Developmental Biology **232,** 157–175 (2001)
doi:10.1006/dbio.2001.0174, available online at http://www.idealibrary.com on IDEAL®

Drosophila Single-minded Represses Gene Transcription by Activating the Expression of Repressive Factors

Patricia Estes,[1] Jack Mosher,[1] and Stephen T. Crews[2]

Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7260

The *Drosophila single-minded* gene controls CNS midline cell development by both activating midline gene expression and repressing lateral CNS gene expression in the midline cells. The mechanism by which Single-minded represses transcription was examined using the *ventral nervous system defective* gene as a target gene. Transgenic-*lacZ* analysis of constructs containing fragments of the *ventral nervous system defective* regulatory region identified sequences required for lateral CNS transcription and midline repression. Elimination of Single-minded:Tango binding sites within the *ventral nervous system defective* gene did not affect midline repression. Mutants of Single-minded that removed the DNA binding and transcriptional activation regions abolished *ventral nervous system defective* repression, as well as transcriptional activation of other genes. The replacement of the Single-minded transcriptional activation region with a heterologous VP16 transcriptional activation region restored the ability of Single-minded to both activate and repress transcription. These results indicate that Single-minded indirectly represses transcription by activating the expression of repressive factors. Single-minded provides a model system for how regulatory proteins that act only as transcriptional activators can control lineage-specific transcription in both positive and negative modes. © 2001 Academic Press

Key Words: **bHLH-PAS; CNS midline; neuroectoderm; *single-minded; ventral nervous system defective; tango.***

INTRODUCTION

The *Drosophila* embryonic CNS consists of a linear array of two identical hemiganglia separated by a discrete set of midline cells. The CNS midline cells differentiate into a diverse set of neurons and glia and provide a number of critical developmental functions: they provide signals that guide the paths of axons during growth and regulate the development of the epidermis via the Spitz signaling pathway and serve as an axis of symmetry for the development of the lateral CNS cells (Crews, 1998; Jacobs, 2000). All steps of midline cell development are dependent upon the expression of the *single-minded* (*sim*) gene (Nambu *et al.*, 1991), which acts as a switch to direct neuroectodermal cells to a midline fate. Sim achieves this by both activating midline gene transcription (Nambu *et al.*, 1990) and repressing lateral CNS gene transcription (Chang *et al.*, 1993; Mellerick and Nirenberg, 1995; Xiao *et al.*, 1996). The ability of Sim to activate transcription has been studied extensively, yet little is known about how it represses gene expression.

The neuroectoderm of the early embryo gives rise to the cells of the CNS and can be divided along the dorsal–ventral axis into four domains. The mesectoderm constitutes the ventralmost cells and gives rise to the CNS midline cells. The three lateral CNS domains are defined by expression of three homeobox-containing proteins: *ventral nervous system defe*ctive (*vnd*) (ventral column), *intermediate neuroblasts defective* (*ind*) (intermediate column), and *muscle segment homeobox* (*msh*) (dorsal column) (D'Alessio and Frasch, 1996; Chu *et al.*, 1998; McDonald *et al.*, 1998; Weiss *et al.*, 1998). The initial expression of these three homeobox proteins is controlled by the Dorsal, Dpp, and Egfr signaling pathways (Von Ohlen and Doe, 2000). The ventral boundary of each homeobox gene is limited, in turn, by the adjacent homeobox protein, such that *ind* represses *msh* in the intermediate column and *vnd* represses *ind* in the ventral

[1] Authors contributed equally.

[2] To whom correspondence should be addressed at the Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7260. Fax: (919) 962-8472. E-mail: steve_crews@unc.edu.

0012-1606/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

column (McDonald *et al.*, 1998; Weiss *et al.*, 1998; Von Ohlen and Doe, 2000). *Sim* functions in a similar manner to limit the ventral expression domain of *vnd*. The relative expression domains of all of these proteins are conserved between *Drosophila* and vertebrates (Crews and Fan, 1999; Weiss *et al.*, 1998). The experiments described here were designed to define the transcriptional mechanisms employed by *sim* to distinguish the midline cells from the adjacent ventral column of lateral CNS cells.

That *sim* is required for midline cell development has been shown through the analysis of *sim* mutants. CNS midline cells fail to develop, and transcription of genes normally expressed in the midline cells is abolished in *sim* mutants (Thomas *et al.*, 1988; Nambu *et al.*, 1990). Ectopic activation of *sim* in the lateral neuroectoderm results in transformation of the entire CNS into midline cells (Nambu *et al.*, 1991). To activate transcription, the Sim bHLH-PAS protein forms a heterodimer with the Tango (Tgo) bHLH-PAS protein (Sonnenfeld *et al.*, 1997) and binds to CNS midline elements (CMEs; ACGTG) residing within target genes. Alteration of target gene CMEs by site-directed mutagenesis coupled with transgene-*lacZ* assays has shown that the CMEs are required *in vivo* for midline gene expression (Wharton *et al.*, 1994; Sonnenfeld *et al.*, 1997; Ohshiro and Saigo, 1997; Zelzer *et al.*, 1997). Sim, like other bHLH-PAS proteins, can be divided into discrete functional domains. Sim has a DNA binding basic region and a HLH domain, which, together with the PAS domain, are required for dimerization with Tgo (Moffet *et al.*, 1997; Probst *et al.*, 1997; Nystrom and Crews, unpublished). The PAS domain also serves to interact with other coregulatory proteins (Zelzer *et al.*, 1997; Ma *et al.*, 2000). Finally, the C-terminus contains multiple regions that act as transcriptional activation domains (Franks and Crews, 1994).

Genetic and ectopic expression experiments have shown that *sim* function is required for repression of lateral CNS gene expression in the midline cells. The *hedgehog* (*hh*), *tartan* (*trn*), *vnd*, and *wingless* (*wg*) genes are all expressed in cells of the lateral CNS, but expression is abolished in the midline cells after the appearance of Sim protein (Chang *et al.*, 1993; Mellerick and Nirenberg, 1995; Xiao *et al.*, 1996). Embryos mutant for *sim* ectopically express all four genes in midline cells, indicating that *sim* is required for their repression in the midline. Further confirmation is provided by experiments in which *sim* is ectopically expressed in the lateral neuroectoderm and eliminates lateral expression of these genes (Chang *et al.*, 1993; Xiao *et al.*, 1996; J. Nambu, personal communication).

The mammalian *Sim* genes show functional similarities to *Drosophila sim* (Epstein *et al.*, 2000), and therefore, insight from the *Drosophila* studies may be generally relevant. Both mammalian homologs of *sim*, Sim1 and Sim2, can dimerize with Tgo orthologs, Arnt or Arnt2, and bind the CME (Moffett and Pelletier, 2000). The *Sim1* and *Sim2* genes are expressed in the developing nervous system (reviewed in Crews and Fan, 1999). *Sim1* appears to activate transcription, as it is required for the expression of several neuropeptide hormones in the hypothalamus and for some regional expression of the *brain-2* gene (Michaud *et al.*, 1998). *Sim2* has been implicated in Down syndrome in humans (Chen *et al.*, 1995; Dahmane *et al.*, 1995; Muenke *et al.*, 1995) and learning behavior in mice (Ema *et al.*, 1999). *Sim2* precedes the expression of *Sonic hedgehog* (*Shh*) in the diencephalon and may regulate its expression. Ectopic expression of *Sim2*, as well as *Sim1* and *Drosophila sim*, causes additional *Shh* expression in this tissue (Epstein *et al.*, 2000). In cell culture experiments, *Sim2* has been shown to repress transcription, although no known repressive targets have been identified (Ema *et al.*, 1996; Moffett *et al.*, 2000; Moffett and Pelletier, 1997).

The question of how *Drosophila sim* controls midline repression deals with the important issue of how master regulatory proteins control cell fate in negative as well as positive modes. To define the mode of action of *sim* in mediating repression, we tested three general models: (1) Sim could directly bind to DNA of target genes and repress in association with corepressors. In this model, Sim:Tgo heterodimers bind CMEs on target genes and interact with corepressors having adjacent genomic binding sites. This model predicts that Sim binds target DNA directly via CMEs. (2) Another model proposes that Sim does not directly bind target gene DNA, but interacts with positively acting regulatory proteins to prevent their function. This mode of repression does not require DNA binding, but would be mediated by direct interactions between Sim and the activator protein. (3) The final model is that Sim represses indirectly by activating transcription of genes encoding repressive factors. DNA binding and transcriptional activation by Sim would be required in this model. The repressed target genes would not need sites for Sim:Tgo binding, but would instead contain binding sites for repressors that are activated by Sim.

To distinguish between these models, we examined *sim*-mediated midline repression of the *vnd* gene (Kim and Nirenberg, 1989; Jimenez *et al.*, 1995), a representative Sim target gene (Mellerick and Nirenberg, 1995). Although differences may occur among genes repressed in the midline, the *vnd* gene is an attractive target since it plays an important role in development of the lateral CNS cells directly adjacent to the midline. The *vnd* gene has multiple CMEs upstream of the transcriptional start site. We have used a transgenic *vnd-lacZ* approach to test whether these Sim:Tgo binding sites are required for midline repression and to identify critical regions required for midline repression. Using an alternative approach, modified forms of Sim were assayed for their ability to repress and activate transcription when ectopically expressed. This analysis can test which regions of Sim are required for repression and test models of Sim function. The results demonstrate that Sim represses by activating transcription of target genes encoding repressive factors. The CMEs present within the *vnd* gene are not required for midline repression, rather independent sequences (lacking any CMEs) are needed. Mutational analysis of Sim demonstrates that the basic region is

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

required for repression as well as activation, indicating that repression requires Sim DNA binding. The Sim C-terminal activation domain is also required for repression. The C-terminal region of Sim can be replaced with a heterologous VP16 activation domain and still activate and repress transcription. These results indicate that Sim mediates midline repression through the activation of gene expression.

MATERIALS AND METHODS

Drosophila Strains

The yw^{67} and w^{1118} strains were used for injections of transgenic constructs and as control strains. Enhancer-Gal4 lines used to misexpress UAS-*sim* transgenes included *engrailed*-Gal4 (*en*-Gal4) and *paired*-Gal4 (*prd*-Gal4). The 4×CME-*lacZ* flies contain a P[w^+; 4×CME-*lacZ*] transgene that has four copies of the *Toll* site 4 CME cloned into the C4PLZ enhancer tester vector (Wharton *et al.*, 1994).

Isolation of the vnd Genomic Upstream Regulatory Region

Genomic clones containing the *vnd* gene were isolated by screening a λEMBL3 genomic library (provided by Ron Blackman) with a 1090-bp *Sac*II fragment derived from a *vnd* cDNA clone (obtained from Kalpana White). One of the genomic clones contained an 11.4-kb *Hin*dIII fragment consisting of 7.7 kb of upstream flanking DNA, exon 1, and most of intron 1. This fragment was cloned into the *Hin*dIII site of pBKS (pBKS-H3) and used to generate the *vnd* transgenes.

DNA Sequencing and Analysis

DNA sequencing was carried out on a 7.7-kb *Hin*dIII–*Sac*II *vnd* fragment containing upstream flanking DNA. Primers corresponding to the sequence were used in automated ABI DNA sequencing carried out by the UNC Sequencing Facility. Computer analysis of predicted binding sites for the *Drosophila* Dorsal, Snail (Sna), Twist (Twi), and Vnd transcription factors was carried out using MatInspector v2.2 (Quandt *et al.*, 1995) and GCG FindPatterns (Womble, 2000). Dorsal binding sites were identified by: (1) MatInspector analysis using the Dorsal high-affinity consensus sequence, GGGTTTTTCC (Pan and Courey, 1992), with a score of 0.75 and matrix of 0.85, and (2) FindPatterns of Dorsal consensus site, GGGRHTYYCC, with a mismatch of 1 (Locker, 1996). Sna sites were identified using the Sna consensus sequence of CACCTGTT (Kasai *et al.*, 1992) with: (1) MatInspector with score of 0.75 and matrix of 0.85 and (2) FindPatterns with a mismatch of 1. Twi sites were noted by using FindPatterns with a mismatch of 0 for sequences identical to known Twi binding sites in the *rhomboid*, *sna*, and *sim* genes (Ip *et al.*, 1992a,b; Kasai *et al.*, 1998). These sequences include CAATTG, CACGTG, CATATG, CATGTG, and CATTTG. Sites for Vnd binding were determined using: (1) MatInspector searches with the closely related mammalian Nkx2.5 protein consensus binding site sequence, TYAAGTG (Chen and Schwartz, 1995), and (2) FindPatterns using the Vnd consensus site CAAGTG with a mismatch of 1 (Gruschus *et al.*, 1997). Sites for Enhancer of split (E(spl)) binding were determined with FindPatterns using the E(spl) consensus sites CACNAG and CACGYG (Nelleson *et al.*, 1999).

Generation of vnd-lacZ Transgenic Constructs

Fragments of the *vnd* gene were cloned upstream of a P-element promoter fused to the *lacZ* structural gene in the C4PLZ enhancer tester vector (Wharton and Crews, 1993), except for 4.3AA, which was inserted into the CaSper-AUG-βgal vector (Thummel *et al.*, 1988). Numbering of the *vnd* sequence is based on that used by Saunders *et al.* (1998) with +1 the presumed start site of transcription. Construction of the different *vnd* transgenes is described below and depicted in Fig. 2.

8.1HV. The *vnd* sequences comprising 8.1HV extend from the −7.7 *Hin*dIII site to the +0.3 *Eco*RV site. The adjacent *vnd* 6.9-kb *Hin*dIII–*Eco*RV and 1.2-kb *Eco*RV fragments were simultaneously ligated into *Hin*dIII/*Eco*RV-digested pBKS. This plasmid was digested with *Not*I and *Kpn*I, liberating the re-created 8.1-kb *vnd* fragment, which was cloned into *Not*I/*Kpn*I-digested C4PLZ.

5.3RS. The 5.3RS *vnd* sequences extend from the −5.3 *Eco*RI site to the +56 *Sac*II site. The *Eco*RI–*Sac*II fragment was subcloned into *Eco*RI/*Sac*II-cut pBKS. The *vnd* insert was removed by *Eco*RI/*Spe*I digestion and inserted into the *Eco*RI–*Spe*I site of C4PLZ.

4.3AA. The 4.3-kb *Afl*II *vnd* fragment (−4.2 to +0.1) was initially ligated into pPCR-Script (Stratagene). This plasmid was digested with *Not*I, cutting at a site located in the pPCR-Script vector at the 3′ end of the *vnd* sequences. The *Not*I site was made blunt using Klenow fragment and the plasmid was digested with *Kpn*I to liberate the 5′ end of the insert. The resulting fragment was directionally inserted into CaSpeR-AUG-βgal that had been digested with *Eco*RI. The *Eco*RI sites were blunted and then digested with *Kpn*I. Unlike the other reporter constructs, this reporter utilized the *vnd* promoter, rather than the P-element transposase promoter provided by the C4PLZ enhancer tester vector.

2.5RB. The *vnd* fragment extending from the −5.3 *Eco*RI site to the −2.8 *Bam*HI site was ligated into *Eco*RI/*Bam*HI-digested C4PLZ.

1.8RN. The *vnd* fragment extending from the −5.3 *Eco*RI site to the −3.5 *Nsi*I site was first inserted into the *Eco*RI–*Nsi*I sites of pBKS. The 1.8-kb *vnd* fragment was removed by *Not*I–*Kpn*I digestion and ligated into *Not*I/*Kpn*I-digested C4PLZ.

1.1RX. The *vnd* fragment extending from the −5.3 *Eco*RI site to the −4.2 *Xho*I site was inserted into the *Eco*RI–*Xho*I sites of pBKS. This fragment was removed by *Not*I–*Kpn*I digestion and ligated into *Not*I/*Kpn*I-digested C4PLZ.

1.1XX. The 1.1-kb *Xho*I *vnd* fragment (−4.2 to −3.1) was ligated into pBKS. This *vnd* fragment was removed by *Not*I–*Kpn*I digestion and ligated into *Not*I/*Kpn*I-digested C4PLZ.

0.8NB. The *vnd* 0.8NB fragment extends from the −3.5 *Nsi*I site to the −2.8 *Bam*HI site. It was inserted into the *Pst*I–*Bam*HI sites of pBKS. The 0.8-kb *vnd* fragment was removed by *Not*I–*Kpn*I digestion and ligated into *Not*I/*Kpn*I-digested C4PLZ.

1.4RBΔXX. This construct deletes the *Xho*I (−4.2 to −3.1) region from 2.5RB. Thus, it includes sequences from the −5.3 *Eco*RI site to the −4.2 *Xho*I site and from the −3.1 *Xho*I site to the −2.8 *Bam*HI site. These two fragments were simultaneously cloned into the *Eco*RI–*Bam*HI sites of pBKS. The 1.4-kb *vnd* fragment was removed by *Eco*RI–*Bam*HI digestion and ligated into *Eco*RI/*Bam*HI-digested C4PLZ.

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

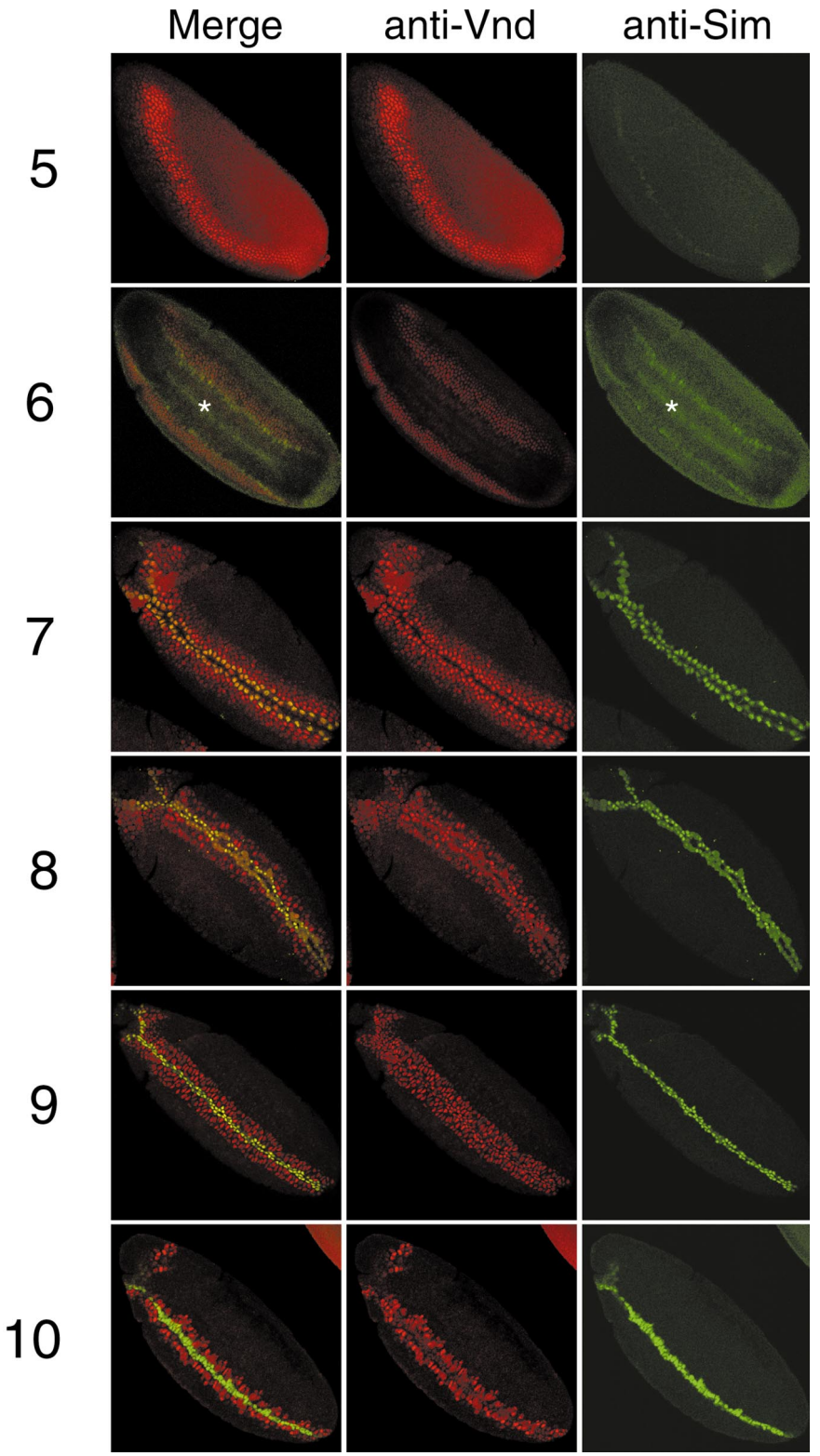

FIG. 1. Localization of Sim and Vnd proteins during early embryogenesis. Whole-mount embryos were double-stained with anti-Vnd (red) and anti-Sim (green) and analyzed by confocal microscopy. The Merge images show both anti-Vnd and anti-Sim staining together, so that the overlap in protein localization is yellow. Ventral views of stages 6–10 and a sagittal view of stage 5 are shown; anterior is to the left.

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

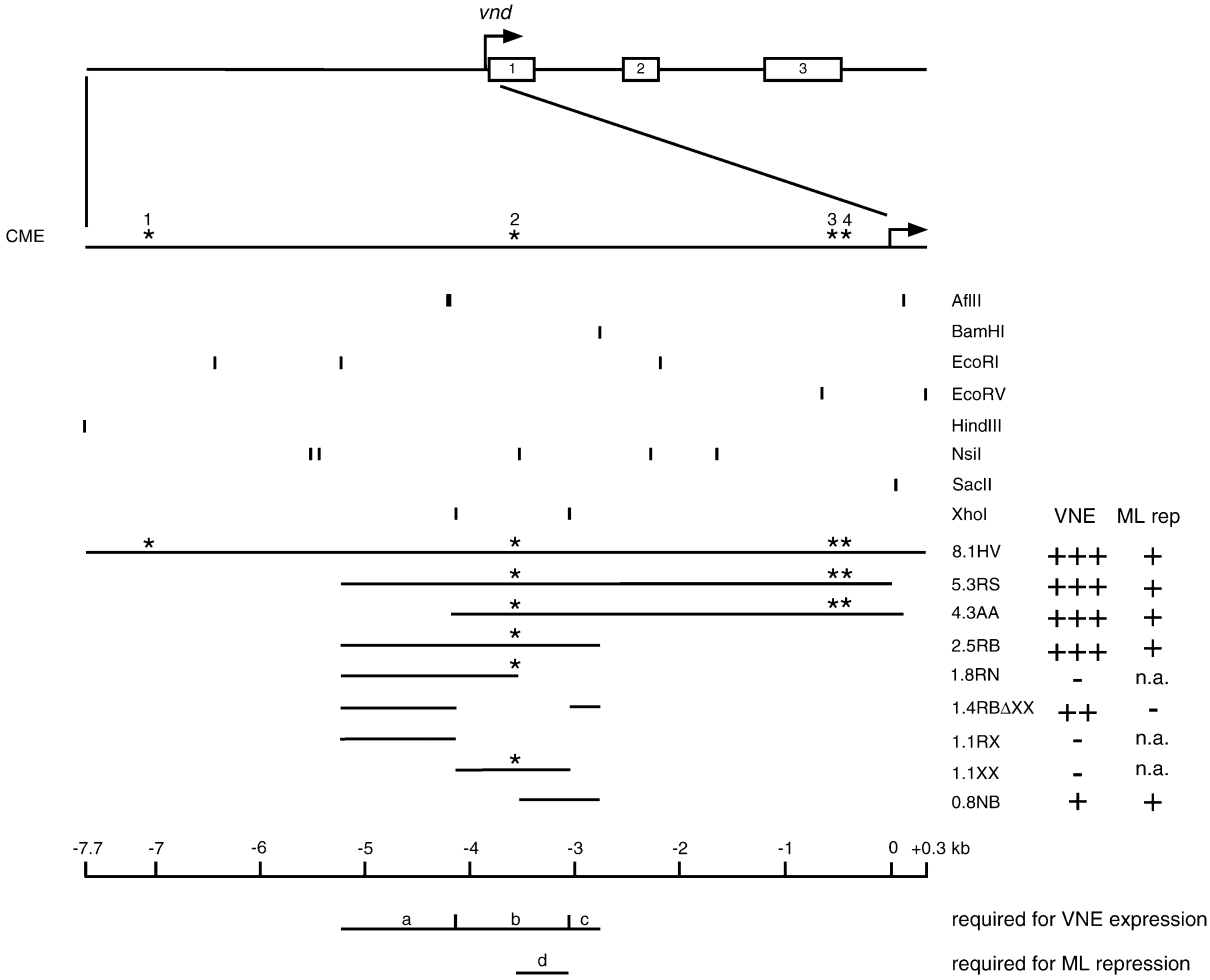

FIG. 2. Regions of the *vnd* gene that control neuroectodermal gene expression and midline repression. The structure of the *vnd* gene is shown at top with 7.7 kb of 5′ flanking sequence shown followed by the three exons (open boxes numbered 1–3) and two introns that constitute the transcription unit. The arrow indicates the location of the transcription start site and direction of transcription. Below are shown an enlargement of the 7.7-kb regulatory region (scale in kb is at bottom) and the locations of the four CME Sim:Tgo binding sites, which are indicated by (*). CMEs are numbered 1–4. The restriction map of the regulatory region is indicated. The fragments used to create the *vnd-lacZ* transgenic constructs are shown below the map. Each fragment was cloned into C4PLZ, except 4.3AA, which was cloned into CaSpeR-AUG-*β*gal. The name of each construct is a combination of its size in kb and the restriction sites that generated the fragment. Expression in the ventral neuroectoderm (VNE) and occurrence of CNS midline cell repression (ML rep) at stage 10 and later are indicated to the right of each construct. (+) in the VNE column indicates expression of *lacZ* resembling endogenous *vnd* expression, and (+) in the ML rep column indicates the presence of midline repression; (−) indicates no expression or an absence of midline repression; (n.a.) indicates not applicable. Below the scale is a summary of the results from this paper indicating the existence of three discrete regions required for lateral CNS transcription (a–c) and a region (d) required for midline repression.

Numbers to the left of the images correspond to embryonic stages as described by Campos-Ortega and Hartenstein (1997). Vnd protein is observed in the mesectoderm and ventral neuroectoderm in the stage 5 blastoderm embryo before the appearance of Sim protein. Vnd remains in cells of the ventral neuroectoderm through stage 10 and later (Mellerick and Nirenberg, 1995). As gastrulation takes place (stage 6), Sim protein is also detected in the mesectoderm. Asterisks indicate the location of the ventral furrow; enhanced green staining at the ventral furrow is background and not Sim. Both Sim and Vnd proteins continue to appear in the midline cells during stages 7–9 (note yellow midline staining in the Merge images). Late stage 9, stage 10, and older embryos are devoid of Vnd midline staining, while Sim protein remains in the midline cells.

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

Site-Directed Mutagenesis of vnd CME Sequences

Sim:Tgo binding sites were mutated by site-directed mutagenesis using an Altered Sites II kit (Promega). DNA fragments to be mutated were initially cloned into pALT. These fragments include a 3.1-kb *Eco*RI fragment containing CME-2, a 1.2-kb *Nsi*I–*Pst*I fragment containing CME-3, and a 0.8-kb *Pst*I–RV fragment containing CME-4. Each mutation was generated individually, the mutated fragments were reassembled into pBKS and subsequently subcloned into C4PLZ as a *Not*I–*Kpn*I fragment to generate 5.3RSΔCME. Generation of 2.5RBΔCME required subcloning the resulting *Eco*RI–*Bam*HI fragment containing the mutated CME-2 into the corresponding site of C4PLZ. In each case, two or three nucleotides of the ACGTG core CME sequence were altered, and a new restriction site was created to assist in the identification of mutant fragments. Primers used for mutagenesis are indicated below. The mutated residues are underlined below the location of the wild-type ACGTG or CACGT complementary sequences. The newly created restriction sites are indicated to the right of the sequence.

	ACGTG	
CME-2	5′-CAGTGTTTAAAAATAGATCTTTTTTTATTTTTCG-3′	*Bgl*II
	CACGT	
CME-3	5′-AACCTCCTCAGGCTCGAGGGATTGACGTTC-3′	*Xho*I
CME-4	5′-TCCTTATGCCGCCCGGGATCCCGCTCTAA-3′	*Sma*I

Generation of UAS-sim Mutant Lines

Mutated *sim* constructs were subcloned into pUAStag (Mosher and Crews, unpublished), a derivative of the pUAST vector that places the insert under control of the Gal4-UAS promoter (Brand and Perrimon, 1993). In addition, pUAStag provides a green fluorescent protein (GFP) tag fused in frame to the 3′ end of introduced coding sequences. Creation of pUAStag involved PCR amplification of GFP cDNA from the phGFP-S65T vector (Clontech) to create a *Bgl*II–GFP–*Bam*HI fragment that was subcloned into the *Bgl*II site of pUAST (Brand and Perrimon, 1993). Full-length *sim* and three *sim* deletion constructs were subcloned into the *Bgl*II site of pUAStag in frame with the 3′ GFP sequences. The full-length Sim construct (FL-Sim) was created by subcloning a 2-kb *Bgl*II fragment excised from the *sim* cDNA clone NB-F1(Bgl) (Franks and Crews, 1994). Δb-Sim deletes the basic region of Sim by removing the first 11 amino acids (aa) of Sim. ΔPAS-Sim removes the PAS domain by deleting residues 88–356. These two mutant forms of Sim were previously made (Franks and Crews, 1994) and moved into pUAStag. ΔAAQ-Sim was made according to the PCR–ligation–PCR mutagenesis technique of Ali and Steinkasserer (1995). Two fragments of *sim* were PCR-amplified, the first extending from the start of the coding sequence up to the AAQ repeat and the second extending 3′ of the AAQ repeat to the 3′ end of the gene. These two fragments were then ligated in frame, creating a *sim* fragment without the AAQ repeat (aa 377–428). The ΔC-Sim construct removes the final 211 amino acids from the C-terminus of Sim. This construct was generated by cleaving *sim* with *Bgl*II and *Bam*HI and cloning it into pUAST without a GFP tag.

Generation of a UAS-sim Heterologous Activation Line

A heterologous activation construct was created by fusing the VP16 activation domain (Triezenberg *et al.*, 1988) from the herpes simplex virus (HSV) onto ΔC-Sim. The UAS-ΔC-Sim-VP16 construct was made by digesting pSJT 1193 CRF3 (provided by Steven Triezenberg) with *Bam*HI and *Bgl*II and subcloning a 360-bp fragment containing the VP16 activation domain into pUAST. The ΔC-Sim fragment was subcloned in frame 5′ to the VP16 activation domain.

Drosophila S2 Cell Transient Transfection Assays

Plasmids for expressing *sim* in cell culture were created by subcloning mutated *sim* cDNA fragments into the pAct5C expression vector (Han *et al.*, 1989; Sonnenfeld *et al.*, 1997). Full-length *sim* and *tgo* expression constructs were previously described (Sonnenfeld *et al.*, 1997). The heterologous activation domain–Sim fusion construct was made by first subcloning the VP16 activation domain into pAct5C, followed by the addition of ΔC-Sim. The reporter plasmid was CME-*luc* (Emmons *et al.*, 1999), which has six copies of the *Toll* site 4 CME fused to the firefly *luciferase* gene in the pGL3 Promoter vector (Promega). Transfections were normalized with a *copia*-LTR-*luc* plasmid that constitutively expresses *Renilla* luciferase. Transfections, dual luciferase assays, and the use of the *copia*-LTR-*luc* transfection normalization plasmid were previously described (Emmons *et al.*, 1999). Each transfection was carried out in triplicate.

Generation of P-element Germ-Line Transformants

Germ-line transformants were generated by microinjection of both transgenic DNA constructs and P-turbo (pUChspΔ2-3wc) transposase helper plasmid using standard procedures (Rubin and Spradling, 1982).

Immunostaining of Embryos

Antibody staining of embryos was carried out according to standard protocols (Patel *et al.*, 1987). Primary antibodies used for staining were: (1) 10×-concentrated monoclonal antibody (mAb) anti-Tgo (Ward *et al.*, 1998), (2) rat anti-Sim (1:200) (Ward *et al.*, 1998), (3) rabbit anti-Vnd (1:200) (McDonald *et al.*, 1998), (4) mAb anti-β-galactosidase (βgal) (1:200; Promega), and (5) rabbit anti-βgal (1:200; Cappel). For the visualization of immune complexes, the secondary antibodies anti-rabbit Texas red, anti-rat Alexa 488 (green), and anti-mouse 594 (red) (all 1:200; Molecular Probes) were used. Embryos were mounted in Aquapolymount (Polysciences, Inc.) and viewed using a Zeiss LSM 410 laser scanning confocal microscope with an argon–krypton laser.

RESULTS

Spatial and Temporal Overlap of Vnd and Sim Gene Expression

The relationship between Sim and Vnd in the CNS midline cells was examined by immunostaining embryos with both anti-Sim and anti-Vnd (Fig. 1). Vnd protein is first seen at embryonic stage 5 in the presumptive mesectoderm

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

and ventral neuroectoderm, preceding the appearance of Sim protein (McDonald *et al.*, 1998). Sim protein appears during gastrulation (stage 6) (Crews *et al.*, 1988) in the mesectoderm and overlaps with Vnd protein. During stages 6–9 both Sim and Vnd are colocalized in the CNS midline cells, while Vnd protein continues to be present in cells of the ventral neuroectoderm. By the end of stage 9 and during stage 10, Vnd protein is absent in the CNS midline cells (McDonald *et al.*, 1998), while Sim protein remains. The absence of Vnd protein is preceded by the reduction of *vnd* RNA at stage 8 (Mellerick and Nirenberg, 1995). These results show that Sim and Vnd proteins overlap within the mesectoderm during several stages and that a considerable lag exists between the appearance of Sim protein and the loss of Vnd protein (~2 h). There is also a substantial lag between the appearance of Sim protein and the loss of midline *vnd* RNA (~1 h). The delay in *vnd* repression after initial Sim appearance is consistent with an indirect mechanism of repression.

Identification of Genomic vnd Sequences Required for Ventral Neuroectodermal Expression and Midline Repression

One model of direct repression of *vnd* by Sim requires that Sim:Tgo heterodimers bind to sites within the *vnd* gene. The only known Sim:Tgo binding sites contain an ACGTG core sequence, the CME. Sequence analysis of 7.7 kb of DNA upstream of the *vnd* start site revealed four Sim:Tgo binding sites (−7121, −3622, −567, −400; Fig. 2). Deletions and mutations of the *vnd* regulatory region were tested by *lacZ* germ-line transformation for proper ventral neuroectodermal expression, CNS midline repression, and the ability of the CMEs to mediate midline repression.

Previous germ-line transformation experiments indicated that an 8.7-kb *vnd* fragment was able to drive *lacZ* expression in a neuroectodermal pattern identical to *vnd* expression, and a 5.6-kb fragment showed *vnd*-like CNS expression but lacked some cephalic expression (Saunders *et al.*, 1998). Similar to these results, a fragment containing 8.1 kb of *vnd* genomic DNA (8.1HV) showed a *lacZ* neuroectodermal pattern resembling authentic *vnd* gene expression, including midline repression (Fig. 3). The 5′ deletion fragments, 5.3RS and 4.3AA, also showed *vnd*-like neuroectodermal expression patterns and midline repression as well as some expression in the dorsal ectoderm. Since CME-1 is not present in these constructs, this indicates that it is not required for midline repression. The 2.5RB construct (−5.3 to −2.8 kb) has the same 5′ end as 5.3RS, but removes 2.8 kb of 3′ DNA. This fragment also had *lacZ* expression resembling *vnd* (Fig. 3). These results indicate that a 2.5-kb upstream fragment of *vnd* can drive normal ventral neuroectodermal gene expression and midline repression. Since 2.5RB contains only CME-2, the other three CMEs are not required for midline repression.

Identification of Sequences Required for vnd Midline Repression

Two constructs that remove 0.7 (1.8RN) and 1.4 kb (1.1RX) from the 3′ end of 2.5RB failed to express *lacZ* in the neuroectoderm (data not shown). This indicates that sequences between −3.6 and −2.8 are required for *vnd* neuroectodermal expression. Another deletion construct (1.1XX) also failed to express *lacZ* (data not shown), but an overlapping construct (0.8NB) showed weak *lacZ* neuroectodermal expression accompanied by midline repression (Fig. 3). These results demonstrated that a region between −3.1 and −2.8 (Fig. 2, "c") is required for neuroectodermal expression. This was confirmed by testing an internal deletion derivative of 2.5RB (1.4RBΔXX). This strain had intermediate levels of *vnd* neuroectodermal expression (Fig. 3). Since 1.4RBΔXX showed neuroectodermal *lacZ* expression and 1.1RX did not, this confirmed that the region from −3.1 to −2.8 (Fig. 2, "c") has sequences necessary for neuroectodermal transcription. However, since 0.8NB is weakly expressed compared to 1.4RBΔXX, there are likely to be additional sequences from −5.3 to −4.2 (Fig. 2, "a") that are required for normal expression levels. Similarly, since 1.4RBΔXX expression is less than that of 2.5RB, sequences between −4.2 and −3.1 (Fig. 2, "b") are necessary for high levels of expression. In summary, robust expression of *vnd* in the neuroectoderm requires at least three separate regions (Fig. 2, "a–c"). Since 0.8NB shows midline repression, yet does not possess CMEs, this demonstrates that CMEs do not contribute to midline repression.

The 1.4RBΔXX construct was expressed in midline cells in stage 11 embryos (Fig. 3), while the other *lacZ*-expressing constructs did not show midline *lacZ* staining at this stage. The 1.4RBΔXX midline staining diminishes after stage 12. Together, the results from the 1.4RBΔXX and 0.8NB constructs indicate that an element required for initial midline repression lies between −3.6 and −3.1 (Fig. 2, "d").

vnd CMEs Are Not Required for Midline Repression

The ability of 0.8NB, which does not possess a CME, to repress *lacZ* in the midline suggested that *vnd* CMEs were not required for midline repression. To directly test the importance of the CMEs, the CMEs in 5.3RS and 2.5RB, which express *lacZ* at high levels, were mutated and transgenic strains assayed for midline repression. The 5.3RS fragment has three CMEs (Fig. 2), and all three were mutated. Five independent lines of the mutated 5.3RS transgene (5.3RSΔCME) were analyzed. Four of the five lines (5.3RSΔCME-4, -10, -57, and -65) showed a *lacZ* pattern indistinguishable from that of the unmutated 5.3RS lines (Fig. 4). The fifth transformant (5.3RSΔCME-1) showed ventral neuroectodermal expression like the others, but also showed strong midline expression throughout embryogenesis (Fig. 4).

Since four of five lines of 5.3RSΔCME were repressed in the midline, the simplest interpretation of the results is

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

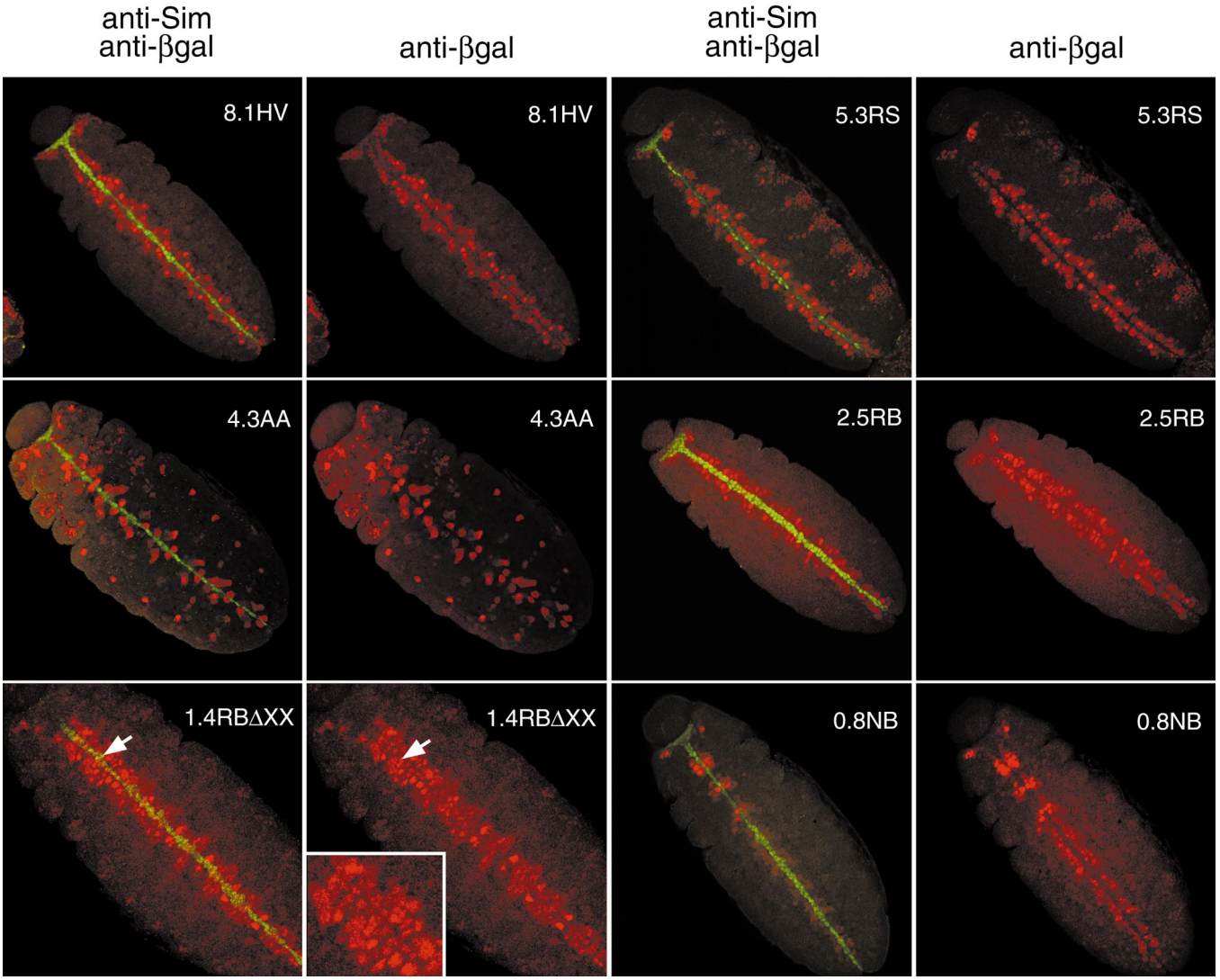

FIG. 3. Expression analysis of *vnd-lacZ* transgenic lines defines *vnd* regulatory regions. Whole-mount embryos derived from *vnd-lacZ* lines containing the reporter genes depicted in Fig. 2 were double-stained with anti-β-galactosidase (anti-βgal; red) and anti-Sim (green) and examined by confocal microscopy. Transgene names are shown. Shown are ventral views of stage 11 embryos; anterior is to the left. Images were taken at focal planes revealing the CNS midline and lateral CNS cells and do not show additional Vnd-positive cells in the adjacent ectoderm. The 8.1HV, 5.3RS, 4.3AA, and 2.5RB transgenic strains all show strong neuroectodermal expression and midline repression. The 1.4RBΔXX transgene is expressed at reduced levels, but also shows expression in CNS midline cells. The midline staining is variable, but reproducible in three independent lines. Highest midline staining is in anterior cells (arrow). Inset shows higher magnification. The 0.8NB line shows neuroectodermal expression and midline repression, although levels of expression are low.

that the 5.3RS CMEs are not required for midline repression, and the exceptional 5.3RSΔCME-1 transgene has inserted into a gene with midline expression. Thus, it acts as an enhancer trap, and the midline staining does not reflect the properties of the mutated *vnd* gene sequences.

The 2.5RB fragment also shows midline repression and has only a single CME (Fig. 2). The CME of 2.5RB was mutated (2.5RBΔCME), and all three transgenic lines showed midline repression (Fig. 4). Thus, site-specific mutation of CMEs in 5.3RS and 2.5RB, along with the results of deletion analysis, demonstrate that *vnd* CMEs are not required for ML repression, but a repressive region located between −3.6 and −3.1 (Fig. 2, "d") is necessary.

In order to test if the midline repression of the reporter transgenes is dependent upon the expression of *sim*, we examined both the 5.3RS and the 2.5RB reporters in a *sim* mutant background. In both cases, the midline repression of the *vnd-lacZ* sequences was abolished (Fig. 4B). This accu-

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

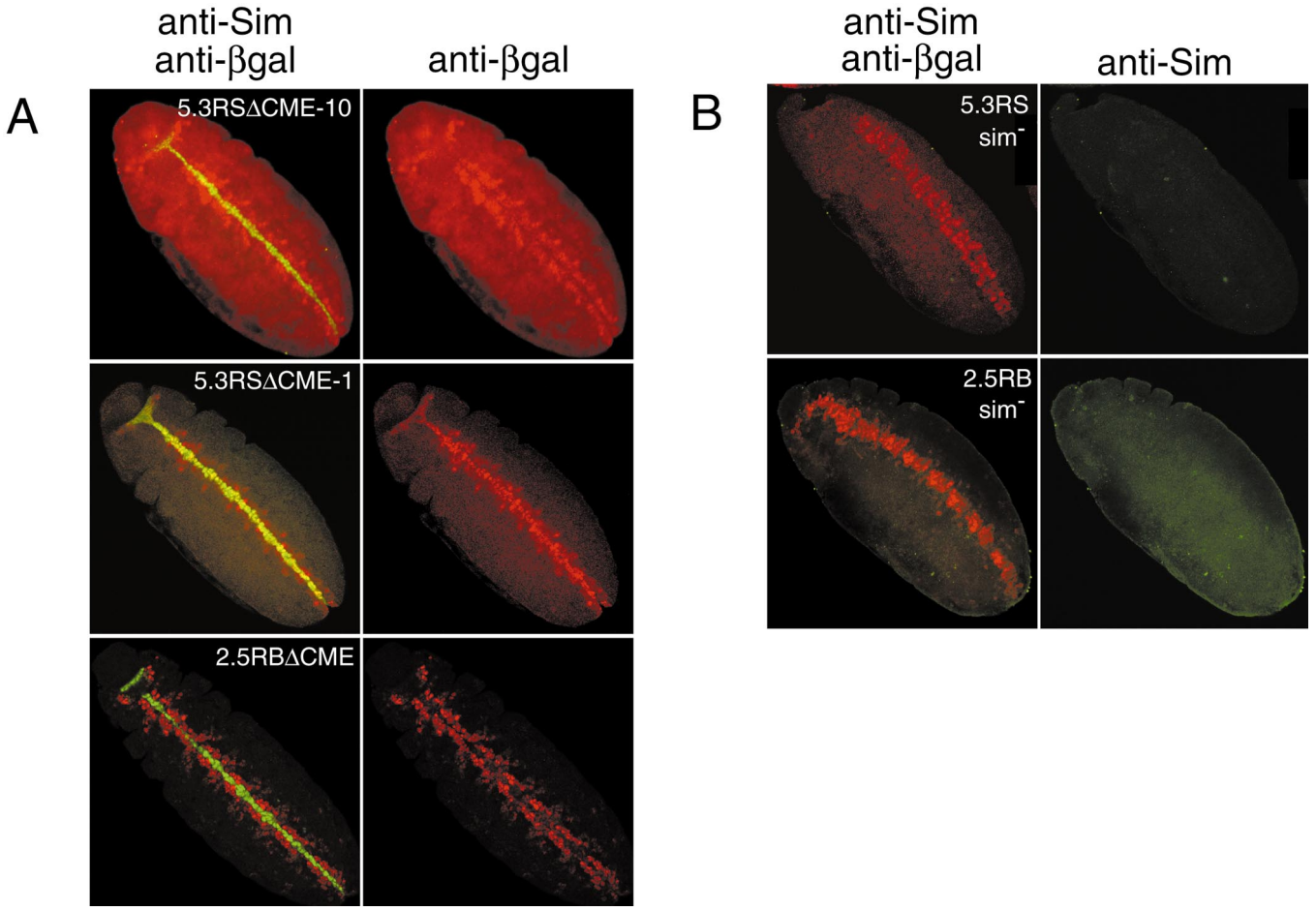

FIG. 4. (A) Elimination of CMEs from the *vnd* regulatory region does not abolish midline repression. The three CMEs (2–4) that reside within the *vnd* 5.3RS fragment were mutated (5.3RSΔCME) and tested *in vivo* for their ability to mediate midline repression. Five independent lines (1, 4, 10, 57, 65) of 5.3RSΔCME were double-stained with anti-βgal (red) and anti-Sim (green). Four lines, 5.3RSΔCME-4, -10, -57, and -65, showed midline *lacZ* repression (5.3RSΔCME-10 is shown) and one line, 5.3RSΔCME-1 (also shown), had high levels of midline *lacZ* expression. Two images are shown for each strain: a merged image and an anti-βgal image. Ventral views of stage 11 embryos are shown. The single CME-2 site in 2.5RB was mutated and the resulting transgene, 2.5RBΔCME, showed midline *lacZ* repression. Three lines of 2.5RBΔCME stained identically to the one shown. (B) Midline repression is lost in *sim* mutants. The 5.3RB and 2.5RB reporter genes were tested in a *sim* mutant background. Embryos were double-stained with anti-βgal (red) and anti-Sim (green). Two images are shown for each tested construct: a merged image and an anti-Sim image. Ventral views of stage 11 embryos are shown.

rately reflects the pattern of the endogenous *vnd* gene in a *sim* background (Mellerick and Nirenberg, 1995) and indicates that midline repression requires *sim* function.

Analysis of UAS-Sim Deletion Mutants Correlates Repression with Activation

The identification of regions of Sim necessary for midline repression required the creation of Sim deletion mutants and an assay to test their function. Previous work established that repression of lateral CNS gene expression could be studied by ectopically expressing *sim* in the lateral CNS. Repression of *wg* and *hh* in the neuroectoderm was observed when *sim* was expressed in ectodermal stripes in *prd*-Gal4 x UAS-*sim* embryos (Xiao *et al.*, 1996). We have employed similar assays to examine the ability of Sim deletion mutants to repress *vnd*. The following deletions were made and injected into embryos to create transgenic lines: (1) basic domain (Δb-Sim; aa 1–11 removed), (2) PAS domain (ΔPAS-Sim; aa 88–356 removed), (3) Ala-Ala-Gln repeat region (ΔAAQ-Sim; aa 377–426 removed), and (4) C-terminal region (ΔC-Sim; aa 463–673 removed) (Fig. 5). Full-length Sim (FL-Sim) and the deletion constructs were cloned into a UAS transformation vector that added GFP onto their C-termini, with the exception of ΔC-Sim, which was cloned into the UAS transformation vector without

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

GFP. Flies containing the UAS-Sim transgenes were crossed to either *en*-Gal4 or *prd*-Gal4 to drive transgene expression in ectodermal stripes. The embryos were assayed for repression of neuroectodermal *vnd* using anti-Vnd immunostaining. These mutant forms of Sim were also tested for the ability to activate transcription from a reporter containing multiple Sim:Tgo binding sites (CME-*lacZ*) and to dimerize with Tgo.

FL-Sim. UAS-FL-Sim × *en*-Gal4 embryos direct ectopic expression to 14 circumferential ectodermal stripes, beginning at stage 10 and continuing through late embryogenesis. At the intersection of these stripes and the lateral neuroectodermal expression of *vnd* (Fig. 6A), Vnd protein was absent or greatly reduced in the cells that overlap FL-Sim expression (Fig. 6A). The expression of FL-Sim also induced expansion of Vnd in neuroectodermal cells adjacent to *en* stripes, likely due to cell fate changes directed by Spitz signaling from the ectopic Sim-positive cells (Kim and Crews, 1993; Raz and Shilo, 1993; Gabay *et al.*, 1996; Golembo *et al.*, 1996). Vnd was also expanded in the head region (data not shown). FL-Sim also activated the CME-*lacZ* reporter and directed the nuclear accumulation of Tgo. UAS-FL-Sim × *en*-Gal4; CME-*lacZ* embryos were stained with anti-βgal (Fig. 6B) and anti-Tgo (Fig. 6C). Embryos stained with anti-βgal showed activation of the reporter in cells overlapping the FL-Sim in the *en* stripes, and FL-Sim also directed Tgo to the nucleus (Fig. 6C). Both results are consistent with previous observations (Ward *et al.*, 1998) and demonstrate that the 3′ GFP tag does not affect normal function.

Δ*b-Sim.* Previous work revealed that mutation of the Sim basic region abolished the ability of Sim to activate gene transcription (Franks and Crews, 1994). Consistent with these results, the Δb-Sim deletion failed to activate CME-*lacZ* transcription (Fig. 6B). Δb-Sim also failed to repress *vnd* expression (Fig. 6A). Tgo was localized to the nucleus (Fig. 6C), indicating that Δb-Sim forms a protein able to dimerize with Tgo and translocate to the nucleus. These results suggest that *vnd* repression requires DNA binding by Sim.

Δ*PAS-Sim.* ΔPAS-Sim failed to repress *vnd* expression (Fig. 6A) and it was unable to activate transcription of the CME-*lacZ* reporter gene (Fig. 6B). Although ΔPAS-Sim accumulated in nuclei, Tgo remained cytoplasmic (Fig. 6C). This result indicated that ΔPAS-Sim fails to dimerize with Tgo, supporting the importance of Sim:Tgo dimerization for repression. Since the Sim PAS domain can bind other proteins like Fish (Ma *et al.*, 2000), the possibility exists that the PAS domain may be binding other proteins necessary for repression. Additionally, the failure of ΔPAS-Sim to repress transcription may be due to a repressive region that is absent in the ΔPAS-Sim although subsequent experiments suggest this is unlikely.

Δ*AAQ-Sim.* Adjacent to the C-terminus of the PAS domain is a repeating stretch of 10 Ala-Ala-Gln repeats followed by several imperfect repeats. ΔAAQ-Sim repressed *vnd* expression as efficiently as FL-Sim (Fig. 6A). Expression from the CME-*lacZ* reporter in ΔAAQ-Sim was at high levels (Fig. 6B), and Tgo protein accumulated to high levels in nuclei of *en* stripes (Fig. 6C). Thus, the AAQ region is not required for repression, activation, dimerization with Tgo, or nuclear accumulation of Sim:Tgo complexes.

Δ*C-Sim.* Expression of ΔC-Sim in *en* stripes failed to repress *vnd* expression (Fig. 6A). ΔC-Sim also failed to activate *lacZ* expression from the CME-*lacZ* reporter (Fig. 6B). Tgo accumulated in the nuclei of stripes, indicating that ΔC-Sim was able to dimerize with Tgo, and the complex accumulated in cell nuclei (Figs. 6A and 6B).

The results described above involved crossing UAS-Sim mutant lines to *en*-Gal4; the same results were obtained when UAS-Sim lines were crossed to *prd*-Gal4, another line that expresses Gal4 in circumferential stripes beginning at stage 8 (data not shown). These experiments reveal that the basic region, the PAS domain, and the C-terminal regions are required for *vnd* repression and also for activation. The basic region is likely required for DNA binding and the PAS domain is required for Tgo dimerization. The C-terminal region is required for repression either because repression requires transcriptional activation or because there are separate repressive domains present in the C-terminus. This issue can be addressed by substituting heterologous activation domains for the C-terminus of Sim and assaying for repression.

Ectopic Expression of Sim Heterologous Activation Domain Chimeric Proteins Represses vnd Expression

The C-terminal region of Sim was replaced with the well-characterized activation domain from the HSV virion protein 16 (VP16) (Regier *et al.*, 1993). The ability of this chimeric protein to activate transcription was initially assessed using transient expression assays and *Drosophila* cell culture (Fig. 7). When cotransfected with Tgo, the ΔC-VP16-Sim activated expression from a multimerized CME-*luc* reporter, which contains six copies of the *Toll* 4 CME fused to *luc*. ΔC-VP16-Sim:Tgo activated CME-*luc* expression at levels 2.4-fold greater than FL-Sim:Tgo and 7.3-fold above ΔC-Sim:Tgo levels. Control experiments indicated that *luc* expression was dependent on the presence of Tgo and CME sequences. These experiments indicate that the VP16 activation domains can efficiently substitute for the Sim activation region in cultured cell assays.

Transgenic lines were created with the UAS ΔC-VP16-Sim heterologous activation construct and crossed to *en*-Gal4 to test for repression, activation of a multimerized CME-*lacZ* reporter, and Tgo nuclear accumulation. In addition, the ability to activate transcription from the 3.7 *sim-lacZ* reporter was tested. This transgene has a 3.7-kb fragment of the *sim* regulatory region fused to *lacZ* and represents a physiologically relevant Sim target gene since it contains five functional Sim:Tgo binding sites in their native context and is a target of *sim* autoregulation (Wharton *et al.*, 1994). ΔC-VP16-Sim was able to both repress *vnd*

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

(Fig. 8A) and activate CME-*lacZ* (Fig. 8B). Similar results were obtained when ΔC-VP16-Sim was expressed at an earlier time with the *prd*-Gal4 driver (data not shown). The *in vivo* results with the CME-*lacZ* reporter parallel the results from transient transfection assays. ΔC-VP16-Sim also activated expression from the 3.7*sim-lacZ* when crossed to *prd*-Gal4 (Fig. 8B), correlating repression with the activation of authentic target genes. One unusual result was that Tgo accumulated to high levels predominantly in the cytoplasm in the presence of ΔC-VP16-Sim (Fig. 8B). Thus, Tgo and ΔC-VP16-Sim can dimerize, but the proteins either inefficiently enter the nucleus or are actively exported. This may be due to the addition of the VP16 activation domain. UAS-Trh, which has a transcriptional activation domain capable of activating Sim target genes when fused to the PAS domain of Sim (Zelzer *et al.*, 1997), did not repress *vnd* when it was expressed in *en* stripes (data not shown). These experiments argue against a model in which the Sim C-terminal residues contain a repressive domain and instead demonstrate that *vnd* repression by Sim requires transcriptional activation.

DISCUSSION

Sim protein appears in mesectodermal cells during gastrulation and, soon after, activates midline gene transcription. Significantly later, lateral CNS gene expression in the midline cells is abolished in a *sim*-dependent manner. This paper addresses the mechanisms by which Sim mediates midline repression. The process by which *sim* regulates cell fate in an activation-dependent fashion may be a common mechanism for lineage-specific gene control.

Sim Is Required for the Transcriptional Activation of Midline Repressive Genes

We tested three general models of Sim-mediated repression: (1) Sim directly repressed target genes by binding their DNA and repressing transcription in association with a corepressor(s), (2) Sim does not bind DNA of target genes but interacts with positively acting factors preventing their action, and (3) Sim represses indirectly by activating transcription of genes encoding repressive factors. Several complementary experiments demonstrate that midline repression requires activation of repressive gene expression by Sim (Model 3).

Ectopic expression experiments utilizing mutant forms of Sim demonstrated that the basic region, PAS domain, and C-terminal regions are all required for both transcriptional activation and repression. Removal of the PAS domain also abolished the ability of Sim to form dimers with Tgo, suggesting that Tgo is necessary for repression. More informative is Δb-Sim. This mutant protein was able to dimerize with Tgo and the protein complex accumulated in the nucleus. However, neither midline transcription nor repression occurred, presumably due to the inability of the Sim:Tgo dimer to bind DNA. This argues against a model in which Sim interacts with an activator protein in a non-DNA-binding mode (Model 2) and instead suggests that DNA binding is required for Sim repression (Model 1 or Model 3). However, analysis of the *vnd* gene using *lacZ* transgenes indicated that Sim:Tgo binding sites were not required for midline repression (Model 1); mutation of the single CME in 2.5RB or mutation of three CMEs in 5.3RS did not affect *lacZ* expression. Transient transfection experiments have shown that CMEs are relevant targets of Sim:Tgo binding (Sonnenfeld *et al.*, 1997), and *in vivo* analyses of five different genes have shown that the CME functions *in vivo* as a Sim:Tgo binding site (Wharton *et al.*, 1994; Ohshiro and Saigo, 1997; Zelzer and Shilo, 2000). However, it remains possible that Sim:Tgo could bind a variant sequence within the *vnd* gene. Arguing against this are the results indicating that Sim represses indirectly by activating transcription.

The C-terminal region of Sim that follows the PAS domain contains multiple transcriptional activation domains (Franks and Crews, 1994). Removal of the C-terminal 211 aa eliminates those activation domains and additional residues. The ΔC-Sim protein was unable to activate midline transcription or repress *vnd* expression, even though it dimerized with Tgo and the complex accumulated in nuclei. This is consistent with Sim repressing *vnd* expression by activating the transcription of repressive factors. However, it is also possible that there is a domain within the C-terminal region that could directly mediate repression. Fusing the VP16 activation domain onto ΔC-Sim and functionally assaying the fusion protein *in vivo* tested this. The results showed that addition of the VP16 activation domain restored the ability of ΔC-Sim to activate transcription and repress *vnd*. These experiments demonstrated that *vnd* repression correlates with the ability of Sim to activate transcription (Model 3). Another construct removed the Sim AAQ repeat region. Its deletion did not affect the ability of Sim to dimerize with Tgo, accumulate in nuclei, activate transcription, or repress *vnd*. Although striking in sequence, its function remains a mystery. The combination of the *vnd-lacZ* and ectopic Sim-mutant experiments demonstrated that Sim does not directly repress or inhibit *vnd* gene expression but, instead, activates transcription of genes that encode repressive factors consistent with the third model of repression. This model is also consistent with the delayed timing of *vnd* repression seen in early embryonic development.

Repressive Element of vnd Resides in a 0.5-kb Region

A comparison between the 1.4RBΔXX and 0.8NB expression patterns revealed a 0.5-kb region necessary for repression that maps between −3.6 and −3.1 (Fig. 2, "d") in the *vnd* regulatory region. This repression of *vnd* seen with 1.4RBΔXX occurs variably throughout the midline but is seen consistently between embryos. The lack of uniform

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

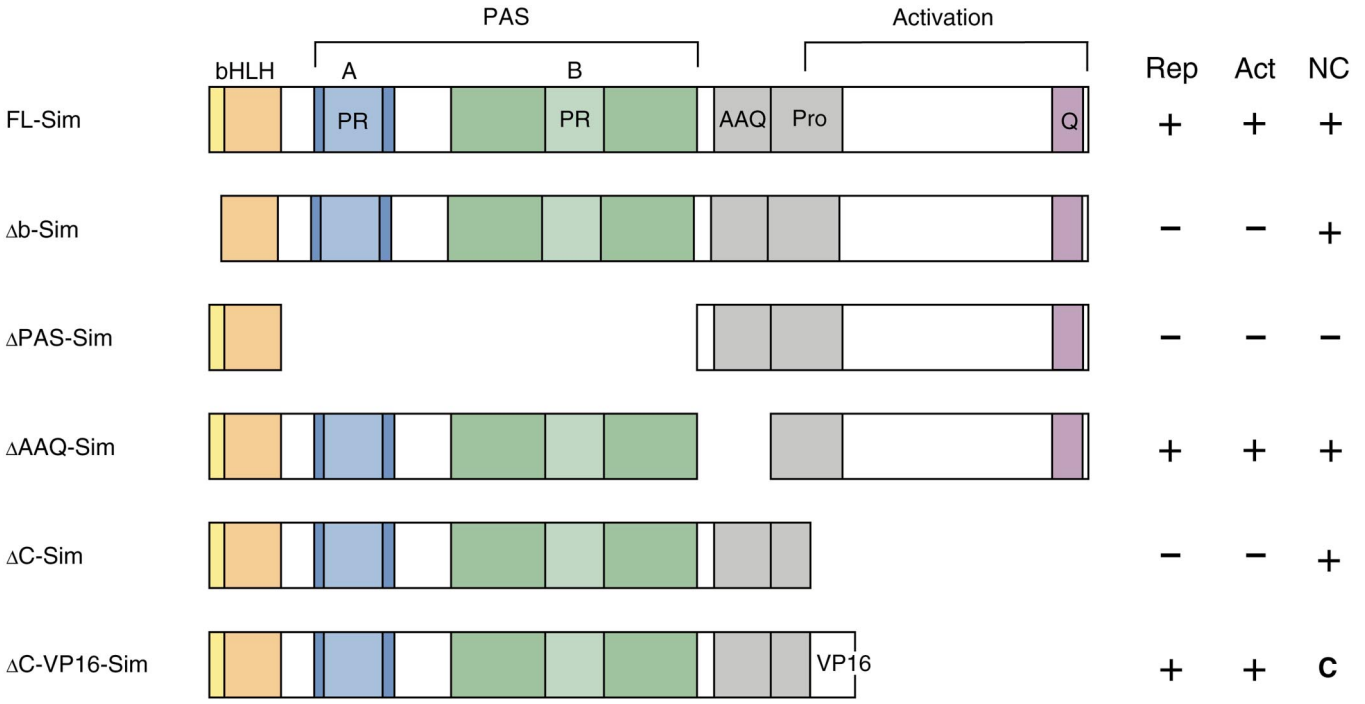

FIG. 5. The structure of Sim proteins used to identify domains important for *vnd* repression. The full-length Sim (FL-Sim) protein, shown at top, is illustrated to highlight Sim structural and functional regions. The bHLH domain consists of a DNA-binding basic region (yellow) and HLH dimerization region (orange). The PAS domain contains two conserved regions, PAS-A (blue) and PAS-B (green), separated by a spacer (white). There is a PAS repeat (PR) found within both PAS-A and PAS-B. Following the PAS domain are an Ala-Ala-Gln (AAQ) repeat region and a proline-rich (Pro) region shown in gray. There are multiple activation regions in the C-terminus of Sim, including a Gln-rich region (Q; purple). Sim mutants assayed include: (1) Δb-Sim, which deletes the basic region; (2) ΔPAS-Sim, which removes the PAS domain and the spacer between the HLH and the PAS domains; (3) ΔAAQ-Sim, which deletes the AAQ region; and (4) ΔC-Sim, which removes the C-terminal region. Also tested was a heterologous activation domain construct, ΔC-VP16-Sim. This construct replaced the C-terminal activation region of Sim with an activation domain from the HSV VP16 protein (VP16). Results of the experiments are shown to the right of the constructs. UAS-FL-Sim- and UAS-Sim-altered transgenic strains were generated and assayed *in vivo* for: (1) repression of *vnd* (Rep), (2) activation of the CME-*lacZ* reporter (Act), and (3) the ability to dimerize with and dictate the nuclear accumulation of Tgo (NC). (+) Sim protein functioned positively in the assay; (−) Sim protein was unable to carry out the function; (C) excess levels of Tgo protein accumulated predominantly in the cytoplasm.

repression suggests that other elements residing in the *vnd* gene may help control the maintenance of repression.

Repressive Targets of Sim

Midline repression by Sim functions by activating transcription of one or more genes that, in turn, represses transcription of genes normally expressed in the lateral CNS. The nature of these repressive factor genes and how they function are unknown, although plausible candidate genes exist. Since E(spl) proteins repress lateral CNS expression, members of this family are candidates for midline repressors, and several are expressed in the CNS midline cells early in development (m5, m7, and m8) (Wech *et al.*, 1999). In this scheme, Sim:Tgo would activate factors that would modify or interact with E(spl) proteins to repress proneural gene activity in the midline. The *vnd* upstream regulatory region contains numerous E(spl) consensus binding sites, although none of the sites lie within the 0.5-kb "d" fragment shown to be important for repression. Since the E(spl) proteins reside in midline cells well before repression occurs, it is unlikely that Sim is required for initial E(spl) transcription, although maintenance of their expression is a possibility. Other potential repressors have not yet been identified.

It is unclear what role the CMEs have within the *vnd* regulatory region. They are not involved in *vnd* repression nor do they seem to play a role in *vnd* embryonic CNS expression. We compared the sequences immediately flanking the four CMEs within the *vnd* regulatory region to CMEs found within genes known to be important for midline activation by Sim. The consensus for sites within genes positively activated by Sim is (A/T)ACGTG, while the consensus for the CMEs within the *vnd* regulatory region is GACGTG (three of four sites had a G at the first residue). Otherwise, the sequences varied widely and no

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

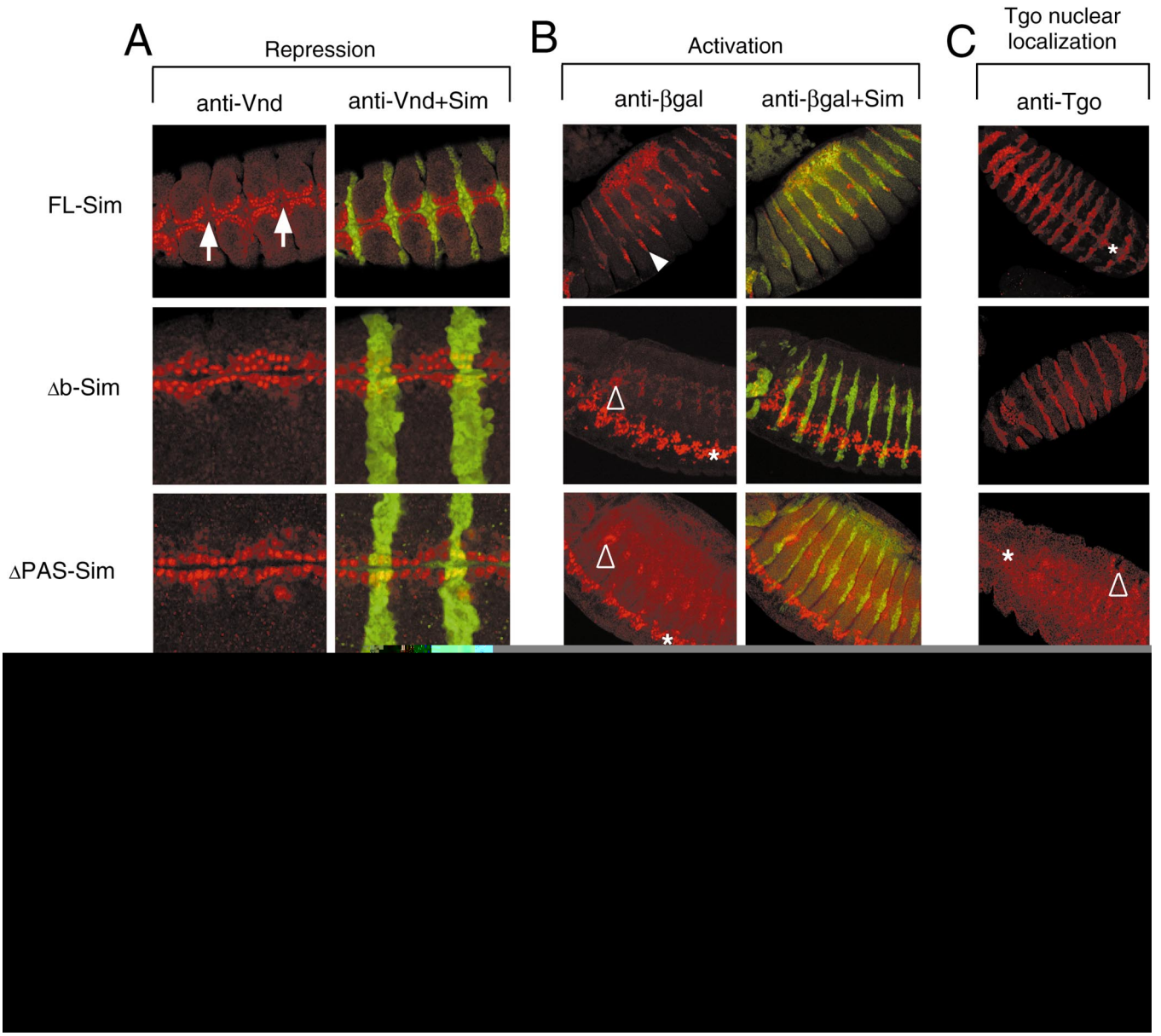

FIG. 6. The basic, PAS, and activation domains of Sim are necessary for repression of *vnd*. UAS-FL-Sim and UAS-Sim deletion constructs were ectopically expressed in circumferential stripes using the *en*-Gal4 driver, and *vnd* expression was assessed in the intersecting regions of the lateral CNS cells. Expression of *vnd* in the lateral CNS was assayed by anti-Vnd immunostaining (red), while Sim in the *en* stripes (green) was determined by: (1) visualizing GFP derived from GFP-tagged forms of Sim (FL-Sim, Δb-Sim, ΔAAQ-Sim), (2) immunostaining with anti-Sim (ΔPAS-Sim), or (3) immunostaining with anti-Tgo (ΔC-Sim). In the latter case, presence of nuclear Tgo corresponding to *en* stripes (open arrow) indicates colocalization with nuclear Sim. *en*-independent nuclear Tgo is also observed in the tracheal pits or branches (open arrowhead) (Ward *et al.*, 1998). (A) UAS FL-Sim and ΔAAQ-Sim repressed *vnd* expression in most cells of the intersecting segments (filled arrows), but also showed some light additional Vnd-positive staining in cells beneath the *en* stripe and an expansion of Vnd-positive cells directly adjacent to the stripe. The Sim deletion mutants, Δb-Sim, ΔPAS-Sim, and ΔC-Sim, failed to repress *vnd* expression. Shown are ventral views of stage 11 embryos; anterior is to the left. (B) The Sim basic, PAS, and activation domains are required to activate transcription of the CME-*lacZ* reporter. The ability of mutant forms of Sim to activate transcription was assayed by examining the expression of a 4×CME-*lacZ* reporter in *en*-Gal4 × UAS mutant-Sim embryos. FL-Sim and mutant-Sim constructs were expressed in *en* stripes (green), and activation of the reporter in lateral stripes was assayed by anti-*β*gal staining (red). Stage 13–15 embryos are shown in ventrolateral views and cells expressing both *lacZ* and *sim* are yellow. FL-Sim and ΔAAQ-Sim activated transcription from the multimerized reporter in *en* stripes (filled arrowheads), while the Δb-Sim, ΔPAS-Sim, and ΔC-Sim constructs failed to activate the reporter despite being expressed at high levels. In embryos that did not activate the reporter, the embryos shown consist of stacked serial sections

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.