctggtaaacagacccgtagatggagctgocctgggcatvttccgtctgctotacggagca 270
N F T S L T N F T S W L N R P V D G A A L G I F R L L Y G A
gccatgctgattgacatccagagggaggtggcggcagctggagctgggtttccaggacccgtacacotgccactttccactgttc 360
A M L I D I A E E R G G G Q L D V R F Q E P L H C H F P L F
aacgggatgagcagctggatccactgagctgggtgctgctatctgtgcaatggctgggctggggtccatgctgggtctatoga 450
N G M R A L D Y P L M G C V Y L C M W L G A W G I M L G Y R
ttcaggaatagctgctggcctctcgtacacactggtacacttctgctggcaaacccagcgtggaaataaacacagctatctat 540
F R T S C L A F L I P Y W Y I F L L D K P T W N N H S Y L F
ggcctggggaaactgtgctcttccacaaagcggagtcactgcttccctggacagctggctaaatccccgaagagtcactgtatt 630
G L V G T L L L F T Q A E R S Y C S L D R W L N P R K S H C I
      position of intron 1
ccctactggaactacttcaatcaaatccagttcttctcactgtaactgtaacagagctcgaagaattctogctggaatgggatog 720
P Y W N Y F L I K F Q F F I L Y M Y A G L K K F S L E W L S
ggatacgcgagtgccaatctgagttaccactgggtctttgocccgtccggcaagtgatgtagccagagatccagcctgactcgtg 810
G Y A M S N L S Y H W V F A P F R Q V I D P E L I D L L I V
cattggttccacgctctctcgaagctgctcattgctttttctcgaatctggaaaagagccctgacccctcttctcagctcag 900
H W F T A F F D V S I A F F M T L E K T R L L V T P F M L S
ttctctcaatgaactccagcctggttgaatagggatggtttccctgggtgctgctggcgaagtcagcctcttctcagctcagctg 990
F H L M N S R L F V I G M F P W V C L A E V P L F F S F D W
      position of intron 2
ccgaggagcttaggcaactggtccaaatacaggaagctccatccaaagatccagaggaagagcttccaaagccgggaactcgtggc 1080
P R R L G N W S K I Q D V H P K D P E E E L S K P E E L C L R A
ctgocagcggcctcactcctgactggaactcctcctgacatactcactcctcctcctcctcctcctcctcctcctcctcctcct 1170
L R T G L I L A Y C G L Q L F L P Y S H F I T K G Y N N W T
      Primer 1140-Forward
aatgggctgtagcgtactcagtgaggacatgagtgcaactcactcagacacogctgcaagcactcactcaggtggtagcaacagagagcgc 1260
N G L Y G Y S W D M M V H S Y D T L Q T S I Q V V D N E S R
caagtgcacaactgaatccatacggcttccagagtagcagccgctggagcaagtaacgacagatggcagtgcaatgacaggtgcatc 1350
Q V H N L N P Y A F T E Y D R W T K Y A D M A V A R C I
gaggagaatttcaagagaaacagccacatattgggaggaacatttccactcactcctgacatctggctccatgacaggaagcttcaag 1440
E E N L Q E K Q P H I G R N I S I Y F D I W C S M N G R F T C
caacgatcattcagatccacgagagatctcctccgggcaagtggtcgcctttgagagacactcctggtgctgctctgctgacagag 1530
Q R S F D P R E D L L R A K W S P F E S T S W S L P L L N E
otgaaocacatgcccacaattgcccacactgagactgaggtgctggcctggaacaaactcactccagatgcaattctgtggcggatttc 1620
L N H M R P K L R T I E T E V L A W N N Y S D V I F V A D F
ccggactaaotcgaacaaotcactcctcggcagcacttcaactgcacgctgacacttcttgaggaatgctcggtagccaggaacagag 1710
P G L T L T N F I S P D L I N C T L T I L E G N V R Y R N E
cgggatccggagcctattcctgacagcggcgaagagatggtctggagagcaacatccccacttggtcaccacaaactggccagaag 1800
R D P E A Y F L T A G K S I G L E S N I T H L V T T I G Q K
ccagcttgatctgttccacatacccaaaaaacagatgctggagcagggcatcaogtagaacagctgatcaocagagagcgtctgctc 1890
P A S Y L F T Y T N K T M L E Q G I T I E Q P D T E E R S V
      Primer 1842-reverse
ctgcccgtgtagcaggttccagcctcgggttaccacaaocacagcagtttttggccacattggcaacogtcaatgtaactgtctctac 1980
L P L W Q E F Q H R V T N Y Q Q F L G H I G N C L M Y I L Y
gatgtgocattccgagccgtaaagggggagattgagaggagaggtcaaacatccgcaotatttccgttttagccgtatctagattc 2019
D V F I P Q A V K G G D Stop codon
gcaaatccaaagatataaatctggttaacotataatgcoctataaatcagaggataaacaatttatattgagaatgaaggtccagatt
gaaggatatacaacaatcctctggttttctctataatcaaatgcaatgcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca
cgctcaaatcaatcaatcctgcaaaaacacccctggaacacagagcagatataacagatgtaggggtgataaaatgataataatgttc
catattgtaccac
    
```

FIG. 2. Full-length *Drosophila* carboxylase cDNA and its amino acid sequence.



Fig. 3. Alignment of the amino acid sequence of γ -glutamyl carboxylase with other known carboxylase amino acid sequences. From top to bottom, human, mouse, bovine, and *Drosophila* sequences are shown. "Similar" residues are assigned based on a PAM250 matrix (8). Those residues called "identical" refer to cases where three or more of the aligned amino acids are identical. "Completely conserved" refers to those cases where the amino acids are identical in all five sequences. The three completely conserved cysteines are shown in violet.

at -80°C . Carboxylase activity was assayed as described previously (7).

RESULTS

Identification of the Drosophila Homologue to Human γ -Glutamyl Carboxylase—A "tblastn" search of GenBank™, using the hGC amino acid sequence as query, identified a 1385-bp putative homologue of γ -glutamyl carboxylase in the region 62A10 of *D. melanogaster* chromosome 3L. In many of the cases where there were conservative amino acid differences between *Drosophila* and human sequences, the same position was variable in the γ -glutamyl carboxylases of other species as well.

To determine whether the γ -glutamyl carboxylase homologue was expressed in *Drosophila*, we performed RT-PCR to amplify a 700-bp sequence from total *Drosophila* adult mRNA (Fig. 1, lane 2). To demonstrate that trace genomic DNA was not responsible for the positive result, there was no PCR product when total RNA was used as a template as shown in Fig. 1 (lane 1). These results indicate that this γ -glutamyl carboxylase-like sequence is expressed in adult *Drosophila*.

cDNA Library Screening—To determine which of the three libraries available to us (embryonic, larval/pupal, or adult head) contained our target gene, we performed PCR reactions using $\sim 10^5$ clones from each cDNA library as template and the same primers (see above) for RT-PCR. The expected 700-bp band was amplified from the adult head library but not from the embryo or larval cDNA libraries (data not shown). Seven positive clones were obtained from the 7×10^5 clones that were screened; of these, the two largest clones were chosen for sequence analysis.

Sequence of the Drosophila γ -Glutamyl Carboxylase Gene—The full-length carboxylase cDNA is about 2300 nucleotides, which predicts a protein of 672 residues and a deduced molecular mass of 78.4 kDa (Fig. 2). Using the alignment tool CLUSTALW in Biology WorkBench, we did multiple sequence alignment for *Drosophila*, mouse, rat, bovine, and human carboxylases (Fig. 3). When the translated *Drosophila* sequence was compared with human, the alignment found 33% identical amino acids and 45% similar amino acids. Those amino acids

designated as "similar" were based on PAM250 setting (8). There are 13 cysteines in dGC compared with the 10 cysteines in hGC. Three of the cysteines were found to be aligned in all of the known γ -glutamyl carboxylases (Fig. 3), thus making it highly likely that they preserve a critical structure or have an important catalytic function. Compared with the human carboxylase gene, the dGC gene contains two very short introns, which correspond to introns 4 and 7 of the human carboxylase gene. In addition, a consensus promoter sequence is present at 525 bp upstream from the first ATG coding for the amino-terminal methionine.

The dGC gene is located in *D. melanogaster* chromosome 3L region 62A10. P1 clone DS02777 contains the dGC gene. Using the MapViewer in Berkeley *Drosophila* Genome Program site, we found that there is a P-element (EP (3)0304) located about 2,000 bp upstream of the 5' end of dGC, but no P-element inside the dGC gene has been reported. Moreover, there is a P-glyco-

protein/multidrug resistance protein (Mdr50) gene about 500 bp away from the 3' end of the dGC gene.

Expression Profile of *Drosophila* γ -Glutamyl Carboxylase—To determine where and when the dGC is expressed, we used the same PCR primers to amplify the 700-bp fragment from *Drosophila* cDNA prepared at different stages of development and diluted over a 3-log range. Each cDNA was diluted in water to a series of three concentrations (labeled 1 \times , 0.1 \times , and 0.01 \times), with the highest concentration (1 \times) being approximately 1 ng of cDNA/well. Fig. 4 demonstrates that γ -glutamyl carboxylase is expressed at least 10 times more in the adult head than in any other tissue and that the first measurable expression is after 4 h of embryonic development.

***Drosophila* Carboxylase Activity**—To examine whether the *Drosophila* homologue to hGC has vitamin K-dependent activity, we expressed the dGC in SF9 cells and examined its activity in an *in vitro* assay. Fig. 5 shows that, in the presence of vitamin KH₂, carboxylase activity is 11-fold higher in dGC transfected cells compared with the mock transfected SF9 cells. Using kinetic studies, the K_m of dGC toward FLEEL was determined to be 4 mM (Fig. 6), which is similar to that of hGC. Through a time course, we found that the carboxylase activity for dGC was linear for at least 2 h (data not shown). It should also be noted that the propeptide of human factor IX fails to activate carboxylation of FLEEL by dGC (data not shown).

DISCUSSION

There are several interesting implications resulting from our observation that γ -glutamyl carboxylase is present in *Drosophila*. First, this discovery should contribute to the understanding of the evolutionary relationship among the coelomate phyla. Second, it may be a key to clarifying the relationship between structure and function in hGC. Third, the genetics of *Drosophila* may facilitate the discovery of additional properties of various important Gla proteins and, furthermore, may promote the identification of the vitamin K-epoxide reductase enzyme.

It is now clear that functional γ -glutamyl carboxylase enzymes exist in the phyla *chordata*, *mollusca* (7), and *arthropoda*. For convenience, a simplified evolutionary tree outlining the phyla discussed in this paragraph is shown in Fig. 7. From the phylogenetic tree it is apparent that the carboxylase arose in evolution before the protostome/deuterostome split. Thus, one would expect to find the γ -glutamyl carboxylase in the *annelida* and *echinodermata* and in some more primitive

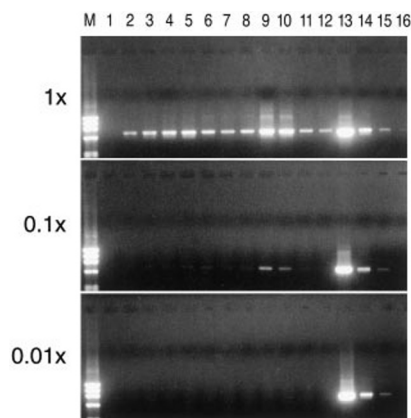
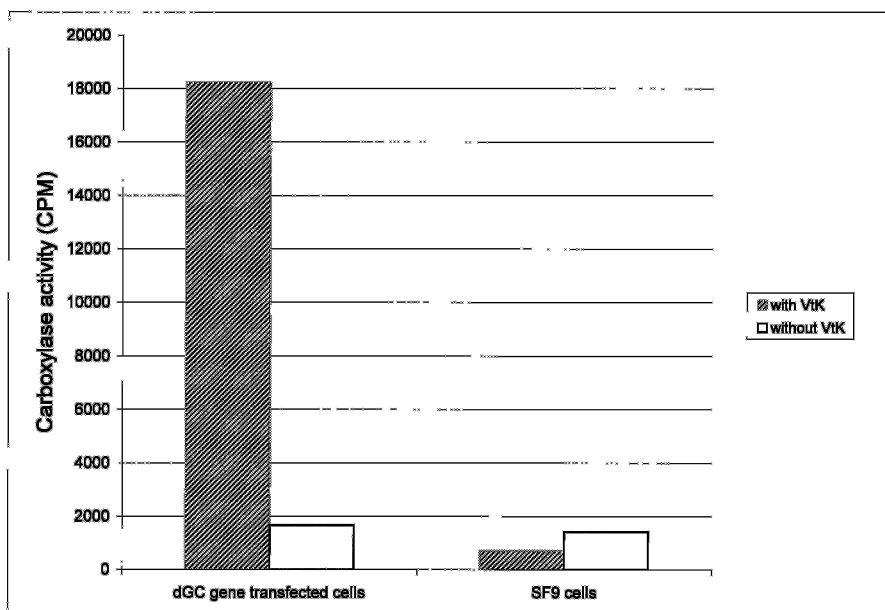


FIG. 4. *Drosophila* carboxylase gene expression level in different developmental stages and different tissues. Three different cDNA concentrations were used: the reverse transcript (1 \times), a 1:10 dilution (0.1 \times), and a 1:100 dilution (0.01 \times). The amount of cDNA in each well at the highest concentration is approximately 1 ng. Lane M, PhiX174/HaeIII DNA maker; lane 1, embryo 0–4 h; lane 2, embryo 4–8 h; lane 3, embryo 8–12 h; lane 4, embryo 12–24 h; lane 5, first instar; lane 6, second instar; lane 7, third instar; lane 8, pupae; lane 9, male head; lane 10, female head; lane 11, male body; lane 12, female body; lane 13, plasmid pSK-Dros700 0.1 pg as template; lane 14, plasmid pSK-Dros700 0.01 pg as template; lane 15, plasmid pSK-Dros700 0.001 pg as template; lane 16, plasmid pSK-Dros700 0.0001 pg as template.

FIG. 5. Carboxylase activity of SF9 cells transfected with the *Drosophila* carboxylase gene. The concentration of the pentapeptide (FLEEL) was 2.4 mM. Incubation was at 20 °C for 1 h. The bars labeled dGC gene transfected cells represent the activity of cells transfected with the dGC gene. Those labeled SF9 cells are wild-type SF9 cells. The shaded bars represent carboxylase activity with vitamin K; the white bars represent carboxylase activity without vitamin K.



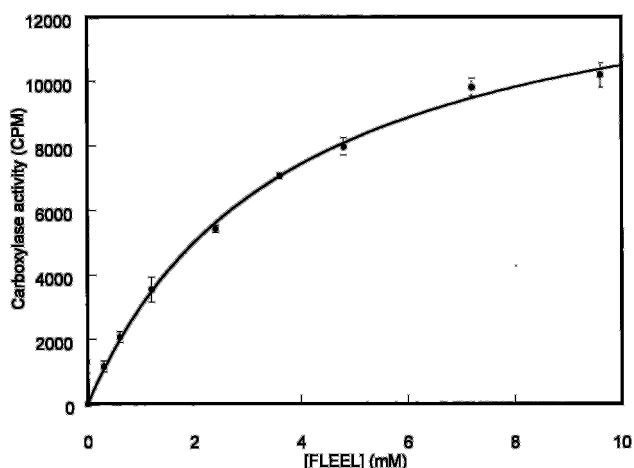


FIG. 6. **Drosophila carboxylase activity.** Solubilized dGC microsomes ($40 \mu\text{l}$) were mixed with various concentrations of FLEEL (0–9.6 mM) and were incubated at 20°C for 1 h. The data were fit to the Michaelis-Menten equation. Points are the averages of three experiments. $K_m = 3.8 \pm 0.3 \text{ mM}$. $V_{\text{max}} = 14,500 \pm 400 \text{ cpm}$.

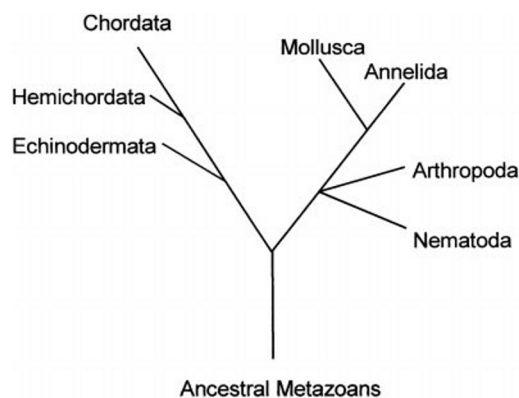


FIG. 7. **Simplified phylogenetic relationships of the species discussed.**

organisms as well. Further research to determine where γ -glutamyl carboxylase first appears in evolution will be important and interesting. A search of the complete genomic sequence of a member (*Caenorhabditis*) of the pseudo-coelomate phylum *nematoda* revealed no significant homology to γ -glutamyl carboxylase. This lack of a homologue to the carboxylase is inconsistent with the notion that nematodes (9) are closely related to the arthropods. This is probably the most interesting evolutionary aspect of our observations. However, one must remember that each organism is only one representative of thousands of diverse organisms within its phylum and that *Caenorhabditis* is rapidly evolving and highly specialized; therefore, it is possible other nematodes may possess γ -glutamyl carboxylase, whereas members of the phyla *arthropoda* and *mollusca* may lack it.

The comparison of sequences of γ -glutamyl carboxylase from different species has additional significance in understanding the relationship between the structure and function of the enzyme. For example, one would expect that the regions of the protein with the most sequence identities would have some crucial purpose relating to either the structure or function of the γ -glutamyl carboxylase. Thus, the substrate recognition site should be conserved because glutamic acid must be recognized by all species. Similarly, one would expect the vitamin K-binding site to be conserved. The expectation that the vitamin K site is conserved is less certain, however, because *Drosophila* may use a slightly different vitamin K hydroquinone.

Vitamin K_1 hydroquinone does stimulate carboxylation of dGC; therefore, if a different quinone is used, it must be very similar to vitamin K_1 . In the mammalian enzyme, substrate affinity appears to be a result of its propeptide docking it to the enzyme rather than the affinity of the Gla domain *per se*. However, the dGC appears not to recognize a mammalian propeptide and in this respect is similar to the carboxylase of *Conus*. It has been shown that the mammalian propeptide is either poorly recognized or not recognized at all by the *Conus* carboxylase. Nevertheless, a hydrophobic propeptide different from the commonly known mammalian propeptides appears to be present in the substrate of the *Conus* carboxylase (1, 10). Therefore, one might expect to find a propeptide-binding site on the dGC that is not homologous to the hGC-binding site but that is located in the same region we previously determined to be the location of the mammalian site (11). One major difference between the *Drosophila* and mammalian γ -glutamyl carboxylases is that the dGC is 86 amino acids shorter than the human enzyme from the carboxyl terminus. Notably, this is approximately the number of amino acids that Roth *et al.* (12) found could be removed from the carboxyl terminus of hGC without seriously affecting its enzymatic activity. A comparison of conserved hydrophobic amino acids between these species is also instructive. It has been known, because a number of related proteins were crystallized, that the packing of amino acids in the core of a protein must be conserved for enzyme stability (13). Also, it has since become clear that many protein-protein interactions take place through hydrophobic patches on the surface of a protein (14). Thus, examining the stability and enzymatic activity of mutations in conserved hydrophobic regions is expected to reveal whether these are surface regions or buried ones and whether they are important in particular substrate or peptide interactions.

It is interesting that the *Drosophila* head has the highest level of γ -glutamyl carboxylase activity. Although Gla proteins traditionally have been associated with coagulation, there has been considerable evidence for several years that certain Gla proteins serve other important functions. For example, it has been recognized since the mid-1970s that warfarin therapy during pregnancy results in characteristic embryopathies, including hypoplasia and a phenotype similar to chondroplasia punctata. In addition, blindness and mental retardation occur in a significant number of cases (15). The importance of carboxylation in bone was implied by the isolation of osteocalcin and matrix Gla protein from bone (2). The recent demonstration that a mouse knockout of matrix Gla protein died shortly after birth because of rupture of calcified arteries (16) provides further proof for the more general importance of γ -glutamyl carboxylase. Still another class of Gla proteins was identified when a gene selected during growth arrest of cells in culture was found to code for a protein containing a Gla domain. This protein, called Gas-6 (17), has subsequently been shown to serve as a ligand for the receptor tyrosine kinase Axl (3). It was also found that protein S, a Gla protein whose absence or defect results in severe thrombotic events (18), is a ligand for Tyro-3, which is closely related to Axl. Both proteins are expressed in neurons (3, 19) as well as a number of other tissues. Two additional Gla proteins, proline-rich Gla proteins 1 and 2 (PRPG-1 and PRPG-2), were identified by searching the human genome data base for sequences homologous to the Gla domain of coagulation proteins (20). Although the function of neither protein is known, both are expressed widely, with PRPG-1 having its highest levels in neural tissues and PRPG-2 having its highest levels in thyroid tissue.

Until now, we have been unable to identify a Gla domain in the *Drosophila* genomic sequence that is recognizably similar

to the Gla domains of most mammalian Gla proteins. A consensus sequence was made using Emotif (21) to scan the domain sequences of blood coagulation factor IX, FX, bone Gla protein, and matrix Gla proteins. This consensus sequence, EXX(E/R)EXCXXXXXXXX(L/F/Y)XXXXXXXX(A/F)(Y/W)XX(F/Y/H) (where *X* equals any amino acid and the letters within parentheses indicate the amino acids found at that position in the pattern), identified all of the known Gla proteins, including osteocalcin and matrix Gla protein. However, a search of the *Drosophila* data base using this pattern failed to identify any recognizable Gla domain. Failure to find a Gla domain similar to the canonical Gla domain of coagulation proteins is not, however, evidence against the presence of Gla proteins in *Drosophila*. It has been shown in mammals that attaching a propeptide of a vitamin K-dependent protein to random substrates results in the carboxylation of those substrates, both *in vivo* and *in vitro* (22, 23). This suggests that the structure of the Gla domain exists, not for recognition by the γ -glutamyl carboxylase, but because it is a characteristic necessary for the function of the particular vitamin K-dependent protein itself. Thus, it is likely that still other vertebrate Gla proteins, which bear no resemblance to the known Gla domains, remain to be identified. Because of its easily accessible genetics, *Drosophila* presents a unique opportunity to identify these new Gla proteins and then, via homology cloning, to identify the same proteins in mammals. The identification of these sequences is likely to lead to an understanding of additional roles for vitamin K-dependent proteins in mammalian cells.

In summary, we have identified and made preliminary characterizations of a functional γ -glutamyl carboxylase in *Drosophila*. This enzyme can utilize reduced vitamin K₁ as a substrate, making it very likely that the powerful genetics of *Drosophila* can be used to find the gene, still not known, for the vitamin K reductase. It will also aid in our understanding of evolution and in understanding the mechanism by which the carboxylase modifies its substrates.

REFERENCES

- Bush, K. A., Stenflo, J., Roth, D. A., Czerwiec, E., Harrist, A., Begley, G. S., Furie, B. C., and Furie, B. (1999) *Biochemistry* **38**, 14660–14666
- Price, P. A. (1989) *Connect Tissue Res.* **21**, 51–57
- Stitt, T. N., Conn, G., Gore, M., Lai, C., Bruno, J., Radziejewski, C., Mattsson, K., Fisher, J., Gies, D. R., Jones, P. F., *et al.* (1995) *Cell* **80**, 661–670
- Malicki, J., Cianetti, L. C., Peschle, C., and McGinnis, W. (1992) *Nature* **358**, 345–347
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., pp. 8.3–8.80, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Wu, S. M., Morris, D. P., and Stafford, D. W. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2236–2240
- Stanley, T., Stafford, D., Olivera, B., and Bandyopadhyay, P. (1997) *FEBS Lett.* **407**, 85–88
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) *Comput. Appl. Biosci.* **8**, 275–282
- Aguinaldo, A., Turbeville, J., Linford, L., Raff, R., and Lake, J. (1997) *Nature* **387**, 489–492
- Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R., Shetty, R., Watkins, M., Bandyopadhyay, P., and Hillyard, D. R. (1999) *Ann. N. Y. Acad. Sci.* **870**, 223–237
- Wu, S.-M., Mutucumarana, V. P., Geromanos, S., and Stafford, D. W. (1997) *J. Biol. Chem.* **272**, 11718–11722
- Roth, D. A., Whirl, M. L., Velazquez-Estades, L. J., Walsh, C. T., Furie, B., and Furie, B. C. (1995) *J. Biol. Chem.* **270**, 5305–5311
- Xu, J., Baase, W. A., Baldwin, E., and Matthews, B. W. (1998) *Protein Sci.* **7**, 158–177
- Clackson, T., Ultsch, M. H., Wells, J. A., and de Vos, A. M. (1998) *J. Mol. Biol.* **277**, 11111–11128
- Hall, J. G., Pauli, R. M., and Wilson, K. M. (1980) *Am. J. Med.* **68**, 122–140
- Luo, G., Ducy, P., McKee, M. D., Pinero, G. J., Loyer, E., Behringer, R. R., and Karsenty, G. (1997) *Nature* **386**, 78–81
- Manfioletti, G., Brancolini, C., Avanzi, G., and Schneider, C. (1993) *Mol. Cell. Biol.* **13**, 4976–4985
- Simmonds, R. E., Ireland, H., Lane, D. A., Zoller, B., Garcia de Frutos, P., and Dahlback, B. (1998) *Ann. Intern. Med.* **128**, 8–14
- Li, R., Chen, J., Hammonds, G., Phillips, H., Armanini, M., Wood, P., Bunge, R., Godowski, P. J., Sliwkowski, M. X., and Mather, J. P. (1996) *J. Neurosci.* **16**, 2012–2019
- Kulman, J. D., Harris, J. E., Haldeman, B. A., and Davie, E. W. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 9058–9062
- Nevill-Manning, C. G., Wu, T. D., and Brutlag, D. L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5865–5871
- Furie, B. C., Ratcliffe, J. V., Tward, J., Jorgensen, M. J., Blaszkowsky, L. S., DiMichele, D., and Furie, B. (1997) *J. Biol. Chem.* **272**, 28258–28262
- Stanley, T. B., Wu, S. M., Houben, R. J., Mutucumarana, V. P., and Stafford, D. W. (1998) *Biochemistry* **37**, 13262–13268