A *Drosophila* gene encoding a protein with similarity to diacylglycerol kinase is expressed in specific neurons

Nicholas HARDEN,* Seow Fong YAP,* Mary-Ann CHIAM* and Louis LIM*†

*Glaxo-IMCB Laboratory, Institute of Molecular and Cell Biology, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511, Republic of Singapore, and †Department of Neurochemistry, Institute of Neurology, 1 Wakefield St., London WC1N 1PJ, U.K.

A *Drosophila* gene encoding a protein with similarity to diacylglycerol kinase (DGK) was isolated by screening a genomic DNA library with a human DGK cDNA under low-stringency hybridization conditions. This putative *Drosophila* DGK gene (dDGK) maps to 43E on the right arm of chromosome 2. A dDGK cDNA with an open reading frame encoding a 517 amino acid protein was obtained in a screen of a 3–12-h embryonic cDNA library. In a 236-amino-acid overlap at their C-termini there is an identity of 52.5% between the dDGK protein and human DGK. The N-terminus of the *Drosophila* protein is not similar to human DGK, and contains

clusters of polar amino acids. dDGK is transcribed in the embryonic, pupal and adult stages, with little expression during the larval stages. Transcripts of 1.7–2.2 kb, 3.5 kb, 3.7 kb and 6.6 kb are seen, although most of the smaller transcripts may be from genes with similarity to dDGK. In stage-16 and stage-17 embryos, dDGK transcripts are limited to the central nervous system and head. There is a particularly high level of expression in the cell bodies of the larval photoreceptor organ, and in the cell bodies of the ventral unpaired median neurons. The dDGK protein may be involved in regulating signal transduction in these specific neurons.

INTRODUCTION

The cleavage of phosphatidylinositols by phospholipase C during receptor activation yields inositol trisphosphate (IP₃) and diacylglycerol (DG) (reviewed in Hokin, 1985). IP₃ leads to the release of calcium from intracellular storage sites (Berridge and Irvine, 1984) and DG activates protein kinase C (PKC), a key modulator of signal transduction (Nishizuka, 1984, 1986). DG causes PKC to become tightly associated with membranes, as do the tumour-promoting phorbol esters which are analogues of DG. Phorbol esters are not metabolized and induce a prolonged association of PKC with membranes. The tumorigenicity of phorbol esters is believed to be due to their action on PKC, but other phorbol ester-binding proteins have been characterized (Ahmed et al., 1990; Maruyama and Brenner, 1991). These proteins presumably also bind DG and may be activated by it. There is also evidence for the involvement of DG in *ras* cell transformation. A cytoplasmic protein that inhibits the GTPase activity of *ras* is activated by DG (Tsai et al., 1990), and transformation of 3T3 cells by *ras* produces an increase in DG levels and a decrease in the membrane-bound activity of the enzyme diacylglycerol kinase (DGK) (Huang et al., 1988; Kato et al., 1988).

Diacylglycerol kinase (DGK) converts DG to phosphatidic acid (PA), leading to the resynthesis of phosphatidylinositols (Berridge and Irvine, 1984; Hokin, 1985). Consequently, DGK may regulate PKC and other DG-binding proteins by controlling the intracellular concentration of DG. The production of PA by DGK may have other effects, as PA is a potential growth factor (Moolenaar et al., 1986; Yu et al., 1988). Until recently, no eukaryotic DGK genes had been cloned, and the involvement of the enzyme in signal transduction remains poorly understood. DGKs appear to be a diverse set of proteins, with immunologically and enzymologically distinct species ranging in size from 80 to 280 kDa (reviewed in Kanoh et al., 1990). Porcine and

human DGK cDNAs have been obtained using partial amino acid sequences derived from purified DGKs of 80 and 86 kDa respectively (Sakane et al., 1990; Schaap et al., 1990). The predicted proteins encoded by these two cDNAs are highly similar, showing 93% identity and differing in length by only one amino acid.

It would be informative to characterize DGK in a species more amenable to genetic and molecular manipulation. An obvious choice is *Drosophila*, a system in which three PKC genes have already been isolated (Rosenthal et al., 1987; Schaeffer et al., 1989). Mutants have been described for one of these PKC genes (Smith et al., 1991), and a *Drosophila* retinal degeneration mutant, *rdgA*, is defective in DGK activity in the eye (Inoue et al., 1989). These mutants could prove useful in the characterization of any DGKs identified in *Drosophila*.

We report here the cloning of a *Drosophila* gene encoding a protein with similarity to the mammalian DGKs, and show that its expression in late embryos is specific to the central nervous system and head.

MATERIALS AND METHODS

Standard procedures

Standard procedures were used in the preparation, blotting and hybridization of DNA and RNA, and in the screening of cDNA libraries (Sambrook et al., 1989).

Low-stringency library screening

The entire human DGK cDNA (Schaap et al., 1990) was random-primed using a Boehringer-Mannheim kit and used to screen a *Drosophila* EMBL3 genomic library (Clontech). Filters were prehybridized and hybridized in a buffer containing 35% formamide, 5 × SSC, 5 × Denhardt's, 50 mM sodium phosphate (pH 6.8) and 100 µg/ml denatured salmon sperm DNA. Filters

Abbreviations used: IP₃, inositol trisphosphate; DG, diacylglycerol; PKC, protein kinase C; DGK, diacylglycerol kinase; PA, phosphatidic acid; 1 × SSC, 0.15 M NaCl/0.015 M sodium citrate, pH 6.8; VUM, ventral unpaired median neuron; CRD, cysteine-rich domain.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under accession no. X67335.

```

                                M Q K V 4
AATTGGCAATTGGCTGTTGGCAATTGCTAAGCCAAGTTTTCCCTTCTTTTTTCGCCGCGAGGTGACGCTGCACTCGAACATGCAGAAGGTG 90

I E L S Q S V V V D K S L M E R F E E I Q R Q S K Q V A T S 34
ATTGAGCTGTGCGAAAGTGTGTGGTCGACAAATCGCTGATGGAACGCTTCGAGGAAATCCAACGGCAGAGCAAGCAGGTGCCACGTCC 180

M G T A A S S T S I M M A S K T E T E M E T M A T M E F G S 64
ATGGGCACGGCGGCTCAAGCACATCGATTATGATGGCCAGCAAGACGGAACGGAATGGAACGATGCCACGATGGAGTTCGGCAGC 270

S T T T T N R T T T T K S I S M S T F E T Q C L Q Q T L R T 94
AGCACGACCACCAACAGAACGACCAGCAAGAGCATTTTCGATGTCCAGCTTCGAGACGCACTGCCTGCAACAGACCCCTGCGGACA 360

A M S S S S S N T S S G S P C N G N Q D A E T E V D S H A A 124
GCGATGAGCAGCAGCAGCAGCAACACGAGCAGCGGAGTCCGTGCAACGGCAACCAGGATGCGGAAACGGAAGTAGATTCCCATGCCGCT 450

A A A D V R E K S V P R R S G E T E K Q S L E T L L L Q H K 154
GCCGCTGCCGATGTCCGGGAGAAGTCCGTGCCAGGAGGTGGGTGAAACGGAAGCAATCACTTGAGACGCTGCTGCTGCAGCAGCAAA 540

Q Q M Q Q Q Q Q Q Q Q Q Q Q Q G V T S L A V E E A A T A T 184
CAGCAGATGCAACACAGCAGCAACAGCAGCAGCAACAGCAGCAAGGAGTCACATCACTAGCTGTGAGGAAAGTCGAACAGCAACG 630

P V G S N Q S D N S S Q R N K Q N N I L K Q Q I T L S L D L 214
CCAGTCGGGAGTAATCAAAGCGATAATAGTAGCCAGCGCAATAAGCAAAACACATCCTTAAACAGCAAATTACATTGTCTATTGGACTTG 720

S D H E D E P K D D G* G G* A G D G* T K S N G N S I P A T P A 244
TCCGACCATGAGGACGAGCCCAAGGACGATGGCGTGGCGTGGCGATGTTACCAAGAGCAACGGCAACAGCATTCCGGCAACACCAGCC 810

T P I T P T T P N A A S S V L Q Q Q Q Q Q H L Q F E Q Q Q K* 274
ACACCAATCACACCCACCACTCCGAATGCCGATCAAGTGTGCTGCAGCAACACAGCAACAGCATCTCCAGTTTCGAGCAGCAACAGAG 900

P I K V Q S D K D C T V P Y N I I N N Y F S V G V D A A I C 304
CCCATCAAAGTGCAATCCGACAAGGACTGCACGGTGCCTACAACATAATCAACAATTATTTCTCCGTCGGCGTGAGCGCCGCTATGTC 990

V K F H L E R E K N P H K F N S R M K N K L W Y F E Y A T S 334
GTCAAGTTCACCTGGAGCGTGAGAAGAACCAGCAAAATTAACAGCCGATGAAGAACAAGCTGTGGTACTTCGAGTATGCAACATCC 1080

E T F A A S C K N L H E S I E I V C D G V A L D L A N G P H 364
GAGACATTTGGCGCATCTGCAAGAATCTCCACGAGAGCATCGAGATTGTGTGCGATGGCGTGGCCCTGGAATTTGGCAACGACCCAC 1170

L Q G V A L L N I P Y T H G G S N L W G E H L S Q K R I R K 394
TTGCAGGAGTGGCTCTGCTCAACATCCCGTACACACAGCGGGTTCACACCTGTGGGCGAGCACCTGTCTCAGAAACGGATCCGCAAG 1260

S A G P F G K S K K L R A G D K E F S A T S F N S V D L S V 424
AGTGCCGGCCCTTCGGCAAGAGTAAGAACTACGCGCCGGCGACAAGGAGTTCTCCGCCACAGCTTCAATTCGCTGGACCTCTCGGTG 1350

A I Q D F G D R L I E V I G L E N C L H M G Q V R T G L R A 454
GCCATTAGGACTTTGAGATCGGCTCATCGAGTAATTGGGCTGGAGAAGTCCCTGCACATGGGCCAGGTGAGGACTGGACTTCGCGCG 1440

S G R R L A Q C S E V I I K T K K T F P M Q I D G E P W M Q 484
TCGGGACGCCCTGGCCAGTCAGCGAAGTGATCATCAAGCAAGAAACATTTCCTATGCAGATAGATGGCGAGCCCTGGATGCAG 1530

M P C T I K V T H K N Q V P M L M A P R S E K G R G F F N L 514
ATGCCCTGCAGATCAAGGTGACCCACAAGAACCAGGTGCCATGCTGATGGCACCAGCGTGGGAAAGGAGCGCGGATTTTCAACCTG 1620

L C S * 517
CTGTGCGACTGATTAGGCGCAACGCATCGAGCGACTTCATGGGCGGAGTGCCAGGAGTCGGCCGATGAGCTGGATGATGACTTTCCG 1710
GTCGTCCTTAGAAGCACCACCGTCCACAACATAGCCTAGCGTCAATAGCTTTAGCCAGTCCACCCGCGATCGCATCGAATTGGTTAAC 1800
CTAGGTTAAGGCTTCCCTTATGATATGATATGATATGATATGATATGATATGATATGATAGTGTAGACTGCTATATGTGTATGATTTCTGTAT 1890
GCATGTATGTCAGTGAGTTACAGAGCTTTATATTAGTGTAGACTAGCCAAATATACAACCGAGTCAAGTCCATTCCAATGGCATCCG 1980
AGTGCGGTGGCTGTATATACATACATAGCAATACTGCGACGGCTTAGGCCACCGGATATTTAAGAACACACTTTAAAGCGGATAGCAT 2070
AGCGAAATTTGTCAAAGATGCAACTGAAATGAGACCAACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2160
GTGCAATTTTATGCCCTAGGCTAAGTACGAGTATGTAATTTATCCGTCAG 2212

```

Figure 1 DNA sequence and conceptual translation of the 2.2 kb dDGK cDNA

Double underlining indicates clusters of polar amino acid residues; single underlining the repeated amino acid sequence P-X-T. The amino acid residues marked with asterisks comprise a putative ATP-binding site. The position of the 5' end of the 3.4 kb cDNA, which is identical with the 2.2 kb cDNA in the region of overlap, is marked with an arrowhead at nucleotide 568.

were prehybridized for 4 h at 42 °C, and then hybridized for 16–18 h at 42 °C after the probe was added. Filters were washed once in 2 × SSC, 0.1 % SDS at room temperature for 20 min, and then once in 2 × SSC/0.1 % SDS at 55 °C for 30 min.

DNA sequence analysis

cDNA and genomic DNA fragments were subcloned into pUC vectors and sequenced on both strands by the dideoxy chain-

termination method using Sequenase (United States Biochemical). Sequences were completed using specific 17-mer oligonucleotides as primers.

In situ hybridization to polytene chromosomes

Polytene chromosomes prepared from wild-type (Canton S) third-instar larvae were hybridized with tritiated DNA probes, as described by Ashburner (1989) except that the RNAase and

acetylation steps were omitted, and LM-1 emulsion (Amersham) was used for autoradiography. Subsequent *in situ* hybridizations to deficiency chromosomes were performed using biotinylated DNA probes with detection by the Detek-I-horseradish peroxidase system (Enzo Biochemicals), again as described by Ashburner (1989).

Whole-mount *in situ* hybridization to embryos

Hybridization of digoxigenin-labelled DNA probes to wild-type (Canton S) embryos was performed as described by Tautz and Pfeifle (1989). Embryos were staged as described by Campos-Ortega and Hartenstein (1985).

RESULTS

A *Drosophila* genomic DNA clone hybridized by the human DGK cDNA was isolated and found to localize to band 43EF on *Drosophila* chromosome 2R (results not shown). The clone has been mapped relative to various deficiencies in the region and on this basis has been positioned at 43E6–7 (results not shown). There are no published mutants in this interval. A 1.9 kb *Bam*HI–*Hind*III restriction fragment which hybridized to the human cDNA was identified within the genomic clone. The fragment was subcloned and sequenced, revealing open reading frames with strong similarity to the C-terminus of human DGK, indicating it to be part of a putative *Drosophila* DGK gene (dDGK). The fragment was used to screen a *Drosophila* adult head λ gt11 cDNA library (Itoh et al., 1985), and a cDNA of 3.4 kb was obtained. DNA-sequence analysis of this cDNA revealed an open reading frame of 1.1 kb, with the rest of the cDNA consisting of 3' untranslated sequences, including a polyadenylation signal and poly(A) tail (results not shown). The cDNA did not appear to be full length, as the open reading frame extended fully to the 5' end and no suitable initiator methionine residue was seen in the conceptual translation. The open reading frame contained the exons seen in the genomic clone and exhibited further similarity to human DGK.

A 1.8 kb *Bam*HI–*Hind*III fragment was subcloned from the 3.4 kb cDNA and used to screen a 3–12-h embryonic λ gt10 cDNA library (Pool et al., 1985). A 2.2 kb cDNA was isolated and sequenced. This cDNA has a much shorter 3' untranslated region than the previous cDNA, and is lacking in a consensus polyadenylation signal and poly(A)⁺ tail (Figure 1). The two cDNAs are identical in sequence in the region of overlap. The 2.2 kb cDNA has a single large open reading frame extending from nucleotide 31 to 1629. The first AUG codon occurs at nucleotide 79, although there are numerous other AUG codons further 3'. None of these AUG codons has a good match to the sequence context expected for a translation initiation codon (Cavener, 1987) and the first AUG has been selected to encode the putative initiator methionine in Figure 1. The predicted protein is 517 amino acids long with a calculated molecular mass of 57 kDa. The C-terminal region of this protein shows strong similarity to the C-terminal region of the predicted human DGK. In a 236 amino acid overlap there is 52.5% identity between the proteins (Figure 2). N-Terminal to the region of strong identity between the dDGK protein and human DGK the similarity between the two proteins breaks down completely. The N-terminus of the dDGK protein is enriched for polar amino acid residues. There are clusters of threonine residues (8 of 10 residues; amino acids 66–75), serine residues (8 of 11 residues; amino acids 97–107) and glutamine residues (16 of 19 residues; amino acids

152–170, and 10 of 14 residues; amino acids 260–273). Aside from proteins with clusters of polar amino acids, the *Drosophila* protein is not significantly similar to proteins other than the mammalian DGKs in searches of current databases. An *Escherichia coli* DGK (Loomis et al., 1985) with no significant

dgdk	282	R	D	C	T	V	F	Y	N	I	N	N	Y	F	S	V	G	V	D	A	A	I	C	V	K	F	H	L	E	R	
hdgk	512	K	S	D	P	V	F	Q	I	N	N	Y	F	S	I	G	V	D	A	S	I	A	H	R	F	H	I	M	R		
dgdk	312	E	K	N	P	H	K	F	N	S	R	M	K	N	K	L	W	Y	F	E	Y	A	T	S	E	T	F	A	A	S	C
hdgk	542	E	K	Y	P	E	K	F	N	S	R	M	K	N	K	L	W	Y	F	E	F	A	T	S	E	S	I	F	S	T	C
dgdk	342	R	N	L	H	E	S	I	E	I	-	V	C	D	G	V	A	L	D	L	A	N	G	P	H	I	Q	G	V	A	L
hdgk	572	K	K	L	E	S	L	T	V	E	I	C	-	G	K	P	L	D	L	S	N	L	S	-	L	E	G	I	A	V	
dgdk	371	L	N	I	P	Y	T	H	G	G	S	N	L	W	G	E	H	L	S	Q	K	R	I	R	K	S	A	G	P	F	G
hdgk	600	L	N	I	P	S	M	H	G	G	S	N	L	W	G	D	-	-	-	T	R	-	R	P	H	G	D	I	Y	G	
dgdk	401	K	S	K	K	L	R	A	G	D	K	E	F	S	A	T	S	F	N	S	V	D	L	S	V	A	I	Q	D	F	G
hdgk	625	I	N	Q	A	L	G	A	T	A	K	V	I	T	D	P	D	-	-	-	I	L	K	T	C	V	P	D	L	S	
dgdk	431	D	R	L	I	E	V	I	G	L	E	N	C	L	H	M	G	Q	V	R	T	G	L	R	A	S	G	R	R	L	A
hdgk	651	D	K	R	L	E	V	I	G	L	E	G	A	I	E	M	G	Q	I	Y	T	K	L	K	N	A	G	R	R	L	A
dgdk	461	Q	C	S	E	V	I	I	K	T	K	K	T	F	P	M	Q	I	D	G	E	P	W	M	Q	H	P	C	T	I	K
hdgk	681	K	C	S	E	I	T	F	H	T	K	T	L	P	M	Q	I	D	V	E	P	W	M	Q	T	P	C	T	I	K	
dgdk	491	V	T	H	K	N	Q	V	P	M	L	M	A	P	R	S	E	K	G	R	G	F	F	N	L	L	C	S	*		
hdgk	711	I	T	H	K	N	Q	M	P	M	L	M	G	P	P	P	-	R	S	T	N	F	F	G	F	L	-	S	*		

Figure 2 Comparison of the C-terminal amino acid sequence of the predicted dDGK protein with the C-terminal amino acid sequence of human DGK

Identical amino acids are boxed and gaps have been introduced to maximize identity. The asterisk indicates the position of the stop codon in the cDNA sequence.

Figure 3 Developmental profile of dDGK transcription

The numbers and lines on the left correspond to the size (kb) and position of RNA standards (Bethesda Research Laboratories). (a) The Northern blot was probed with a ³²P-labelled 1.8 kb *Bam*HI–*Hind*III fragment from the 3.4 kb dDGK cDNA (this fragment does not contain the polyglutamine-encoding sequences). Each lane contains 2 μ g of poly(A)⁺ RNA prepared from embryonic stages (0–6 h, 6–12 h and 12–24 h after egg laying), larval stages (L1, first instar; L2, second instar; L3, third instar), pupae (P), adult head (H) and adult body (B). (b) The blot was reprobed with actin to demonstrate approximately equal loading of RNA in each lane.

Figure 4 Expression of dDGK transcripts in late embryos

Stages are according to Campos-Ortega and Hartenstein (1985). (a) Lateral view of a stage-15 embryo; anterior is to the left and dorsal is at the top. Expression is widespread, but strongest in the central nervous system and head. (b) Lateral view of stage-17 embryo; orientation is as before. A particularly high level of expression is seen in the cell bodies of Bolwig's nerve (small arrow) and the VUM cells (large arrow). (c) Dorsal view of stage-17 embryo (anterior is to the left) showing heavy expression in the cell bodies of Bolwig's nerve (arrow). (d) Ventral view of a dissected stage-17 ventral nerve cord showing expression in the VUM cells. The arrow indicates the midline and anterior is at the top. (e) Higher-magnification view of transcript in the cell bodies of Bolwig's nerve; orientation is as in (c). The arrow indicates the pharyngeal musculature.

similarity to the mammalian DGKs is not similar to the dDGK protein.

Northern blots bearing poly(A)⁺ RNA from various developmental stages of *Drosophila* (Figure 3) were hybridized with dDGK probes. Smearly multiple transcripts are seen in the size range of about 1.7–2.2 kb. From additional blots it appears that there are at least three transcripts in this size range (results not shown). Faint bands are seen at 3.5 and 3.7 kb in addition to a transcript of 6.6 kb. The 6.6 kb transcript is expressed in the adult head but not in the body. Expression of all transcripts is weak or non-existent in the larval stages. Despite extensive screening of cDNA libraries from embryonic, larval and pupal stages, no transcripts of 6.6 kb have been isolated.

The spatial distribution of dDGK mRNA during development was examined by RNA *in situ* hybridization experiments using whole-mount embryos from various stages. First observed around the cellular blastoderm stage (stage 5), dDGK transcripts are fairly evenly distributed throughout the embryo during the early stages of embryogenesis (results not shown). Beginning around late stage 14/early stage 15 (Figure 4a), and proceeding to the end of embryogenesis, expression is particularly strong in the developing head and central nervous system. From late stage 16 onwards specific clusters of cells in the ventral nerve cord and head exhibit increased levels of dDGK transcription (Figures 4b–4e).

Within each abdominal segment of late (stage 16/stage 17) embryos there is a cluster of at least six cells showing high levels of dDGK transcript (Figures 4b and 4d). These clusters of cells are located ventrally on the midline in the ventral nerve cord, and we believe that they are the cell bodies of the ventral unpaired median neurons (VUMs). The VUMs exist as groups of six neurons and two glial support cells, and earlier in development they are involved in the formation of the commissures (Klämbt et al., 1991).

The bilateral clusters of the cells expressing dDGK mRNA in the head appear to be the cell bodies of Bolwig's nerve, the larval visual organ (Bolwig, 1946; Steller et al., 1987). The 12 cell bodies of the Bolwig's photoreceptor neurons are clustered tightly near the pharyngeal muscles. We can detect transcript in at least eight of these cell bodies, but not in their axons.

Work on mammalian DGKs suggests that they comprise a diverse set of proteins. Thus it is quite likely that there will be more than one DGK in *Drosophila*. During our cDNA screens it became apparent that there are transcripts other than those from dDGK which are similar to dDGK probes. Numerous cDNAs of 2.2 kb or less in size have been isolated which weakly hybridize with dDGK, and their corresponding mRNAs may account for much of the signal seen on Northern blots in the 1.7–2.2 kb size range. The strong signal in this size range is probably a reflection of high levels of expression. We do not yet know if these transcripts encode proteins with similarity to the dDGK protein. The *in situ* hybridizations to embryos have been repeated with probes containing various pieces of the dDGK cDNA to check that the expression pattern is due to dDGK transcripts and not related sequences. In all cases, a pattern of cell-specific labelling identical with that seen with whole dDGK probes is observed (results not shown). Therefore we are confident that the cell-specific transcripts are derived from dDGK. It may be that some of the weaker signals seen throughout the head and central nervous system are due to cross-homology to other transcripts.

DISCUSSION

We have isolated and characterized a *Drosophila* gene (dDGK) with strong similarity to the mammalian DGKs. A cDNA of 2.2 kb in length has been sequenced and shown to encode an open reading frame of 517 amino acids. The last 236 amino acids of this predicted protein show 52.5% identity with the C-

terminal region of human DGK. The identity becomes 64% if the following groups of conservative changes are included: MILV; FYW; ST; AG; DE; NQ; RK. Although the *Drosophila* protein has not been assayed for DGK activity, this degree of similarity with a biochemically characterized DGK suggests a similar if not identical function, and it is likely that the catalytic domain of the DGKs is located at the C-terminus. Regulatory domains and possibly substrate-binding sites would then lie in the N-terminal region, the structure of which could differ between DGKs depending on their particular DG substrate. It is likely that DGKs vary in their substrate specificities, as the acyl chain composition of the DG substrate can affect the level of DGK activity (MacDonald et al., 1988). The N-terminus of the dDGK protein is very different from that of human and porcine DGK, and it will be of interest to see if proteins with similarity to this domain can be found in other species.

There are a number of distinctive features in the dDGK protein N-terminal region, including clusters of the polar amino acids serine, threonine and glutamine. Stretches of polyglutamine occur in numerous *Drosophila* genes (Wharton et al., 1985), many of them transcription factors. Glutamine-rich and serine/threonine-rich sequences are seen in a number of transcriptional activation domains, and may be involved in protein-protein interactions (Mitchell and Tjian, 1989; Pugh and Tjian, 1990; Ruvkun and Finney, 1991). In addition to being rich in polar amino acids, the N-terminus has the motif P-X-T repeated four times from amino acid residue 240 to 251 (Figure 1). There are at least four other proteins with similar repeats. The herpes simplex virus-infected cell protein ICP34.5 (Chou and Roizman, 1986) has ten copies of the motif P-A-T and a *Drosophila* hobo transposable element protein has six copies of the repeat P-E-T (Streck et al., 1986). The COUP transcription factor (Wang et al., 1989) has six copies of the motif P-X-T, although two of the threonine residues are substituted by other amino acids. *Drosophila* TFIID has three copies of P-X-T, close to a stretch of polyglutamine (Hoey et al., 1990). The significance of the P-X-T motif in these proteins is unknown, but COUP and TFIID are DNA-binding proteins, and it is likely that the hobo protein is as well. The function of ICP34.5 is unknown.

It is interesting to note that the N-terminus of the dDGK protein has sequences similar to those seen in various DNA-binding proteins, as the N-termini of the mammalian DGKs also bear sequences reminiscent of structures found in transcription factors. The mammalian DGKs (Sakane et al., 1990; Schaap et al., 1990) have two copies of a cysteine-rich domain (CRD) very similar to that seen in the phorbol ester-binding proteins, PKC (Parker et al., 1986), n-chimaerin (Hall et al., 1990) and the product of the *unc-13* gene (Maruyama and Brenner, 1991). This CRD is necessary for high-affinity phorbol ester binding by PKC (Ono et al., 1989; Burns and Bell, 1991) and n-chimaerin (Ahmed et al., 1990). As phorbol ester is an analogue of DG, the CRD is almost certainly the site of DG binding in these proteins. It is tempting to speculate that the CRDs of the mammalian DGKs also bind DG even though they are not high-affinity phorbol ester receptors (Ahmed et al., 1991). As the DGKs would be binding DG as a substrate rather than a ligand, it is perhaps not surprising that their CRDs do not bind phorbol ester with high affinity. The CRD has a similar distribution of cysteines to that of the DNA-binding zinc fingers of the steroid hormone receptors (Hall et al., 1990; Ahmed et al., 1991), and it binds phorbol ester in a zinc-dependent fashion (Ahmed et al., 1991). It is intriguing that such apparently similar structures as zinc fingers and CRDs can be used to bind molecules as different as DNA and phorbol ester. The dDGK protein does not have CRDs and may bind DG using a different motif, perhaps one of those described

above. If the dDGK protein binds phorbol ester, it should be possible to identify DG-binding sequences.

Another structure seen in the mammalian DGKs but not in the dDGK product is the putative Ca^{2+} -binding E-F hand motif. Both porcine and human DGK have two such motifs. Although porcine DGK appears to be regulated by Ca^{2+} (Sakane et al., 1990), no Ca^{2+} -dependence has been observed for human DGK (Schaap et al., 1990). The mammalian DGKs have two amino acid stretches that match the protein kinase ATP-binding site consensus sequence Gly-X-Gly-X-X-Gly-X₁₅₋₂₀-Lys (Hunter, 1987). Although the dDGK protein does not contain a perfect match to this consensus, it does have a glycine-rich stretch with a neighbouring lysine [amino acid residues 225–274 (Figure 1)]. It is not known if the putative ATP-binding sites in the mammalian DGKs actually bind ATP.

We have shown with RNA blots that dDGK transcripts are predominantly expressed in the embryonic, pupal and adult stages. dDGK probes detect transcripts of 1.7–2.2 kb, 3.5 kb, 3.7 kb and 6.6 kb. We believe that many of the small transcripts detected on Northern blots are from genes with similarity to dDGK. The 2.2 kb dDGK cDNA has no consensus polyadenylation signal or poly(A) tail and may be derived from a transcript larger than 2.2 kb with a longer 3' end. Another cDNA, identical with the 2.2 kb cDNA in the region of overlap, but truncated at the 5' end, has a long 3' untranslated region of about 2.3 kb, complete with polyadenylation signal and poly(A) tail. If this 3' tail occurs in conjunction with the 517 amino acid open reading frame, it would result in a transcript of about 3.9 kb in size. No cDNAs have yet been isolated which would correspond to the 6.6 kb transcript.

In late (stage 16/ stage 17) embryos, dDGK probes detect particularly high levels of transcript in specific cells in the central nervous system and head. It is possible that the dDGK protein regulates a developmental signal-transduction pathway involved in determining the fate of these cells. If dDGK transcription is maintained in these neurons in the larval stages, this would suggest a role for the dDGK protein in their activity. The lack of transcript in the larval stages on blot hybridization does not necessarily mean that the cell-specific pattern of expression is not maintained, since these neurons constitute such a small number of cells, and their transcripts may be undetectable on Northern blots. It is possible that the dDGK protein is involved in both the differentiation and functioning of specific cells.

dDGK mRNA is highly expressed in the cell bodies of the VUMs of the ventral nerve cord. The VUMs are motor neurons which establish projections to larval muscles during late embryogenesis (Sink and Whittington, 1991). Interestingly, a gene encoding a *Drosophila* 5-hydroxytryptamine receptor (5HT-dro2A) is expressed in the VUMs, beginning at stage 16 (Saudou et al., 1992). The 5HT-dro2A receptor activates phospholipase C and may be involved in motor control. It is conceivable that the dDGK protein could act downstream of this 5-hydroxytryptamine receptor by regulating the levels of DG produced by phospholipase C.

The expression of dDGK mRNA in the cell bodies of the larval photoreceptor organs suggests that the dDGK protein may be involved in regulating phototransduction. It will be of great interest to ascertain if the dDGK protein is present in the adult photoreceptors. The *Drosophila* retinal degeneration mutant, *rdgA*, is defective in DGK activity in the adult eye (Inoue et al., 1989). We know from the cytological location of dDGK that it is not the *rdgA* gene. The *rdgA* gene may encode a DGK with little or no similarity to the dDGK protein, or it may be involved in the expression of DGK in the adult photoreceptors. If dDGK is expressed in the adult photoreceptors, or if it can be

used to isolate the *rdgA* gene, it may prove very useful in unravelling DGK function. There is a photoreceptor-specific PKC in *Drosophila* (Schaeffer et al., 1989), and mutants have been identified in the structural gene (Smith et al., 1991). These mutants may provide both genetic and molecular routes to understanding how DGK participates in signal transduction.

Note added in proof (received 30 October 1992)

A transcript from the gene as dDGK, encoding a larger protein of 791 amino acids, has recently been described (Masai et al., 1992). The transcript of Masai et al. is expressed in the same embryonic pattern as ours, and their larger predicted protein exhibits additional similarity to mammalian DGK.

We thank Bill Chia and John Roote for fly stains, Bernadette Oei-Murugasu for the developmental Northern blot used in Figure 3, Ben Li for oligonucleotides, Dick Schaap for the human DGK cDNA, members of the Lim and Chia labs for advice and discussions, and the Glaxo-Singapore Research fund for financial support. We thank Sohail Ahmed for his encouragement during the course of this work.

REFERENCES

- Ahmed, S., Korma, R., Monfries, C., Hall, C., Lim, H. H., Smith, P. and Lim, L. (1990) *Biochem. J.* **272**, 767–773
- Ahmed, S., Korma, R., Lee, J., Monfries, C., Harden, N. and Lim, L. (1991) *Biochem. J.* **280**, 233–241
- Ashburner, M. (1989) *Drosophila: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Berridge, M. J. and Irvine, R. F. (1984) *Nature (London)* **313**, 315–321
- Bolwig, N. (1946) *Vidensk. Medd. fra Dansk. Naturh. Foren.* **109**, 80–212
- Burns, D. J. and Bell, R. M. (1991) *J. Biol. Chem.* **266**, 18330–18338
- Campos-Ortega, J. A. and Hartenstein, V. (1985) *The Embryonic Development of Drosophila melanogaster*, Springer-Verlag, Berlin
- Cavener, D. R. (1987) *Nucleic Acids Res.* **15**, 1353–1361
- Chou, J. and Roizman, B. (1986) *J. Virol.* **57**, 629–637
- Hall, C., Monfries, C., Smith, P., Lim, H. H., Kozma, R., Ahmed, S., Vannaisingham, V., Leung, T. and Lim, L. (1990) *J. Mol. Biol.* **211**, 11–16
- Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. and Tjian, R. (1990) *Cell* **61**, 1179–1186
- Hokin, L. E. (1985) *Annu. Rev. Biochem.* **54**, 205–235
- Huang, M., Chida, K., Kamata, N., Nose, K., Kato, M., Homma, Y., Takenawa, T. and Kuroki, T. (1988) *J. Biol. Chem.* **263**, 17975–17980
- Hunter, T. (1987) *Cell* **50**, 823–829
- Inoue, H., Yoshioka, T. and Hotta, Y. (1989) *J. Biol. Chem.* **264**, 5996–6000
- Itoh, N., Salvaterra, P. and Itakura, K. (1985) *Dros. Inf. Serv.* **61**, 89
- Kanoh, H., Yamada, K. and Sakane, F. (1990) *Trends Biochem. Sci.* **15**, 47–50
- Kato, H., Kawai, S. and Takenawa, T. (1988) *Biochem. Biophys. Res. Commun.* **154**, 959–966
- Klämbt, C., Jacobs, R. J. and Goodman, C. S. (1991) *Cell* **64**, 801–815
- Loomis, C. R., Walsh, J. P. and Bell, R. M. (1985) *J. Biol. Chem.* **260**, 4091–4097
- MacDonald, M. L., Mack, K. F., Williams, B. W., King, W. C. and Glomset, J. A. (1988) *J. Biol. Chem.* **263**, 1584–1592
- Maruyama, I. N. and Brenner, S. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5729–5733
- Masai, I., Hosoya, T., Kojima, S.-I. and Hotta, Y. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6030–6034
- Mitchell, P. J. and Tjian, R. (1989) *Science* **245**, 371–378
- Moolenaar, W. H., Kruyer, W., Tilly, B. C., Verlaan, I., Bierman, A. J. and de Laat, S. W. (1986) *Nature (London)* **323**, 171–173
- Nishizuka, Y. (1984) *Nature (London)* **308**, 693–698
- Nishizuka, Y. (1986) *Science* **233**, 305–313
- Ono, Y., Fujii, T., Igarashi, K., Kuno, T., Tanaka, C., Kikkawa, U. and Nishizuka, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4868–4871
- Parker, P. J., Coussens, L., Totty, N., Rhee, L., Young, S., Chen, E., Stabel, S., Waterfield, M. D. and Ullrich, A. (1986) *Science* **233**, 853–859
- Pool, S. J., Kauvar, L. M., Drees, B. and Kornberg, T. (1985) *Cell* **40**, 37–43
- Pugh, B. F. and Tjian, R. (1990) *Cell* **61**, 1187–1197
- Rosenthal, A., Rhee, L., Yadegari, R., Paro, R., Ullrich, A. and Goeddel, D. V. (1987) *EMBO J.* **6**, 433–441
- Ruvkun, G., and Finney, M. (1991) *Cell* **64**, 475–478
- Sakane, F., Yamada, K., Kanoh, H., Yokoyama, C. and Tanabe, T. (1990) *Nature (London)* **344**, 345–348
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Saudou, F., Boschert, U., Amlaiky, N., Plassat, J.-L. and Hen, R. (1992) *EMBO J.* **11**, 7–17
- Schaap, D., de Widt, J., Van der Wal, J., Vandekerckhove, J., van Damme, J., Gussow, D., Ploegh, H. L., van Blitterswijk, W. J. and van der Bend, R. L. (1990) *FEBS Lett.* **275**, 151–158
- Schaeffer, E., Smith, D., Mardon, G., Quinn, W. and Zuker, C. (1989) *Cell* **57**, 403–412
- Sink, H. and Whittington, P. M. (1991) *J. Neurobiol.* **22**, 298–311
- Smith, D. P., Ranganathan, R., Hardy, R. W., Marx, J., Tsuchida, T. and Zuker, C. S. (1991) *Science* **254**, 1478–1484
- Steller, H., Fischbach, K. F. and Rubin, G. M. (1987) *Cell* **50**, 1139–1153
- Strecker, R. D., MacGaffey, J. E. and Beckendorf, S. K. (1986) *EMBO J.* **5**, 3615–3623
- Tautz, D. and Pfeifle, C. (1989) *Chromosoma* **98**, 81–85
- Tsai, M.-H., Yu, C.-L. and Stacey, D. W. (1990) *Science* **250**, 982–985
- Wang, L. H., Tsai, S. Y., Cook, R. G., Beattie, W. G., Tsai, M. J. and O Malley, B. W. (1989) *Nature (London)* **340**, 163–166
- Wharton, K. A., Yedvobnick, B., Finnerty, V. G. and Artavanis-Tsakonas, S. (1985) *Cell* **40**, 55–62
- Yu, C.-L., Tsai, M.-H. and Stacey, D. W. (1988) *Cell* **52**, 63–71