Available online at www.sciencedirect.com
SCIENCE DIRECT®
Developmental Biology 266 (2004) 374–387
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Transcriptional control of glial and blood cell development in *Drosophila*: *cis*-regulatory elements of *glial cells missing*

Bradley W. Jones,* Matthew Abeysekera, Jolanta Galinska, and Ethel M. Jolicoeur

Molecular Neurobiology Program, Skirball Institute of Biomolecular Medicine, and Department of Pharmacology, New York University School of Medicine, New York, NY 10016, USA

Received for publication 9 September 2003, revised 10 October 2003, accepted 10 October 2003

Abstract

In *Drosophila*, glial cell differentiation requires the expression of *glial cells missing* (*gcm*) in multiple neural cell lineages, where *gcm* acts as a binary switch for glial vs. neuronal fate. Thus, the primary event controlling gliogenesis in neural progenitors is the transcription of *gcm*. In addition, *gcm* is also required for the differentiation of macrophages, and is expressed in the hemocyte lineage. This dual role of *gcm* in glial cell and blood cell development underscores the need for the precise temporal and spatial regulation of *gcm* transcription. To understand how *gcm* transcription is regulated, we have undertaken an analysis of the *cis*-regulatory DNA elements of *gcm* using *lacZ* reporter activity in transgenic embryos, testing the activity of approximately 35 kilobases of DNA from the *gcm* locus. We have identified several distinct DNA regions that promote most of the elements of *gcm* expression. These include elements for general neural expression, *gcm*-independent and *gcm*-dependent glial-specific expression, as well as early and late hemocyte expression. We show that expression of an abdominal glial-specific element is dependent on the homeotic gene *abdominal-A*. Our results indicate that *gcm* transcription is controlled by a combination of general and lineage-specific elements, positive autoregulation, and neuronal repression.

Keywords: *gcm*; Glia; Hemocyte; Macrophage; *Drosophila*; CNS; *cis*-regulatory elements; Homeotic genes; *abd-A*

Introduction

A functional nervous system requires the correct specification and precise organization of a large number of neural cell types. A central problem of developmental neurobiology is to understand mechanisms by which neural progenitor cells give rise to this diversity of cell types. Neural cells can be classified into two distinct groups, neurons and glia. One general rule that has emerged from lineage analysis of neurogenesis is that both neurons and glia are generated from multipotent neural progenitors or neural stem cells (e.g., Bossing et al., 1996; Davis and Temple, 1994; Frank and Sanes, 1991; Schmid et al., 1999; Schmidt et al., 1997; Stemple and Anderson, 1992; Turner and Cepko, 1987; Udolph et al., 1993; Williams and Price, 1995). While much effort has been made to identify such progenitors and mechanisms controlling their fates, the mechanisms that control whether neural progenitor cells will adopt glial vs. neuronal fates are poorly understood.

In vertebrates, neural stem cells respond to multiple cell-intrinsic, cell-extrinsic and temporal cues that induce glial cell fates, yet no single developmental pathway leading to gliogenesis has been discovered (Anderson, 2001). In *Drosophila*, however, a single gene, *glial cells missing* (*gcm*, also known as glide), is the primary regulator of glial cell determination. *gcm* encodes a novel transcription factor that is transiently expressed in all embryonic glia except for the midline/mesectoderm-derived glia (Akiyama et al., 1996; Hosoya et al., 1995; Jones et al., 1995; Schreiber et al., 1997; Vincent et al., 1996). *gcm* loss-of-function mutant embryos lack nearly all lateral glial cells, and presumptive glial cells are transformed into neurons (Hosoya et al., 1995; Jones et al., 1995; Vincent et al., 1996). Conversely, when *gcm* is ectopically expressed, presumptive neurons are transformed into glia (Hosoya et al., 1995; Jones et al., 1995). Thus, within the nervous system, *gcm* acts as a binary genetic switch, with Gcm-positive cells becoming glia and Gcm-negative cells becoming neurons.

* Corresponding author. Molecular Neurobiology Program, Skirball Institute of Biomolecular Medicine, and Department of Pharmacology, New York University School of Medicine, 5th Floor, 540 First Avenue, New York, NY 10016. Fax: +1-212-263-8214.
E-mail address: bjones@saturn.med.nyu.edu (B.W. Jones).

0012-1606/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2003.10.013

Currently, *gcm* occupies the most upstream position in lateral glial cell determination. In addition, the expression of *gcm* mRNA and protein closely profiles the initiation of gliogenesis in the various neural lineages that generate glia. These observations lead to the conclusion that the primary event controlling gliogenesis is the transcriptional regulation of *gcm*. It also follows that *gcm* is likely to be regulated by a combination of factors whose expression is not restricted to glia.

Glia are derived from neural progenitors that originate in the ventral neurogenic ectoderm and the peripheral ectoderm lateral to either side of the midline (Bossing et al., 1996; Goodman and Doe, 1993; Jan and Jan, 1993; Schmid et al., 1999; Schmidt et al., 1997). In the PNS, neural stem cells called sensory organ precursors (SOPs) undergo a series of cell divisions that generate specific types of neurons, glia, and other support cells. In the CNS, two types of progenitors generate glia: glioblasts (GBs) that give rise to only glial cells, and neuroglioblasts (NGBs) that produce mixed glial/neuronal lineages. All of the neural lineages in the PNS and CNS give rise to unique and stereotypic patterns of neuronal and glial progeny.

How each neural stem cell generates stereotypic patterns of multiple cell types in their lineage is not well understood. What is known is that neural progenitors express different combinations of transcription factors depending on their dorsal–ventral and anterior–posterior position within each segment, the timing of their delamination from the neuroectoderm and as they generate early-, mid- and late-born daughter cells (Brody and Odenwald, 2000; Doe et al., 1998; Goodman and Doe, 1993; Isshiki et al., 2001; Jan and Jan, 1993; Kambadur et al., 1998; Novotny et al., 2002). In addition, as neural progenitors divide, they inherit localized determinants such as Prospero and Numb, and are subject to signaling from the Notch pathway, which differentiates sibling cell fates (Doe et al., 1991; Hirata et al., 1995; Knoblich et al., 1995; Rhyu et al., 1994; Udolph et al., 2001; Uemura et al., 1989; Umesono et al., 2002; Vaessin et al., 1991). How all these factors come together to produce different patterns of glial vs. neuronal progeny is poorly understood. In the case of glial cells, the resulting readout of combinations of temporal and spatial cues somehow leads to the activation of *gcm* and the initiation of glial cell differentiation in precise and stereotypic patterns in the nervous system.

In addition to acting as a regulator of glial cell differentiation, *gcm*, together with its closely related homolog *gcm2* (also known as *glide2*), is required for the proper differentiation of the plasmatocyte/macrophage lineage of blood cells, or hemocytes, that are derived from the procephalic mesoderm (Alfonso and Jones, 2002; Bernardoni et al., 1997; Kammerer and Giangrande, 2001; Lebestky et al., 2000). Before *gcm* is activated in glial progenitors, *gcm* is expressed in the hemocyte primordium and in hemocyte precursors where it acts to trigger the maturation of hemocytes into macrophages. The ability of *gcm* to promote either glial cell differentiation or blood cell differentiation in different developmental contexts suggests that it cooperates with other tissues specific factors, which underscores the need for the precise temporal and spatial regulation of *gcm* transcription in the mesoderm and neuroectoderm.

The transcriptional regulation of *gcm* provides a model for understanding how diverse genetic inputs are integrated in different neural lineages to give rise to stereotypic patterns of glial vs. neuronal progeny, as well as the differentiation of the hematopoietic lineages. To identify genetic inputs controlling the transcriptional regulation of *gcm*, we have undertaken a systematic analysis of the *cis*-regulatory DNA elements of *gcm* using *lacZ* reporter activity in transgenic embryos. Recent studies aimed at rescuing *gcm* mutant phenotypes with *gcm* locus genomic clones demonstrate that as much as 9 kilobases (kb) of upstream flanking sequences are required to partially rescue glial differentiation in most CNS glia cell lineages in the embryo (Ragone et al., 2003); however, they also suggest that additional *cis*-regulatory DNA elements are required for complete rescue. Here we describe experiments testing the reporter activity of approximately 35 kilobases of DNA flanking the *gcm* transcription unit. We follow by further dissecting the DNA 9 kb upstream and 3 kb downstream of *gcm*, where we have identified several distinct regions that promote most of the elements of the *gcm* expression pattern; these can be further subdivided into general neural expression elements, *gcm*-independent and *gcm*-dependent glial-specific elements, as well as early and late hemocyte elements. *gcm*-independent activity in most CNS glial lineages lies within the first 5 kb upstream of the *gcm* promoter, as reporter expression is still initiated in *gcm* mutant embryos. *gcm*-dependent glial-specific activity lies further distal to the *gcm* promoter, where the presence of Gcm-binding sites implicates an autoregulatory mechanism. We have also identified a 3′ element that drives expression in a small subset of abdominal and maxillary segment glia that are derived from a single neural progenitor in each hemisegment. We show that the expression of this element in the abdominal region is dependent on the homeotic selector gene *abdominal-A*, indicating that anterior–posterior segmental differences in glial progenitor lineages are under the control of homeotic genes of the Hox complex. Our results indicate that *gcm* transcriptional regulation is controlled by a combination of both general and lineage-specific elements, positive autoregulation, and active repression in neurons.

Materials and methods

Construction of gcm-lacZ reporter genes

Two different transformation vectors were used to construct *gcm-lacZ* reporter genes: pCasPeR-hs43-lacZ

(abbreviated *hs43-lacZ*) (Thummel and Pirrotta, 1992) and pCasPeR-gcmp-lacZ (abbreviated *gcmp-lacZ*). The *hs43-lacZ* vector contains the *lacZ* reporter gene transcribed from a minimal promoter consisting of the *hsp70* TATA box, deleted at −43, and leader sequences, and has been used extensively for testing enhancers and tissue-specific regulatory elements. The *gcmp-lacZ* vector is a derivative of pCasPeR-β-gal (Thummel et al., 1988) that contains 424 base pairs (bp) from the *gcm* promoter region, including the first 17 amino acid codons of *gcm* fused in frame with the *lacZ* gene. PCR was used to generate the *gcm* promoter fragment, which includes the natural *Eco*RI site located 368 bp upstream of the ATG, as well as unique restriction sites (*Pst*I, *Xba*I, *Not*I, *Sac*II, and *Bam*HI). A *Sma*I site was engineered at amino acid codon 17 and the PCR fragment, cut with *Xba*I and *Sma*I, was cloned into pCasPeR-β-gal at the *Xba*I and blunted *Bam*HI sites. As a reference point, we refer to +1 as the position of the start of the longest published *gcm* cDNA (Hosoya et al., 1995, GenBank Accession No. , D64040); thus, as well as the endogenous ATG, *gcmp-lacZ* includes the 150-bp 5′ UTR and 223 bp of flanking sequence.

The genomic DNA clones used in this study to generate reporter constructs (Fig. 1) came from previously described genomic walks of the region spanning 30B–30C (λcc2, λg14 are from Lane and Kalderon, 1993; cos3.12.16 is from Jones et al., 1995). The endpoints of the genomic clones were sequenced to map them precisely onto the published *Drosophila melanogaster* genomic sequence (Adams et al., 2000). Fragments of *gcm* region DNA used in *gcm-lacZ* reporter genes are shown in Fig. 1. Genomic fragments were cloned into either *hs43-lacZ* or *gcmp-lacZ* with the normal 5′ to 3′ orientation with respect to the *gcm* promoter preserved. All constructs were introduced into flies by P-element-mediated transformation.

D. melanogaster stocks and genetics

w^{1118} and $y^{1}w^{67c23}$ were used for generating transgenic lines. The *gcm* null allele was $gcm^{\delta P1}$ (Jones et al., 1995); the *abd-A* null allele was $abd\text{-}A^{MX1}$ (Sanchez-Herrero et al., 1985). All reporter lines crossed into mutant backgrounds were analyzed in homozygous condition so as to easily identify misexpression in mutant embryos.

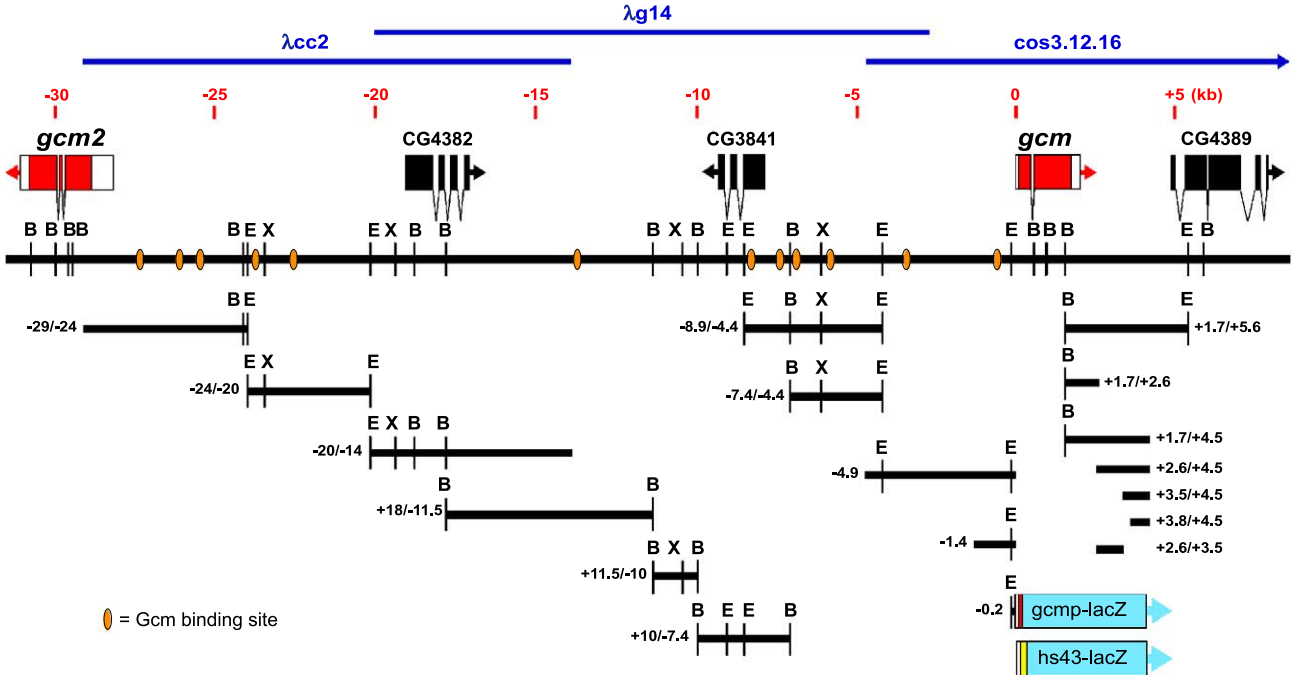

Fig. 1. Structure of the genomic region at the *gcm* locus; *gcm*-reporter constructs. Map of the genomic region 30B–30C, showing locations *gcm*, *gcm2*, and three other predicted genes, derived from sequences and gene predictions from the *Drosophila* Genome Project (GenBank Accession No. AE003625; Adams et al., 2000). Distal to proximal orientation on chromosome 2 is from left to right. From top to bottom. Genomic DNA clones λcc2, λg14, and cos3.12.16 are represented by three overlapping (blue) lines. Arrowheads indicate that cloned DNA extends beyond the bounds of the map. Distance in kilobases is marked in red, with 0 denoting the transcription start site of *gcm*. Predicted transcripts are represented by rectangles, arrows represent direction of transcription; *gcm* and *gcm*2 exons are shown with the shaded part (red) representing coding sequence. Restriction map of genomic DNA indicating sites: B = *Bam*HI, E = *Eco*RI, X = *Xba*I. Ovals represent location of putative Gcm-binding sites (Miller et al., 1998; Ragone et al., 2003). Fragments used for *lacZ* reporter constructs tested in transgenic animals are shown below the restriction map, designated by names to the left or right. *lacZ* vectors that were used to make reporter constructs are shown: *gcmp-lacZ* and *hs-43-lacZ*. Reporter constructs and expression patterns associated with them are described in detail in Table 1.

Immunohistochemistry

Horseradish peroxidase (HRP) immunohistochemistry and embryo dissection were done as previously described (Patel, 1994). Rabbit anti-β-galactosidase (β-gal) antibody (Cappel) was used at 1:10,000 dilution. Mouse anti-Engrailed/Invected monoclonal antibody 4D9 (Patel et al., 1989) was used at a 1:5 dilution. Mouse anti-Repo Monoclonal antibody 8D12 (Alfonso and Jones, 2002) was used at a 1:5 dilution. Secondary antibodies conjugated to HRP (Jackson Immunoresearch) were used at 1:300 dilutions. Secondary antibodies were detected using HRP/Diaminobenzidine (DAB) reaction. DAB reactions were enhanced to give a black color by addition of 0.067% $NiCl_2$.

Results

gcm expression in the embryo

The expression pattern of *gcm* in the embryo has been well described (Alfonso and Jones, 2002; Bernardoni et al., 1997; Freeman and Doe, 2001; Hosoya et al., 1995; Jones et al., 1995; Umesono et al., 2002; Vincent et al., 1996). These studies followed *gcm* expression by in situ localization of *gcm* transcripts, by antisera against Gcm protein, and by the expression of the more persistent β-gal protein found in the gcm^{rA87} enhancer trap embryos that marks cells that transiently express *gcm*. We summarize these data as follows.

gcm transcripts and protein are first detected in the embryo at stage 5 in an anterior ventral region of the blastoderm that comprises the hemocyte anlagen. During gastrulation, these cells invaginate at the end of the ventral furrow to give rise to the hemocyte primordium. *gcm* continues to be expressed in hemocytes as they develop into macrophages and begin to migrate from the cephalic region through stage 12, after which *gcm* expression fades. Persistent expression of β-gal in gcm^{rA87} enhancer trap embryos shows that *gcm* is expressed in all hemocytes of the plasmatocyte/macrophage lineage.

gcm is first detected in glial progenitors at early stage 10 in the longitudinal glioblast or glial precursor (GP) that gives rise to the longitudinal glia (Jacobs et al., 1989; Jones et al., 1995; Hosoya et al., 1995; Schmid et al., 1999; Schmidt et al., 1997). Shortly thereafter, *gcm* is detected in the neural precursor 6–4; in thoracic segments this precursor generates both glia and neurons, and is called NGB 6–4T, whereas in abdominal segments it produces only glia and is called GB 6–4A (Akiyama-Oda et al., 1999; Bernardoni et al., 1999; Freeman and Doe, 2001; Hosoya et al., 1995; Schmid et al., 1999; Schmidt et al., 1997). From late stage 10 through stage 13, *gcm* is expressed in the glial progeny of all other neural precursors that generate mixed lineages of glia and neurons (except for the midline glia); these include thoracic and abdominal NGBs 1–1A, 1–3, 2–2T, 2–5, 5–6, and 7–4 that produce CNS cell body and subperineurial glia and the CNS-derived peripheral glia that ensheath the axons of the PNS nerve tracks. *gcm* is also expressed in glial precursors of the head and gnathal segments. *gcm* continues to be expressed in all developing lateral glia as they differentiate through early stage 16, after which expression fades.

In addition, *gcm* is expressed in a subset of PNS lineages derived from sensory organ precursors (SOPs). In the abdomen, *gcm* is expressed at stages 12 through 13 in the ligament cells of the of the lateral chordotonal organs, the glial support cells of the dorsal bipolar dendrite (dbd) neuron, and the SOP-derived dorsolateral peripheral glia (dlPG, also known as PG3) that ensheaths the distal portion of the intersegmental nerve (Sepp et al., 2000; Umesono et al., 2002). Expression of *gcm* in the PNS of the thorax and head has not been well described.

Finally, *gcm* is transiently expressed in a lateral stripe of cells in the epidermis of each segment from stages 12 through 15. By position and morphology this expression is likely to be in the primordium of the intersegmental apodemes, a single row of cells in the segmental furrow that serve as attachment sites for muscles (Campos-Ortega and Hartenstein, 1997). Expression of β-gal in the intersegmental apodemes of gcm^{rA87} enhancer trap embryos supports this conclusion (Granderath et al., 2000).

Identification of cis-regulatory elements of gcm

To map transcriptional control elements that produce the *gcm* expression pattern, we dissected approximately 35 kb of DNA flanking the *gcm* transcription unit into fragments, fused these to *lacZ* reporter genes, and introduced the gene fusions into flies via P-element-mediated transformation. We then assayed the protein expression patterns driven by these gene fusions in transgenic embryos using anti-β-gal antibodies.

Many fragments were initially tested using a vector containing a minimal *hsp70* promoter fused to the *lacZ* gene (abbreviated *hs43-lacZ*). Concerned that we may be missing important expression elements without the endogenous *gcm* promoter, which lacks a TATA box, we also constructed a vector containing *gcm* promoter sequences fused to *lacZ* (abbreviated *gcmp-lacZ*). In *gcmp-lacZ*, the first 17 amino acid codons of *gcm* are fused in frame with β-gal, which includes the endogenous ATG, 5′ UTR and 223 base pairs of flanking DNA (see Materials and methods). We found no significant differences in the expression patterns of constructs using these two different vectors, though we consistently observed more robust β-gal expression with *gcmp-lacZ–gcm* genomic DNA gene fusions. In addition, *gcmp-lacZ* has its own consistent expression pattern elements (described below).

Fig. 1 depicts a map of approximately 40 kb of DNA from the *gcm* region and the fragments used in this study.

This genomic region spans from *gcm2*, located 28 kb 5′ to *gcm*, to CG4389, located 4 kb 3′ of *gcm*. *gcm2* encodes a homolog of *gcm* that is expressed at very low levels in a similar pattern to *gcm* (Alfonso and Jones, 2002; Kammerer and Giangrande, 2001). CG4389 encodes a long-chain enoyl-CoA hydratase that is maternally expressed (E.M. Jolicoeur, unpublished data). Evenly spaced between *gcm* and *gcm2* are two predicted genes encoding for carboxylesterases, CG4382 and CG3841; Northern analysis indicates that these genes are not expressed at significant levels in the embryo (B.W. Jones, unpublished data).

In the 28 kb separating *gcm* and *gcm2*, there are many sequences that either match or carry at least one mismatch from the consensus Gcm binding site (GBS), AT(G/A)CGGG(T/C). Most of these potential GBS variants can bind Gcm protein and promote Gcm-dependent transcription to some degree in tissue culture, yet they have not been tested in the organism (Akiyama et al., 1996; Miller et al., 1998; Ragone et al., 2003; Schreiber et al., 1997, 1998). GBSs in the 5′ region of *gcm* have been presented as evidence for autoregulation (Miller et al., 1998). In the 8.9 kb 5′ region, between *gcm* and CG3841, we find three GBSs that are a perfect match to consensus, and up to six more that carry at least one mismatch; three of these mismatched GBSs have been shown to bind with moderate to high affinity to Gcm protein (Ragone et al., 2003). A second cluster of GBSs is near the promoter of *gcm2*, suggesting regulation of *gcm2* by *gcm*. We show in Fig. 1 only those GBSs that are a perfect match, or have been shown to bind Gcm in vitro.

A summary of the fragment/vector combinations generated for this study, as well as the number of lines analyzed, and their expression patterns, is shown in Table 1. Most fragments from −8.9 to +5.6 kb with respect to the *gcm* promoter were tested as gene fusions with both *hs43-lacZ* and *gcmp-lacZ*, yielding similar results. Fragments further upstream (−29 kb to −7.4 kb) were tested only with the *gcmp-lacZ* vector.

Table 1
Expression of *gcm-lacZ* gene fusion constructs shown in Fig. 1

Genomic fragment	Reporter vector	No. of lines	Early embryonic stages (5–9)	Late embryonic stages (9–16)
gcm −0.2	gcmp-lacZ	8	*st. 8,9: hemocyte primordium (weak, 4 out of 8 lines)*	*st. 9–16: lateral epidermis (weak), CNS (weak)*
gcm −1.4	gcmp-lacZ	4	st. 8,9: hemocyte primordium	st. 9–16: lateral epidermis (weak), CNS (stronger) CNS expression strongest at st. 12
gcm −4.9	hs43-lacZ	7	st. 8,9: hemocyte primordium	st. 9–16: lateral epidermal stripe at intersegmental furrow st. 11–16: CNS-derived glia, dorsal PG
gcm −4.9	gcmp-lacZ	6	st. 6–9: hemocyte primordium	st. 9–16: lateral epidermal stripe at intersegmental furrow st. 11–16: CNS-derived glia, dorsal PG
gcm −7.4/−4.4	hs43-lacZ	8	–	st. 9–16: hemocytes, macrophages st. 11–16: CNS-derived glia, no dorsal PG
gcm −7.4/−4.4	gcmp-lacZ	6	st. 6–9: hemocyte primordium	st. 9–16 lateral epidermal stripes (weak) st. 9–16: hemocytes, macrophages st. 11–16: CNS-derived glia, no dorsal PG
gcm −8.9/−4.4	hs43-lacZ	8	–	st. 9–16: hemocytes, macrophages st. 11–16: CNS-derived glia, (no dorsal PG)
gcm −8.9/−4.4	gcmp-lacZ	3	st. 6–9: hemocyte primordium	st. 9–16 lateral epidermal stripes (weak) st. 9–16: hemocytes, macrophages st. 11–16: CNS-derived glia, (no dorsal PG)
gcm −10/−7.4	gcmp-lacZ	3	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)*
gcm −11.5/−10	gcmp-lacZ	4	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)*
gcm −18/−11.5	gcmp-lacZ	7	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)* SOP glia: lig., dbd, and dorsal PG (weak, 4 out of 7 lines)
gcm −20/−14	gcmp-lacZ	6	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)* SOP-derived glia: lig., dbd, and dorsal PG (weak, 2 out of 6 lines)
gcm −24/−20	gcmp-lacZ	4	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)*
gcm −29/−24	gcmp-lacZ	3	*st. 8,9: hemocyte primordium (weak)*	st. 9–16: epidermal stripes—intersegmental apodemes
gcm +1.7/+5.6	hs43-lacZ	5	–	–
gcm +1.7/+5.6	gcmp-lacZ	3	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)* st. 11–16: GB6–4A glia (M-CBG, MM-CBG) (weak)
gcm +1.7/+2.6	gcmp-lacZ	2	*st. 8,9: hemocyte primordium (weak)*	*st. 9–16: lateral epidermis (weak), CNS (weak)*
gcm +1.7/+4.5	hs43-lacZ	5	–	st. 11–16: GB6–4A glia (M-CBG, MM-CBG)
gcm +2.6/+4.5	hs43-lacZ	10	–	st. 11–16: GB6–4A glia (M-CBG, MM-CBG)
gcm +3.5/+4.5	hs43-lacZ	3	–	st. 11–16: GB6–4A glia (M-CBG, MM-CBG)
gcm +3.8/+4.5	hs43-lacZ	6	–	st. 11–16: GB6–4A glia (M-CBG, MM-CBG)
gcm +2.6/+3.5	hs43-lacZ	8	–	–

For each transgenic line, the vector used, the number of lines analyzed, and a brief description of tissues expressing *lacZ* at early and late embryonic stages are listed. Expression patterns that are associated with *gcmp-lacZ* vector are denoted with italics.

Our results are summarized as follows. Most of the activity reproducing elements of the *gcm* expression pattern—in hemocyte progenitors, in CNS derived glial cells (including NGB derived peripheral glia), and in epidermal stripes—can be found in the DNA regions immediately flanking *gcm* bounded by CG3841 and CG4389, from −8.9 kb to +5.6 kb with respect to the *gcm* promoter. None of these elements, however, reproduced expression in SOP-derived glia, which include the ligament cells of the lateral chordotonal organs, and the dbd glia. Our survey of fragments of the region further upstream, from CG3841 to *gcm2* (−8.9 to −29 kb) showed little consistent activity except that associated with the *gcmp-lacZ* vector. While we found two large overlapping fragments that promote weak activity in SOP-derived glia between −20 and −11.5 kb relative to the *gcm* promoter, the weak expression and variability in these lines do not provide much confidence that they represent bona fide regulatory elements for the SOP-derived glia (data not shown).

The only other upstream fragment gene fusion with reproducible expression is *gcm −29/−24-gcmp-lacZ*, which has part of the *gcm2* gene, including the first 28 amino acid codons, transcription start site, and 3 kb of flanking DNA with orientation preserved in opposite direction to the *gcmp-lacZ* promoter; these lines consistently express β-gal in the lateral epidermal stripes that become the intersegmental apodemes (data not shown). Despite the presence of three consensus GBSs, and up to six more GBSs with one mismatch, we observed no glial expression in *gcm −29/−24* embryos.

In the following sections, we describe our results in more detail, dividing the focus on three regions of *gcm* flanking DNA that reproduce most of the expression pattern elements of *gcm*: (1) −4.9 kb relative to the *gcm* promoter; (2) −8.9 to −4.4 kb; and (3) +1.7 to +5.6 kb.

gcm −4.9 lacZ reporters drive expression in lateral glial cells, hemocyte primordium, and epidermis

Two separate constructs were analyzed containing sequences −4.9 kb upstream of *gcm*. *gcm −4.9-hs43-lacZ* (with genomic sequence −4873 to +32) includes start site and part of the 5′ UTR cloned into *hs43-lacZ*, and *gcm −4.9-gcmp-lacZ* (with genomic sequence −4873 to +201) includes the endogenous promoter and part of the ORF fused to *lacZ*. Both constructs drive β-gal expression in near identical patterns, with *gcmp-lacZ* transgenes showing slightly earlier and stronger expression.

A line carrying the *gcm −4.9*-hs43-lacZ construct is shown in Figs. 2D–J. β-gal expression in hemocyte progenitors is detected at stage 8 (stage 6 in *gcmp-lacZ* lines), and becomes strongest between stages 9 (Fig. 2D) and 11, after which it fades. Glial expression of β-gal in *gcm −4.9* embryos is first detected at late stage 10. Patches of epidermal expression begin at stage 11, and by stage 13 (Fig. 2G) they resolve into broad stripes, several cell bodies wide, spanning the segmental furrow in each segment; by contrast, normal *gcm* expression is restricted to only a single row of cells in the segmental furrow of each segment. By stage 13, β-gal is expressed in nearly all CNS-derived glia cells, including NGB-derived peripheral glia (Fig. 2G). Both glial expression and epidermal expression persist through stage 16 (Figs. 2H,J and 3A).

Specific CNS expression of gcm −4.9 lacZ does not require endogenous gcm activity

Given the presence of two high-affinity GBSs in the *gcm −4.9* sequence, it was important to test the expression of *gcm −4.9* reporters in a *gcm* mutant background. In *gcm* mutant embryos that also carry the *gcm −4.9* reporter, we found that β-gal expression is initiated in the CNS, but instead of glial-specific expression, β-gal is expressed in the cell bodies of specific neurons (Fig. 3B). In *gcm* mutant embryos, CNS glia are transformed into neurons (Hosoya et al., 1995). We conclude that *gcm −4.9* expression is initiated independent of *gcm*, in presumptive glial cells that are transformed into neurons, and that factors that can initiate *gcm* expression in most CNS glia must act on the −4.9-kb fragment.

Proximal elements drive pan-neural, weak epidermal, and early hemocyte expression

We also made constructs that delete distal parts of the *gcm −4.9* fragment. Deleting 3.5 kb of *gcm −4.9 gcmp-lacZ* to create *gcm −1.4* (Fig. 2A) causes glial-specific expression to be lost; however, β-gal is now expressed at significant levels in all CNS cells except for the midline; this broad CNS expression peaks at stages 12 through 13, with expression strongest in the lateral CNS (Figs. 2B,C). *gcm −1.4* embryos also show weak epidermal expression, as well as expression in the hemocyte primordium similar to *gcm −4.9* (data not shown). Finally, our *gcmp-lacZ* promoter construct, which contains the endogenous ATG, 5′ UTR and 223 base pairs of flanking DNA (*gcm −0.2*, Fig. 2A), has very weak broad expression in the nervous system and epidermis, and weak early hemocyte expression (data not shown); in addition, most of the constructs we made using the *gcmp-lacZ* construct have weak neural, epidermal, and sometimes stronger early hemocyte expression (see Table 1).

We conclude that proximal to the *gcm* promoter are elements for head mesoderm/hemocyte primordium-specific expression, as well as weak neural and epidermal promoting elements; slightly more distal are elements for weak general epidermal expression and stronger general CNS expression; and, in a region more distal to the *gcm* promoter between −4.9 and −1.4 kb lie elements that promote *gcm*-independent CNS glial-specific activity, as well as repressing general neuronal activity.

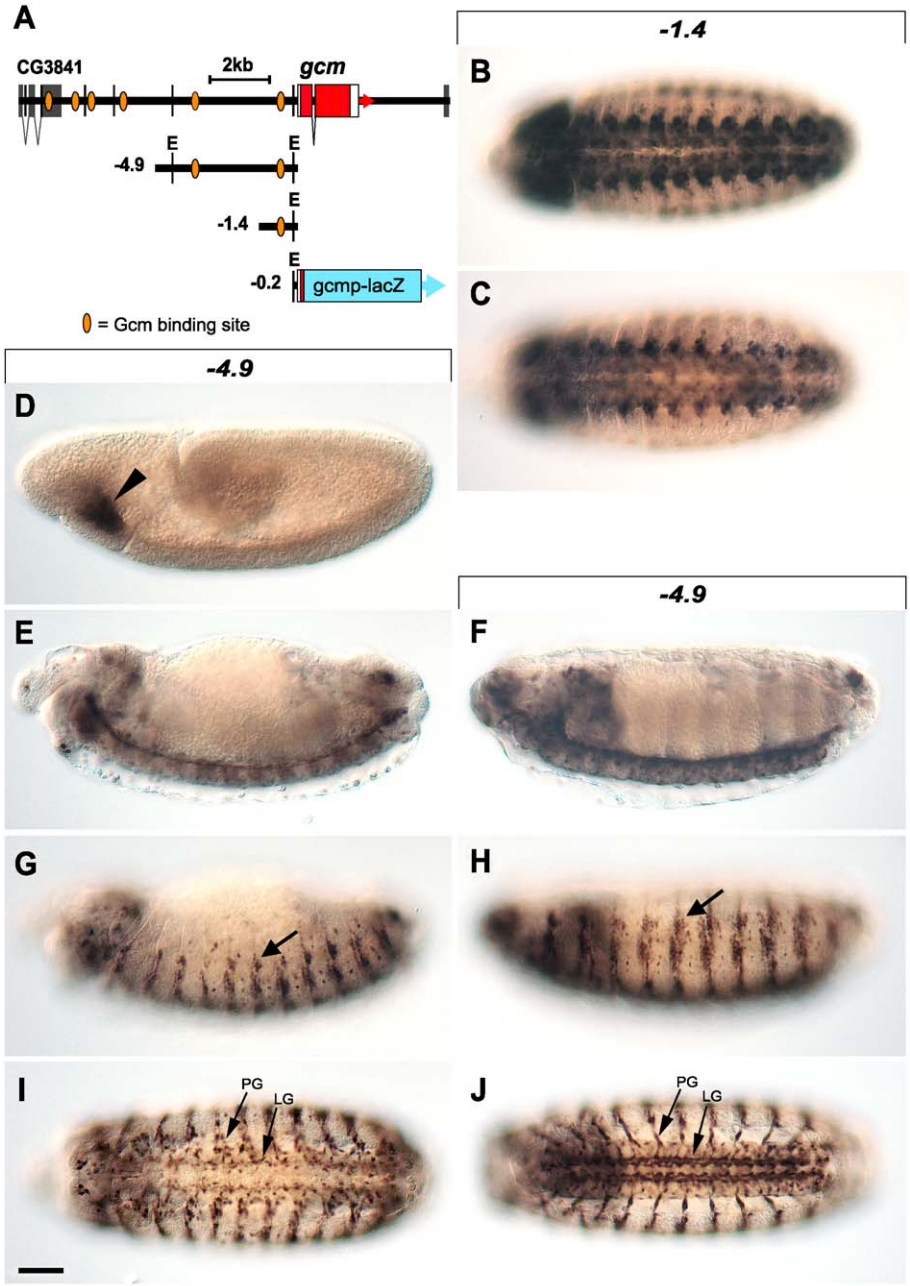

Fig. 2. β-gal expression of *gcm-lacZ* reporter genes: 5′ proximal region constructs. *gcm −4.9* transgenes recapitulate many elements of *gcm* expression. *gcm −1.4* transgenes show general neural expression. (A) Genomic map and reporter constructs. (B–J) Whole-mount transgenic embryos labeled with anti-β-gal (anterior to left). (B,C) Ventral views of *gcm −1.4-gcmp-lacZ* embryos at stage 12 (B) and stage 13 (C) show general neural expression, strongest in the lateral CNS. (D–J) *gcm −4.9-hs43-lacZ* embryos. (D) Stage 9, lateral view. Arrowhead indicates expression in head mesoderm/hemocyte primordium. (E, G, I) Stage 13 embryos. (F, H, J) stage 16 embryos. (E, F) Lateral views, focused on CNS, show glial expression. (G, H) Lateral views, focused on epidermis, show expression in epidermal stripes (arrows). (I, J) ventral views show expression in CNS-derived glia, including peripheral glia (PG) and longitudinal glia (LG). Scale bar, 50 μm.

Elements between −8.9 and −4.4 kb drive gcm-dependent glial expression and macrophage expression

Reporter constructs with genomic DNA between −8.9 and −4.4 kb—further distal to the *gcm* promoter—drive strong CNS glial-specific expression and strong hemocyte expression. In *gcm −8.9/−4.4-hs-43* embryos, β-gal is expressed strongly in hemocytes as they develop into macrophages (Fig 4A, arrowhead); expression persists in macrophages in subsequent stages (Figs. 4B–E). In *gcm −8.9/−4.4-hs-43* embryos, β-gal expression in CNS-glia can first be seen at stage 11 (Fig 4B, arrow), and persists through later stages (Figs. 4B–E). *gcm −7.4/−4.4-hs-43* reporters, deleting the distal part of the −8.9-kb region

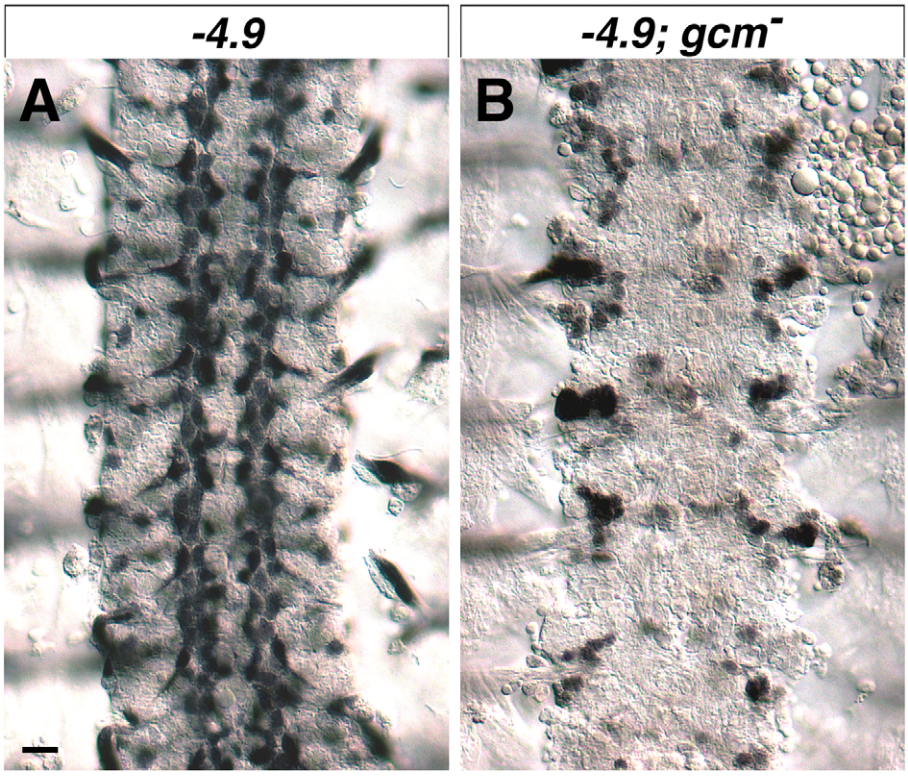

Fig. 3. *gcm −4.9* transgene expression is not dependent on *gcm*. Dissected stage 16 embryos showing seven adjacent segments of the CNS labeled with anti-β-gal (anterior up). (A) *gcm −4.9* transgenic embryos show β-gal expression in CNS-derived glia. (B) *gcm −4.9; gcm* null mutant embryos: β-gal labels specific neurons. In *gcm* mutant embryos, expression is initiated in the CNS in presumptive glia that are transformed into neurons. Scale bar, 10 μm.

(containing the first exon of the neighboring gene and two GBSs), showed identical staining to *−8.9/−4.4-hs-43* embryos (data not shown). Constructs of both fragments using *gcmp-lacZ* vector showed similar expression patterns, but with much stronger and earlier hemocyte-specific expression (data not shown), suggesting that there is synergy between early hemocyte elements proximal to the promoter and late hemocyte/macrophage-specific elements in the −7.4 to −4.4 kb region; in addition, these lines show epidermal staining associated with the *gcmp-lacZ* vector.

When *gcm −8.9/−4.4* and *gcm −7.4/−4.4* reporters are crossed into a *gcm* null background, all CNS expression disappears (Figs. 4G–I); unlike *gcm −4.9*, there is no staining in CNS neurons, indicating that *gcm* is required for initiating CNS expression in these reporters. Hemocyte/macrophage expression persists in *gcm* mutant embryos, revealing the existence of a *gcm*-independent hemocyte enhancer in the −7.4 kb to −4.4 kb region. The presence of two high-affinity GBSs and as many as three lower-affinity GBSs in the −7.4 to −4.4 sequence suggests that Gcm directly regulates glial-specific expression from this region. We conclude that the region between −7.4 and −4.4 kb contains both *gcm* autoregulatory elements and *gcm*-independent hemocyte elements.

Identification of GB 6−4A element in the gcm 3′ region

We initially tested a 3.9-kb fragment 3′ of *gcm* for activity using the *hs43-lacZ* vector (*gcm +1.7/+5.6*, Fig. 5A); we found no activity from this fragment. However, after realizing that it contained the promoter region of the adjacent gene CG4389, we cut back this element to +4.5, and found that it drives β-gal expression in a subset of abdominal glia, and maxillary glia (Figs. 5B–C). Using several overlapping constructs (Fig. 5A), we narrowed this activity down to a 742-bp fragment that is located midway between *gcm* and its neighboring 3′ gene (from +3777 to +4519 relative to the *gcm* promoter). We also cloned the original fragment (which did not show expression in the *hs43-lacZ* vector) into the *gcmp-lacZ* vector. *gcm+1.7/+5.6-gcmp-lacZ* embryos show β-gal expression in the same subset of glia as the other fragments from this region; in addition, these embryos show early hemocyte expression and epidermal expression associated with *gcmp-lacZ* (Table 1, data not shown). We conclude that enhancer activity in the *gcm +1.7/+5.6* fragment can work preferentially on the *gcm* promoter.

Expression from a reporter line showing this glial-specific activity is depicted in Fig. 5. β-gal is first detected at stage 11 in a single neural precursor in each

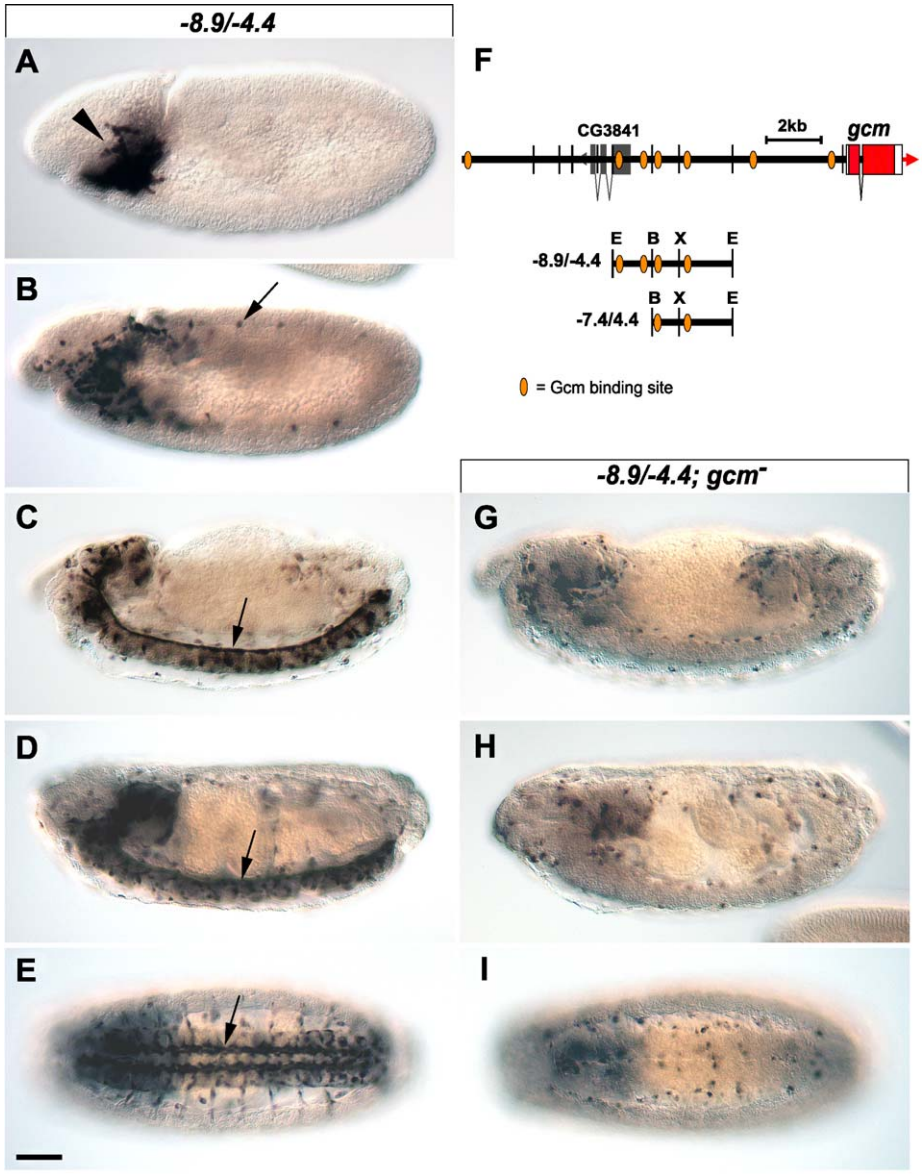

Fig. 4. *gcm*-dependent glia expression elements and hemocyte elements in 5′ distal region constructs. (A–E) Whole-mount *gcm −8.4/−4.4−hs43 lacZ* embryos labeled with anti-β-gal (anterior to left). (A) Stage 10, lateral view. Arrowhead indicates hemocyte expression. (B) Stage 11, lateral view. Glial precursor expression (arrow). (C) Stage 14, lateral view, shows CNS-derived glial expression (arrow). (D) Stage 16, lateral view and (E) ventral view. Expression persists in all CNS-derived glia (arrows) and hemocytes. (F) Genomic map and reporter constructs. (G–I) *gcm −8.4/−4.4*; *gcm* null mutant embryos; same stages and views as (C–E). In *gcm* mutant embryos, expression is not initiated in the CNS; they lack glial and CNS expression, but retain hemocyte expression. Scale bar, 50 μm.

hemisegment as it divides, in abdominal segments 1 through 7, and in a similar neural precursor in the maxillary segment (Figs. 5B and 6A,C). We identify this precursor in abdominal segments as GB 6–4A, by position, colabeling with Engrailed and Repo (data not shown), and by its glial progeny. 6–4A is a pure glioblast that is positioned as the most anterior and lateral Engrailed expressing cell in the neuroblasts array, which divides once to give rise to two glia—the most-medial cell body glia (MM-CBG) and the medial cell body glia (M-CBG) (Broadus et al., 1995; Ito et al., 1995; Schmidt et al., 1997). Expression driven by the *gcm* 3′ enhancer element follows this pattern, and is specific to the 6–4A lineage (Figs. 5B–F).

The homeotic gene abdominal-A is required for expression of the gb 6–4A element

The homeotic gene *abdominal-A* (*abd-A*) is expressed in the abdominal segments from the posterior Engrailed stripe of A1 through A7 (Macias et al., 1990; Karch et al., 1990) and is required for abdominal segment identity

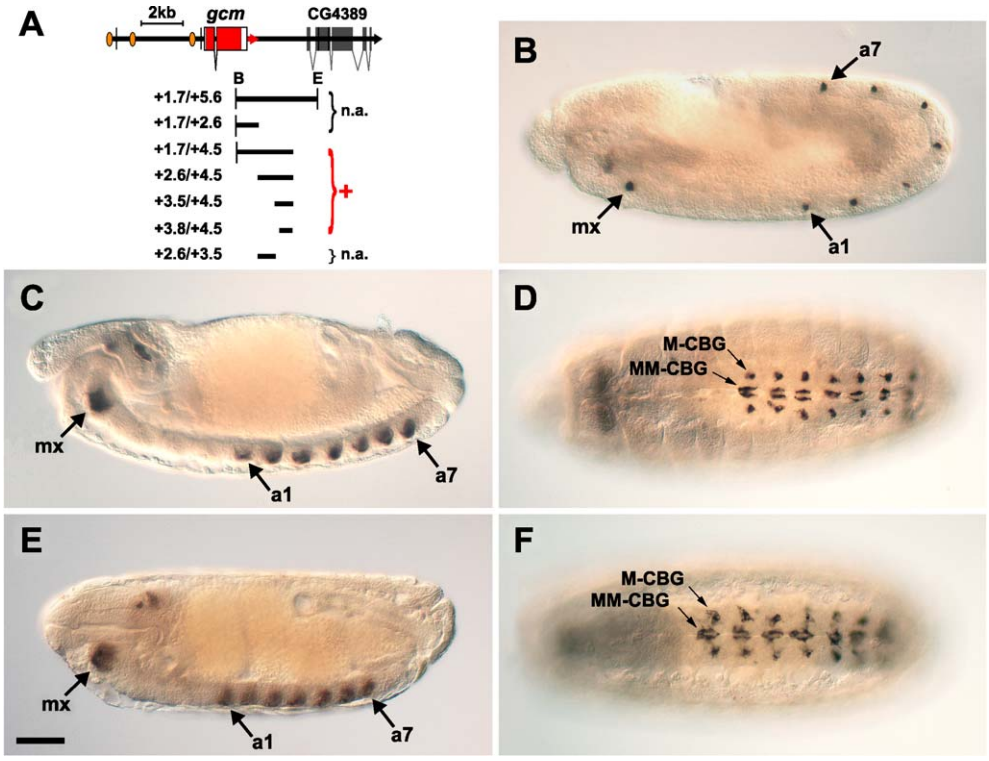

Fig. 5. Abdominal CNS glioblast-specific enhancer located 3′ of *gcm*. (A) Genomic map and reporter constructs. Constructs with no activity are designated "n.a.", and constructs with identical activity are designated with a red plus symbol. (C–F) Whole-mount *gcm+1.7/+4.5-hs43-lacZ* embryos labeled with anti-β-gal (anterior to left). Expression first appears at late stage 10 (B) in GB 6-4A, and in a similar precursor in the maxillary segment, as they begin to divide to give rise to abdominal glia M-CBG and MM-CBG and a pair of maxillary glia. (C, D) Stage 14, lateral (C) and ventral (D) views. (E, F) Stage 16, lateral (E) and ventral (F) views. mx, maxillary segment; a1, abdominal segment 1; a7, abdominal segment 7. Scale bar, 50 μm.

(Sanchez-Herrero et al., 1985). Since the *gcm* 3′ element is expressed in the GB 6–4A lineage in A1 through A7 in the same domain as *abd-A*, we decided to test if *abd-A* is required for this expression. We found that *abd-A* mutant embryos carrying the *gcm +1.7/+4.5* reporter show a complete loss of β-gal expression in abdominal segments, while maintaining expression in the maxillary segment (Figs. 6B,D).

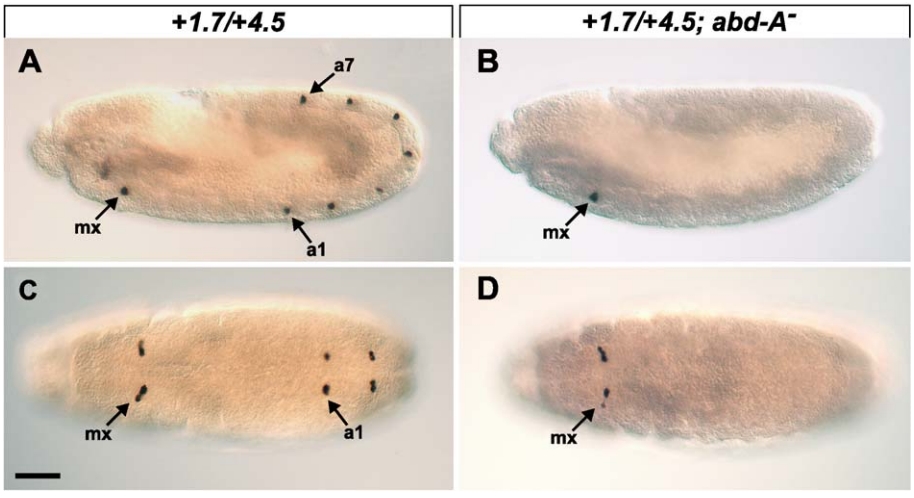

Fig. 6. Expression of 3′ enhancer in GB 6–4A requires the homeotic gene *abd-A*. Whole-mount embryos at late stage 10 labeled with anti-β-gal (anterior to left) (A, C) *gcm +1.7/+4.5-hs43-lacZ* embryos, lateral view (A), and ventral view (C). (B, D) *gcm +1.7/+4.5-hs43-lacZ; abd-A* null mutant embryos, lateral view (B), and ventral view. In *abd-A* mutant embryos, expression in the GB 6–4A lineage is lost, while expression in the maxillary segment lineage is maintained. mx, maxillary segment; a1, abdominal segment 1; a7, abdominal segment 7. Scale bar, 50 μm.

Discussion

We have completed a survey of the *gcm* locus for *cis*-regulatory enhancer activity using *lacZ* reporter fusion genes in vivo, and have mapped DNA regions that promote most of the elements of the *gcm* expression pattern (summarized in Fig. 7). These elements can be separated into distinct modules that reveal the logic underlying the transcriptional regulation of *gcm*. During development, *gcm* is primarily expressed in two separate tissues: neuroectoderm-derived glial cells and mesoderm-derived hemocytes. Expression in these tissues appears to require both general control elements and lineage-specific control elements. Close to the *gcm* promoter are weak and broadly expressed tissue-specific control elements, while more distant from the promoter lie lineage-specific control elements.

In the case of hemocyte expression, we find two distinct regions that promote expression in early and late phases of hemocyte development, respectively. Close to the *gcm* promoter (between −0.2 kb and the transcription start site) are sequences that promote the early, germ-layer phase of *gcm* expression in the procephalic mesoderm domain that comprises the hemocyte primordium. Expression in this procephalic mesoderm domain is weak and transient. Based on the timing and position of expression, activity driven by this early element likely depends on position-specific cues that pattern the blastoderm along the dorsal–ventral and anterior–posterior axis. More distal to the promoter, between −7.4 and −4.4 kb, are sequences that promote later hemocyte lineage-specific expression. Based on the timing and specificity of this expression, activity driven by this late element is likely dependent on factors already expressed in the hemocyte lineage. Reporter constructs that carry both elements act synergistically to produce more robust hemocyte expression.

In the case of glial cell expression, we find at least three different components promoting activity: a general neural component, a lineage-specific component, and an autoregulatory component. Close to the *gcm* promoter (between −1.4 and −0.2 kb) are sequences that promote broad, but weak, pan-neural expression. More distal to the promoter, as well as downstream of the *gcm* transcription unit, are lineage-specific elements. It appears that most of the activity required to initiate expression in CNS glial lineages is in the distal 3-kb portion of the first 5 kb (between −4.4 and −1.4) upstream of the *gcm* promoter. Adding this 3-kb region to the proximal pan-neural promoting region causes a loss of neuronal expression, indicating that, besides promoting glial-specific expression, sequences in this element are required to repress general neuronal activity. These data support a model whereby glial-specific *gcm* expression in the CNS requires both lineage-specific activation and general neuronal repression.

Additional lineage-specific elements are in the 3′ downstream region, and, perhaps, in upstream regions as far as −18 kb from the promoter. In the 3′ region of *gcm* (+3.8 to +4.5 kb), we found an element that promotes expression in the 6–4A glioblast lineage as well as a similar glial lineage in the maxillary segment. These precursors generate two glial cells in each hemisegment: the most-medial cell body glia (MM-CBG) and the medial cell body glia (M-CBG). Interestingly, expression is absent in the analogous precursor, NGB 6–4T, found in the thoracic segments; NGB 6–4T undergoes a slightly different pattern of divisions, generating both neurons and glia. Here we show that expression driven by the GB 6–4 element in the abdomen is dependent on the homeotic gene *abdominal-A*, which demonstrates that some of the segmental differences in the specification of glial precursors along the anterior–posterior axis are under the control of the homeotic genes of the *Hox* complex; these data hold out the possibility that they do so by directly modulating *gcm* expression.

None of the elements described above account for expression in PNS glia that are derived from SOPs. Our data suggests that expression in SOP-derived glia may be driven by other enhancer elements, and we speculate that they may be in a region as distant as −18 kb from the promoter, as several of our reporter constructs from the region midway between *gcm* and *gcm2* show weak expression in SOP-derived glia (data not shown). Whether these are bona fide enhancer elements of *gcm* is not clear, as different transgenic lines show much variability in expression driven from these constructs. It is possible that additional *gcm* proximal promoter sequences are required to enable a strong response from these distant enhancer elements. In a study of the large and complex regulatory region of the mitotic regulator encoded by the *string* locus, it was found that promoter proximal sequences as far away as −525 base pairs from the transcription start site are required to allow a response specifically from distant

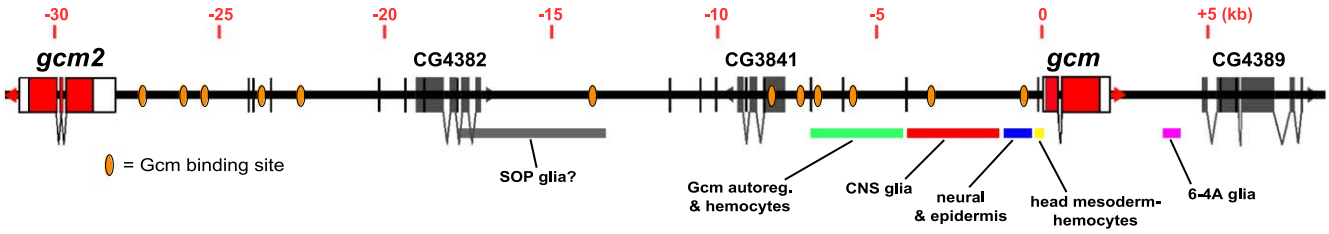

Fig. 7. Summary: DNA regions promoting elements of the *gcm* expression pattern. Regions promoting specific activities are shown as colored bars beneath the restriction map, and are described in detail in the text.

enhancers (Lehman et al., 1999). Alternatively, SOP-derived glia expression may require the synergistic or additive interaction between multiple elements that are in separate regions.

The third component contributing to *gcm* glial expression is autoregulation. Previous studies implicate a positive autoregulatory feedback mechanism in *gcm* transcriptional regulation (Miller et al., 1998). These studies showed that *gcm* transcription is initiated, but not maintained, in a mutant allele ($gcm^{N7\text{-}4}$) that carries a point mutation in the DNA binding domain. This point mutation creates an amino acid change that abolishes the ability of Gcm protein to bind DNA in vitro. In addition, ectopic expression of *gcm* is able to induce ectopic β-gal in a *gcm* enhancer trap line. This evidence, and the presence of multiple Gcm protein binding sites (GBSs) in the 5′ region of *gcm*, implicates direct autoregulation of *gcm*. In support of these observations, we describe elements that drive *gcm*-dependent glial cell expression; these elements are located further distal (−7.4 to −4.4 kb) to the region that initiates *gcm*-independent glia cell expression, and the sequence contains two high-affinity GBSs and as many as three or more lower-affinity GBSs. Interestingly, hemocyte lineage-specific elements that are also found in this region are not dependent on *gcm* for their expression, suggesting that autoregulation is also context dependent and may require the cooperation of additional factors.

Our expression data using *gcm-lacZ* reporter gene fusions agree with a recently published study that used genomic transgenes to rescue *gcm* mutant phenotypes (Ragone et al., 2003). These genomic rescue experiments used four constructs, containing the *gcm* coding sequences, 1.7 kb of downstream, and 2, 4, 6, and 9 kb of upstream sequences, respectively. They showed that the 4-kb upstream transgene was able to rescue glial cell differentiation in most CNS glial lineages; however, the ability to rescue was not fully penetrant, and required additional upstream DNA to provide "quantitative" elements. Full rescue in nearly all CNS glial lineages was only achieved with 9 kb of upstream DNA. These data, together with our studies, suggest that this quantitative regulation is provided by *gcm* autoregulatory elements. Interestingly, the genomic rescue transgenes gave very poor rescue of the 6–4 lineages, and we observe that they also lack the 3′ 6–4A regulatory elements identified in our study. A comparison of these two studies suggests there are as yet unidentified elements required for high levels of *gcm* expression in the NGB 6–4T lineage. Another interesting comparison between these two studies is the rescue provided by the 2-kb genomic transgene vs. the expression driven by our *gcm −1.4* reporter construct. In the genomic rescue studies, the 2-kb upstream transgene rescued glial differentiation in a defined subset of CNS glial lineages, but also showed precocious *gcm* expression and glial transformation of several nongliogenic neuroblasts. Our *gcm −1.4* reporter construct, which differs in length from the 2-kb genomic transgene by 740 base pairs, shows broad low-level neural expression. Taken together, we conclude that there are CNS lineage-specific elements in the 740 base pair region that differs between these two constructs, as well as a requirement for neuronal repression provided by elements distal to this region.

While our study was not directed toward understanding the regulation of the *gcm* homolog, *gcm2*, it is interesting to note that we did not find any glial-specific or hemocyte-specific enhancer activity proximal to the *gcm2* promoter. *gcm2* is expressed at very low levels in glial cells and at more pronounced levels in developing hemocytes (Alfonso and Jones, 2002; Kammerer and Giangrande, 2001). Reflecting this expression pattern, the role of *gcm2* in glial cell development is very minor, while it has more significant role in hemocyte development (Alfonso and Jones, 2002). Our data indicates, as previously suggested by Kammerer and Giangrande, 2001, that *gcm2* derives its expression pattern from enhancer elements shared with *gcm* and/or direct regulation by *gcm*.

gcm occupies the most upstream position in lateral glial cell determination. It follows that *gcm* is regulated by a combination of factors not specific to glia. Candidates include a variety of transcriptional regulators expressed in the neuroectoderm, GBs, NGBs, ganglion mother cells, SOPs, and their progeny. Thus, the regulation of *gcm* provides a model for understanding how these factors combine to give rise to stereotypic patterns of glial vs. neuronal progeny. One model for *gcm* transcription is that its regulatory DNA is modular, where each module integrates a combination of inputs to activate *gcm* in different lineages. At one extreme, each glial lineage and glial subtype possibly have separate regulatory modules with unique modes of regulation. Alternatively, *gcm* may have a limited number of regulatory elements that respond to signals present in multiple developmental environments. Our results indicate that *gcm* regulation is controlled by a combination of general and lineage-specific elements. We have found one lineage-specific element that is expressed in a single neural precursor (GB 6–4A), indicating that at least one neural precursor lineage has a unique mode of regulation. In principle, a fine dissection of the 3-kb region that promotes most CNS glial cell expression will possibly reveal additional neural precursor specific elements, as is suggested by the genomic rescue data.

The comparison of regulatory regions between related species has been shown to be a powerful tool to identify important evolutionary conserved *cis*-regulatory DNA sequences. Comparison of the *gcm* locus in *D. melanogaster* to the recently released sequence of a sibling *Drosophila* species, *D. pseudoobscura*, shows the *gcm* locus to be highly conserved, both with respect to gene order and distances between genes. There is high degree of intermittent sequence similarity in the 9 kb upstream of *gcm* in both species, and the gaps of dissimilarity may serve as a guide to further identify the boundaries of

specific *cis*-regulatory sequences. When we compare the 3′ downstream region of *gcm* in both species, we find that the only significantly conserved sequence is located within the 742 base pair fragment that we identified by our *gcm-lacZ* reporter analysis as the GB 6–4A enhancer; this sequence similarity consists of a continuous stretch of approximately 500 base pairs that is 80% conserved, which confirms the predictive value of interspecies sequence comparisons.

We also note that in the first 5 kb of upstream *gcm* sequence in both species are five highly conserved consensus Suppressor of Hairless [Su(H)] binding sites (Bailey and Posakony, 1995). Su(H) is the primary effector of the Notch signaling pathway (reviewed by Artavanis-Tsakonas et al., 1999). Four of these Su(H) sites are clustered in the 3-kb region (−4.4 to −1.4 kb) that promotes most CNS gliogenesis. Notch signaling has been shown to influence *gcm* transcription and glial cell differentiation in binary cell fate decisions, with the general rule that it promotes gliogenesis in the case of neuronal/glial sibling pairs, but has the opposite effect on secondary precursor/sibling pairs (Udolph et al., 2001; Umesono et al., 2002; Van De Bor and Giangrande, 2001). The presence of Su(H) binding sites suggests that Notch-mediated regulation may be directly influencing *gcm* transcription, and future analysis may provide an opportunity to dissect the context-dependent utilization of Notch signaling in glial cell differentiation.

In summary, we have performed a systematic characterization of the *cis*-regulatory regions controlling *gcm* transcription by reporter gene fusions. These studies represent a first step in elucidating how a diverse set of positional, temporal, and lineage-specific factors are integrated at the *gcm* promoter to produce a complex pattern of expression in a precise and stereotypic manner during neural development and hematopoiesis.

Acknowledgments

We thank Lisa Kim, Berenice Alfonso, Melissa Yee, and Debra Liu for technical help; Miyuki Yussa and Bruce Lee for helpful discussions; Corey Goodman for antibodies; and the Bloomington Stock Center for fly stocks. This research was funded in part by the City of New York Council Speaker's Fund for Biomedical Research: Toward the Science of Patient Care, and NIH grant NS39373.

References

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al., 2000. The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195.

Akiyama, Y., Hosoya, T., Poole, A.M., Hotta, Y., 1996. The gcm-motif: a novel DNA-binding motif conserved in *Drosophila* and mammals. Proc. Natl. Acad. Sci. U. S. A. 93, 14912–14916.

Akiyama-Oda, Y., Hosoya, T., Hotta, Y., 1999. Asymmetric cell division of thoracic neuroblast 6–4 to bifurcate glial and neuronal lineage in *Drosophila*. Development 126, 1967–1974.

Alfonso, T.B., Jones, B.W., 2002. gcm2 promotes glial cell differentiation and is required with glial cells missing for macrophage development in *Drosophila*. Dev. Biol. 248, 369–383.

Anderson, D.J., 2001. Stem cells and pattern formation in the nervous system: the possible versus the actual. Neuron 30, 19–35.

Artavanis-Tsakonas, S., Rand, M.D., Lake, R.J., 1999. Notch signaling: cell fate control and signal integration in development. Science 284, 770–776.

Bailey, A.M., Posakony, J.W., 1995. Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to notch receptor activity. Genes Dev. 9, 2609–2622.

Bernardoni, R., Vivancos, B., Giangrande, A., 1997. glide/gcm is expressed and required in the scavenger cell lineage. Dev. Biol. 191, 118–130.

Bernardoni, R., Kammerer, M., Vonesch, J.L., Giangrande, A., 1999. Gliogenesis depends on glide/gcm through asymmetric division of neuroglioblasts. Dev. Biol. 216, 265–725.

Bossing, T., Udolph, G., Doe, C.Q., Technau, G.M., 1996. The embryonic central nervous system lineages of *Drosophila melanogaster*: I. Neuroblast lineages derived from the ventral half of the neuroectoderm. Dev. Biol. 179, 41–64.

Broadus, J., Skeath, J.B., Spana, E.P., Bossing, T., Technau, G., Doe, C.Q., 1995. New neuroblast markers and the origin of the aCC/pCC neurons in the *Drosophila* central nervous system. Mech. Dev. 53, 393–402.

Brody, T., Odenwald, W.F., 2000. Programmed transformations in neuroblast gene expression during *Drosophila* CNS lineage development. Dev. Biol. 226, 34–44.

Campos-Ortega, J.A., Hartenstein, V., 1997. The Embryonic Development of *Drosophila melanogaster*. Springer-Verlag, Berlin.

Davis, A.A., Temple, S., 1994. A self-renewing multipotential stem cell in embryonic rat cerebral cortex. Nature 372, 263–266.

Doe, C.Q., Chu-LaGraff, Q., Wright, D.M., Scott, M.P., 1991. The prospero gene specifies cell fates in the *Drosophila* central nervous system. Cell 65, 451–464.

Doe, C.Q., Fuerstenberg, S., Peng, C.Y., 1998. Neural stem cells: from fly to vertebrates. J. Neurobiol. 36, 111–127.

Frank, E., Sanes, J.R., 1991. Lineage of neurons and glia in chick dorsal root ganglia: analysis in vivo with a recombinant retrovirus. Development 111, 895–908.

Freeman, M.R., Doe, C.Q., 2001. Asymmetric prospero localization is required to generate mixed neuronal/glial lineages in the *Drosophila* CNS. Development 128, 4103–4112.

Goodman, C.S., Doe, C.Q., 1993. Embryonic development of the *Drosophila* central nervous system. In: Bate, M., Martines-Arias, A. (Eds.), The Development of *Drosophila melanogaster*. Cold Spring Harbor Press, New York, pp. 1131–1206.

Granderath, S., Bunse, I., Klambt, C., 2000. gcm and pointed synergistically control glial transcription of the *Drosophila* gene loco. Mech. Dev. 91, 197–208.

Hirata, J., Nakagoshi, H., Nabeshima, Y., Matsuzaki, F., 1995. Asymmetric segregation of the homeodomain protein prospero during *Drosophila* development. Nature 377, 627–630.

Hosoya, T., Takizawa, K., Nitta, K., Hotta, Y., 1995. glial cells missing: a binary switch between neuronal and glial determination in *Drosophila*. Cell 82, 1025–1036.

Isshiki, T., Pearson, B., Holbrook, S., Doe, C.Q., 2001. *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. Cell 106, 511–521.

Ito, K., Urban, J., Technau, G.M., 1995. Distribution, classification and development of *Drosophila* glial cells during late embryogenesis. Roux. Arch. Dev. Biol. 204, 284–307.

Jacobs, J.R., Hiromi, Y., Patel, N.H., Goodman, C.S., 1989. Lineage, mi-

gration, and morphogenesis of longitudinal glia in the *Drosophila* CNS as revealed by a molecular lineage marker. Neuron 2, 1625–1631.

Jan, Y.N., Jan, L.Y., 1993. The peripheral nervous system. In: Bate, M., Martines-Arias, A. (Eds.), The Development of *Drosophila melanogaster*. Cold Spring Harbor Press, New York, pp. 1207–1244.

Jones, B.W., Fetter, R.D., Tear, G., Goodman, C.S., 1995. glial cells missing: a genetic switch that controls glial versus neuronal fate. Cell 82, 1013–1023.

Kambadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S.J., Odenwald, W.F., 1998. Regulation of POU genes by castor and hunchback establishes layered compartments in the *Drosophila* CNS. Genes Dev. 12, 246–260.

Kammerer, M., Giangrande, A., 2001. Glide2, a second glial promoting factor in *Drosophila melanogaster*. EMBO J. 20, 4664–4673.

Karch, F., Bender, W., Weiffenbach, B., 1990. abd-A expression in *Drosophila* embryos. Genes Dev. 4, 1573–1587.

Knoblich, J.A., Jan, L.Y., Jan, Y.N., 1995. Asymmetric segregation of numb and prospero during cell division. Nature 377, 624–627.

Lane, M.E., Kalderon, D., 1993. Genetic investigation of cAMP-dependent protein kinase function in *Drosophila* development. Genes Dev. 7, 1229–1243.

Lebestky, T., Chang, T., Hartenstein, V., Banerjee, U., 2000. Specification of *Drosophila* hematopoietic lineage by conserved transcription factors. Science 288, 146–149.

Lehman, D.A., Patterson, B., Johnston, L.A., Balzer, T., Britton, J.S., Saint, R., Edgar, B.A., 1999. *Cis*-regulatory elements of the mitotic regulator, string/Cdc25. Development 126, 1793–1999.

Macias, A., Casanova, J., Morata, G., 1990. Expression and regulation of the abd-A gene of *Drosophila*. Development 110, 1197–1207.

Miller, A.A., Bernardoni, R., Giangrande, A., 1998. Positive autoregulation of the glial promoting factor glide/gcm. EMBO J. 17, 6316–6326.

Novotny, T., Eiselt, R., Urban, J., 2002. Hunchback is required for the specification of the early sublineage of neuroblast 7–3 in the *Drosophila* central nervous system. Development 129, 1027–1036.

Patel, N.H., 1994. Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. In: Goldstein, L.S.B., Fyrberg, E.A. (Eds.), Methods in Cell Biology, vol. 44. Academic Press, San Diego, pp. 445–487.

Patel, N.H., Martin-Blanco, E., Coleman, K.G., Poole, S.J., Ellis, M.C., Kornberg, T.B., Goodman, C.S., 1989. Expression of engrailed proteins in arthropods, annelids, and chordates. Cell 58, 955–968.

Ragone, G., Van De Bor, V., Sorrentino, S., Kammerer, M., Galy, A., Schenck, A., Bernardoni, R., Miller, A.A., Roy, N., Giangrande, A., 2003. Transcriptional regulation of glial cell specification. Dev. Biol. 255, 138–150.

Rhyu, M.S., Jan, L.Y., Jan, Y.N., 1994. Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. Cell 76, 477–491.

Sanchez-Herrero, E., Vernos, I., Marco, R., Morata, G., 1985. Genetic organization of *Drosophila* bithorax complex. Nature 313, 108–113.

Schmidt, H., Rickert, C., Bossing, T., Vef, O., Urban, J., Technau, G.M., 1997. The embryonic central nervous system lineages of *Drosophila melanogaster*: II. Neuroblast lineages derived from the dorsal part of the neuroectoderm. Dev. Biol. 189, 186–204.

Schmid, A., Chiba, A., Doe, C.Q., 1999. Clonal analysis of *Drosophila* embryonic neuroblasts: neural cell types, axon projections and muscle targets. Development 126, 4653–4689.

Schreiber, J., Sock, E., Wegner, M., 1997. The regulator of early gliogenesis glial cells missing is a transcription factor with a novel type of DNA-binding domain. Proc. Natl. Acad. Sci. U. S. A. 94, 4739–4744.

Schreiber, J., Enderich, J., Wegner, M., 1998. Structural requirements for DNA binding of GCM proteins. Nucleic Acids Res. 26, 2337–2343.

Sepp, K., Schulte, J., Auld, V.J., 2000. Developmental dynamics of peripheral glia in *Drosophila melanogaster*. Glia 30, 122–133.

Stemple, D.L., Anderson, D.J., 1992. Isolation of a stem cell for neurons and glia from the mammalian neural crest. Cell 71, 973–985.

Thummel, C.S., Pirrotta, V., 1992. Technical notes: new pCaSpeR P-element vectors. Drosoph. Inf. Serv. 71, 150.

Thummel, C.S., Boulet, A.M., Lipshitz, H.D., 1988. Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. Gene 74, 445–456.

Turner, D.L., Cepko, C.L., 1987. A common progenitor for neurons and glia persists in rat retina late in development. Nature 328, 131–136.

Udolph, G., Prokop, A., Bossing, T., Technau, G.M., 1993. A common precursor for glia and neurons in the embryonic CNS of *Drosophila* gives rise to segment-specific lineage variants. Development 118, 765–775.

Udolph, G., Rath, P., Chia, W., 2001. A requirement for Notch in the genesis of a subset of glial cells in the *Drosophila* embryonic central nervous system which arise through asymmetric divisions. Development 128, 1457–1466.

Uemura, T., Shepherd, S., Ackerman, L., Jan, L.Y., Jan, Y.N., 1989. Numb, a gene required in determination of cell fate during sensory organ formation in *Drosophila* embryos. Cell 58, 349–360.

Umesono, Y., Hiromi, Y., Hotta, Y., 2002. Context-dependent utilization of notch activity in *Drosophila* glial determination. Development 129, 2391–2399.

Vaessin, H., Grell, E., Wolff, E., Bier, E., Jan, L.Y., Jan, Y.N., 1991. Prospero is expressed in neuronal precursors and encodes a nuclear protein that is involved in the control of axonal outgrowth in *Drosophila*. Cell 67, 941–953.

Van De Bor, V., Giangrande, A., 2001. Notch signaling represses the glial fate in fly PNS. Development 128, 1381–1390.

Vincent, S., Vonesch, J.L., Giangrande, A., 1996. Glide directs glial fate commitment and cell fate switch between neurones and glia. Development 122, 131–139.

Williams, B.P., Price, J., 1995. Evidence for multiple precursor cell types in the embryonic rat cerebral cortex. Neuron 14, 1181–1188.