REVIEW

Control of cell lineage-specific development and transcription by bHLH–PAS proteins

[1]Stephen T. Crews

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

We dance round in a ring and suppose
But the secret sits in the middle and knows

Robert Frost
The Secret Sits

The basic–helix–loop–helix-PAS (bHLH–PAS) proteins comprise a prominent class of transcriptional regulators that control a variety of developmental and physiological events including neurogenesis, tracheal and salivary duct formation, toxin metabolism, circadian rhythms, response to hypoxia, and hormone receptor function. The bHLH–PAS proteins have a number of similarities with other bHLH protein subfamilies (Littlewood and Evan 1995). bHLH–PAS proteins usually function as dimeric DNA-binding protein complexes; although some bHLH–PAS proteins can form homodimers, the most common functional unit is comprised of heterodimers. These heterodimers consist of one partner that is broadly expressed, and another whose expression or function is restricted spatially, temporally, or by the presence of inducers. Just as other vertebrate and invertebrate bHLH proteins control cell lineage specification (Weintraub et al. 1991; Jan and Jan 1993), bHLH–PAS proteins are also important cell lineage regulators (Thomas et al. 1988, Isaacs and Andrew 1996; Wilk et al. 1996). The combinatorial and interactive properties of bHLH–PAS proteins provide a variety of potential mechanisms to control their function as transcriptional regulators, which may help explain their widespread use in complex biological events. The purpose of this review is to describe characteristics of the bHLH–PAS protein subfamily, in particular, how bHLH–PAS proteins control lineage-specific gene transcription and development of the *Drosophila* CNS midline cells and respiratory system, and to discuss the evolutionary implications of the bHLH–PAS/Arnt regulatory cassette. The underlying mechanisms employed by the bHLH–PAS developmental regulatory proteins discussed here may prove to be common in both vertebrates and invertebrates, and provide a general understanding into how regulatory proteins control the formation of cell lineages.

[1]E-MAIL steve_crews@unc.edu; FAX (919) 962-3155.

bHLH–PAS proteins share a conserved sequence structure

The sequence organization of bHLH–PAS proteins is remarkably similar (Fig. 1). The bHLH domain is located near the amino terminus. The basic region binds DNA and the HLH domain promotes dimerization. These residues are followed closely by the PAS domain. The carboxy-terminal residues contain transcriptional activation domains (Franks and Crews 1994; Jain et al. 1994; Li et al. 1994) or repression domains (Moffett et al. 1997). The unique feature of bHLH–PAS proteins is the PAS domain, named for the first three proteins identified with this motif: the *Drosophila* Period (Per), human Arnt, and *Drosophila* Single-minded (Sim) (Nambu et al. 1991). The PAS domain found in bHLH–PAS proteins is ~260–310 amino acids long (Crews et al. 1988) (Fig. 1); it is subdivided into two well-conserved regions, PAS-A and PAS-B, separated by a poorly conserved spacer. Within both the A and B regions lies a copy of a 44-amino acid repeat referred to as the PAS repeat (Crews et al. 1988; Nambu et al. 1996). The repeat begins with a nearly invariant Phe residue and terminates with a His *X X* Asp motif (Wang et al. 1995; Nambu et al. 1996). Overall, the PAS domain is not well-conserved; nonorthologous family members are often <25% identical in amino acid sequence. It is not surprising, given its size and diversity in sequence, that the PAS domain can mediate a number of biochemical functions. It is used for dimerization between PAS proteins (Huang et al. 1993), small molecule binding (Dolwick et al. 1993; Coumailleau et al. 1995), and interactions with non-PAS proteins (Coumailleau et al. 1995; Gekakis et al. 1995).

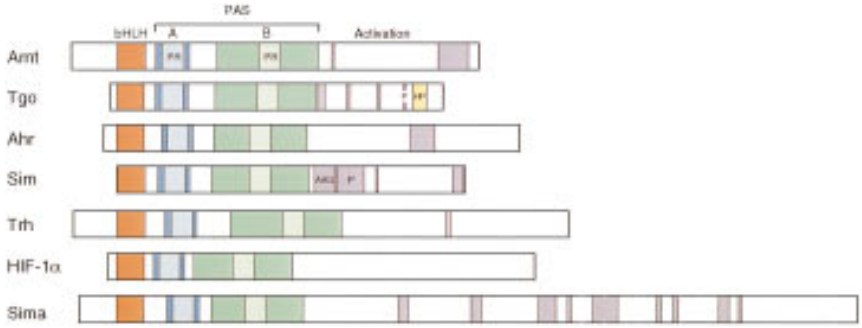

Figure 1. The structure of bHLH–PAS proteins is highly conserved. Shown are representative of the human aryl hydrocarbon receptor nuclear translocator (Arnt), *Drosophila* Tango (Tgo), murine aryl hydrocarbon receptor (Ahr), *Drosophila* Single-minded (Sim), human hypoxia inducible factor-1α, (HIF-1α), and *Drosophila* Similar (Sima) proteins. The bHLH domain (red) is near the amino terminus followed closely by the PAS domain. PAS consists of two conserved regions: A (blue) and B (green) separated by a relatively unconserved spacer. Within each PAS region is a 44-amino acid PAS repeat (PR). The carboxyl termini of these bHLH–PAS proteins function as transcriptional activation domains. Shown in unlabeled purple blocks are poly[glutamine] repeats associated with activation function. Also shown are a His–Pro-rich (HP) Paired repeat in Tgo (orange), and an Ala–Ala–Gln (AAQ) repeat and Pro-rich (P) region in Sim.

Drosophila sim and mammalian aryl hydrocarbon receptor: paradigms for bHLH–PAS protein function

The first two bHLH–PAS proteins extensively studied were the *Drosophila sim* gene and the mammalian Ahr. Genetic and cellular analysis of *sim* provided the initial evidence that bHLH–PAS proteins could act as lineage-specific developmental regulatory proteins. These experiments showed that *sim* function is required for all midline transcription and development (Thomas et al. 1988; Nambu et al. 1990, 1991). Numerous target genes of *sim* were identified and transgenic experiments identified a regulatory element that acts as a Sim-binding site and is required for CNS midline cell transcription (Wharton and Crews 1993; Wharton et al. 1994). This established a foundation for further molecular genetic analysis of bHLH–PAS protein control of developmental processes.

The biochemistry of bHLH–PAS protein function has been described in greatest detail for the mammalian aryl hydrocarbon receptor complex (AHRC; also referred to as the dioxin receptor) (Fig. 2) (for review, see Swanson and Bradfield 1993; Harkinson 1995; Whitlock et al. 1996; Rowlands and Gustafsson 1997). This complex activates transcription of genes that encode proteins involved in toxin metabolism, such as cytochrome P450IA1 and glutathione *S*-transferase (GST). The functional DNA-binding complex consists of the Ahr ligand-binding bHLH–PAS protein (Burbach et al. 1992; Ema et al. 1992) and another bHLH–PAS protein, Arnt (Hoffman et al. 1991). Transcriptional control involves AHRC binding to the xenobiotic response element (XRE) that contains a GC-GTG core binding sequence. The induction of AHRC function is controlled by ligand (e.g., dioxin) binding to Ahr, and thus AHRC constitutes a regulated signaling pathway. Ahr is found in the unliganded state in the cytoplasm complexed to heat shock protein 90 (Hsp90) and Ahr interaction factor (AIF) (Ma and Whitlock 1997). These proteins are thought to keep the unliganded Ahr in a state responsive to ligand binding and interaction with Arnt. Ligand passes through the plasma membrane and binds to a site in the Ahr PAS domain. Although the sequence of events is controversial, Ahr dissociates from Hsp90 and AIF, binds to Arnt, and the Ahr::Arnt complex enters nuclei, where it activates transcription. Analysis of AHRC function has established a paradigm for bHLH–PAS protein function: signal transduction by small molecule binding, control of nuclear localization, bHLH–PAS protein heterodimerization with Arnt, DNA binding to XRE-related sequences, and transcriptional activation. This paradigm was instrumental in investi-

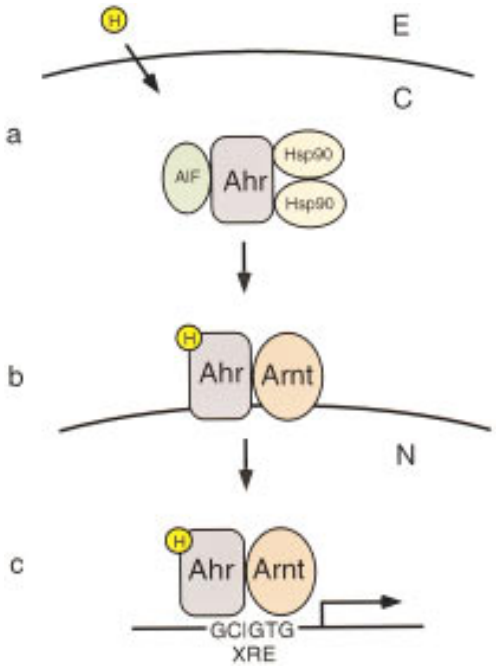

Figure 2. Ligand-controlled regulation of bHLH–PAS protein function by AHRC. (*a*) Ahr resides in the cytoplasm (C) complexed with AIF and two molecules of Hsp90. Aryl hydrocarbons (yellow) diffuse through the membrane from the extracellular (E) side and bind Ahr. (*b*) Ligand-bound Ahr dimerizes with Arnt and accessory proteins dissociate from Ahr. (*c*) The Ahr::Arnt complex enters the nucleus (N), binds the XRE, and activates target gene transcription. On the XRE, Ahr binds the GC half-site and Arnt binds the GTG half-site.

gating the molecular genetics of how *Drosophila sim* controls CNS midline development and transcription.

Development and function of the *Drosophila* CNS midline cells

The *Drosophila* embryonic CNS (for review, see Goodman and Doe 1993) consists of a brain and ventral nerve cord (vnc). The vnc is comprised of 14 fused ganglia, each consisting of 400 neurons and additional glia. CNS neurons extend axons that join together to form axon bundles. Longitudinal axon bundles connect the ganglia and run along the anterior–posterior axis of the vnc, whereas within each ganglion two commissural axon bundles cross the midline and connect each side of the vnc. Each hemiganglion is separated by a set of CNS midline cells that are best considered as a discrete tissue distinct from the rest of the CNS (Nambu et al. 1993). They have a different developmental origin, are specified by distinct regulatory genes, and play important roles in controlling the formation of adjacent tissues and guiding commissure formation.

The existence of the insect CNS midline cells was recognized over a century ago (Wheeler 1893), and the development and function of these cells first characterized using modern techniques in the grasshopper embryo (Bate and Grunewald 1981; Goodman 1982). More recently, the development and function of the CNS midline cells (Fig. 3) have been extensively studied in *Drosophila* by a number of laboratories (for review, see Nambu et a. 1993; Bossing and Technau 1994). The mature midline cells consist of two to six midline glia, two midline precursor 1 (MP1) interneurons, two unpaired median interneurons (UMI), six ventral unpaired median (VUM) motorneurons and interneurons, and five to eight interneuronal and motorneuronal progeny of the median neuroblast (MNB). Midline cells have an unusual origin: They are derived from cells that are initially separate in the embryo. In the blastoderm embryo, precursors to the CNS midline cells form two single-cell-wide groups of cells (approximately four cells per hemisegment) that lie between the presumptive mesoderm and lateral neuroectoderm (Fig. 4; see also Thomas et al. 1988). These cells, referred to as the ''mesectoderm,'' join together at the end of gastrulation to form seven to eight midline precursor cells per segment.

The midline precursor cells undergo a synchronous cell division and then a cell shape change, in which the nuclei migrate internally and leave a cytoplasmic projection joined to the surface of the embryo (Nambu et al. 1991). Most precursor cells will not divide again; they differentiate into neurons and glia (Bossing and Technau 1994). The current view is that two precursor cells give rise to the midline glia, one to the pair of MP1s, one to the UMIs, three to the six VUMs, and one to the median neuroblast and its progeny (Bossing and Technau 1994). The cellular simplicity of the 22–26 nerve cells and glia that reside at the midline of each ganglion has made them an attractive system for studying the molecular genetics of neural development and function. However,

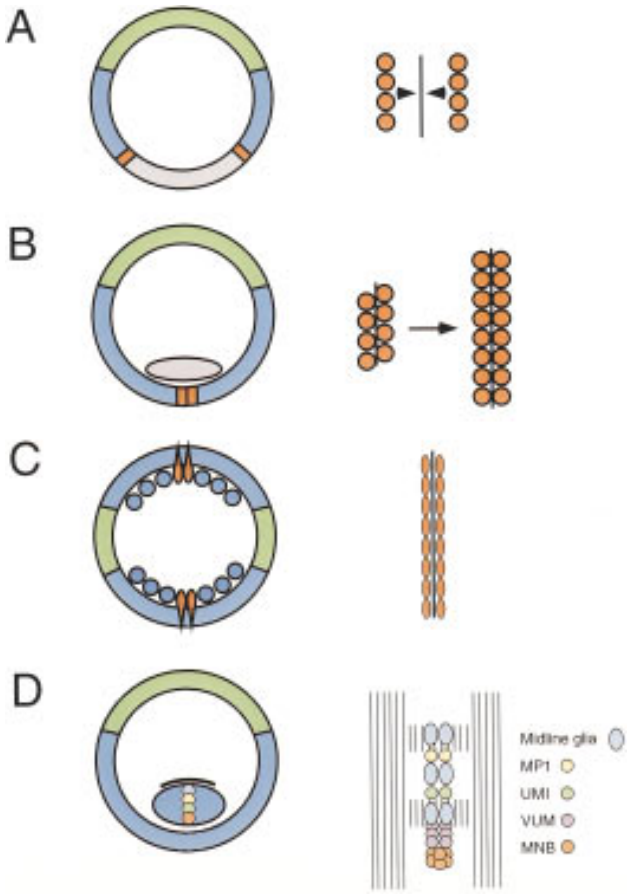

Figure 3. Development and identity of the CNS midline cells. (*Left*) A series of schematized cross sections of *Drosophila* embryos at different developmental stages (Campos-Ortega and Hartenstein 1985) with the dorsal side at the *top*. (*Right*) A horizontal view of an individual segment or ganglion depicting midline cell development. (*A*) Stage 5 blastoderm embryo showing the primordia of the mesoderm (lavendar), mesectoderm (red), ventrolateral neuroectoderm (blue), and dorsal ectoderm and extraembryonic membranes (green). The midline precursors consist of single cell wide strips of cells on either side of the mesoderm. Each hemisegment has about four midline precursors. At gastrulation the midline cells migrate (arrowheads) towards the ventral midline of the ectoderm (line). (*B*) After gastrulation, the mesoderm has migrated internally and the midline precursor cells (red) are joined together at the ventral midline. The midline precursors then undergo a synchronous cell division. (*C*) During the germband elongation phase, the midline cells (red) send their nuclei internally and maintain a cytoplasmic projection at the surface of the embryo. The lateral neuroblasts (blue circles) emerge from the underlying ventrolateral ectoderm and form a neural precursor layer. (*D*) After germband retraction, the midline and lateral neural precursor cells differentiate into neurons and glia and form the mature vnc. The midline cells are depicted as colored circles and the lateral neurons of the vnc are shown in blue. Axon bundles form above the vnc. (*Right*) A single ganglion illustrating the three pair of midline glia and the four groups of midline neurons. The midline cell positions along the anterior–posterior axis are idealized (see Bossing and Technau 1994 for exact positions within the ganglion). The vertical-line structure represents the orthogonal axon scaffold showing two longitudinal connectives joined by the anterior and posterior commissures that cross the midline.

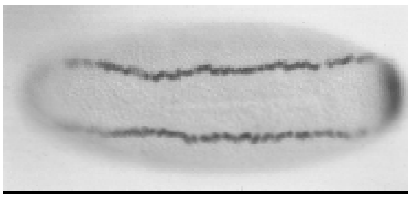

Figure 4. Expression of *sim* in the mesectodermal cells. Ventral view of a gastrulating embryo (stage 6) hybridized in situ to a *sim* cDNA probe showing the two single-cell wide stripes of *sim* transcription. Anterior is to the *left*.

the truly remarkable aspect of the CNS midline cells are their additional developmental roles.

Early in embryonic development when cells are acquiring their tissue-specific fates, mesectodermal cells are in contact with the adjacent lateral neuroectoderm and mesoderm. The lateral neuroectodermal cells give rise to both ventral epidermis and the lateral CNS. Genetic studies have shown that the mesectoderm instructs the adjacent ectoderm to form the ventral epidermis (Kim and Crews 1993). In addition, some ectodermal cells give rise to lateral neuroblasts whose proper development is dependent on a signal from the mesectoderm (Menne et al. 1997; Y. Lee, S.T. Crews, and S.H. Kim, in prep.). The adjacent mesoderm also requires an influence from the mesectoderm for proper development (Lüer et al. 1997; Zhou et al. 1997). All of these effects are mediated by a signal emanating from the midline (Mayer and Nüsslein-Volhard 1988; Kim and Crews 1993; Golembo et al. 1996; Xiao et al. 1996). The mesectodermal cells secrete the Spitz protein, which is related to vertebrate transforming growth factor-α. Spitz acts a ligand for the *Drosophila* ortholog of the epidermal growth factor receptor (DER), which is present on the adjacent ectoderm and mesoderm (Raz and Shilo 1992). Midline cells also influence the migration of a subset of muscle precursor cells (Lewis and Crews 1994), although it is not known how this is accomplished.

Another important function of the insect CNS midline cells, which is shared with vertebrate ventral midline or floor plate cells, is attraction of commissural axons to the midline (Goodman 1996). Approximately 90% of *Drosophila* CNS neurons extend axons across the midline to the contralateral side of the CNS, where they join with other axons and migrate to their synaptic targets. The midline cells secrete Netrin proteins (Harris et al. 1996; Mitchell et al. 1996) that attract axons expressing the Netrin receptor, the product of the *frazzled* gene (Kolodziej et al. 1996), to the midline. The midline cells also act as a barrier, repelling axons that are either not programmed to cross the midline or to prevent those that have crossed from migrating back (Seeger et al. 1993; Tear et al. 1996). Midline glia, which ensheath the commissural axon bundles (Jacobs and Goodman 1989), physically separate the anterior and posterior commissures as they migrate to their final positions (Klämbt et al. 1991). In summary, the CNS midline cells are a functionally rich set of cells that not only act as motorneurons, interneurons, and glia, but also influence axon guidance and the development of the epidermis, mesoderm, and the lateral CNS.

Mesectodermal specification results from the initial activation of *sim* transcription by dorsal–ventral patterning genes

The *sim* gene acts as a simple genetic switch for midline cell development. When the gene is activated in ventral–lateral ectodermal cells around the time of gastrulation, it drives those cells into the CNS midline cell lineage. Specification of the CNS midline lineage is dependent on precise expression of *sim* in the mesectodermal precursor cells. Initial *sim* transcription is restricted to these two single-cell-wide rows of ectodermal cells that separate mesoderm from lateral neuroectoderm; there is no refinement from an initial broader domain of expression (Thomas et al. 1988). This represents the most extreme example of initial dorsal-ventral patterning in that high levels of *sim* expression occur in a single row of cells while it is undetectable in the adjacent cells.

Biochemical, genetic, and molecular studies suggest how this is achieved (Fig. 5). Genetic studies implicate the *dorsal*, *snail* (*sna*), *twist* (*twi*), *scute* (*sc*), *daughterless* (*da*), and *Notch* genes in *sim* activation (Fig. 5A) (Kosman et al. 1991; Leptin 1991; Rao et al. 1991; Kasai et al. 1992; Lewis 1994). All are transcription factors or, in the case of Notch, presumably function through transcription factors. The Dorsal protein, an NF-κb relative, forms a morphogenetic gradient that is the key regulator of tissue specification along the dorsal–ventral axis of the embryo (for review, see Rusch and Levine 1996). Dorsal forms a nuclear gradient with highest concentrations at the ventral surface of the embryo (Roth et al. 1989; Rushlow et al. 1989; Steward 1989). Dorsal, in turn, activates *twi* and *sna* expression. The Twi bHLH protein also forms a gradient along the ventral side of the embryo. Both Dorsal and Twi proteins are found in the mesectodermal cell anlage, and in both cases these cells lie in a steep region of their gradients (Leptin and Grunewald 1990; Kosman et al. 1991). The distribution of the Sna zinc finger protein is more highly restricted; it is found at high concentrations in the mesoderm but is absent in the adjacent mesectoderm (Kosman et al. 1991; Leptin 1991). Da and Sc bHLH proteins are expressed at this time throughout the embryo, forming an E-box (ACNNGT) binding heterodimer (Jiang and Levine 1993). Thus, Dorsal, Twi, and Da::Sc act together to activate *sim* ventrally in the mesoderm and mesectoderm. Sna, which is restricted to the mesoderm, represses *sim* in those cells (Nambu et al. 1990; Rushlow and Arora 1990), leading to *sim* activation in the mesectoderm. In addition, the *Notch* signaling pathway positively regulates *sim* transcription (Lewis 1994; Menne and Klämbt 1994; Martin-Bermudo et al. 1995).

Biochemical and germ-line transformation studies indicate that the dorsal–ventral patterning proteins directly control *sim* transcription (Fig. 5B) (Kasai et al. 1992; Y. Kasai, M. Sonnenfeld, J. Lewis, S. Stahl, and S.

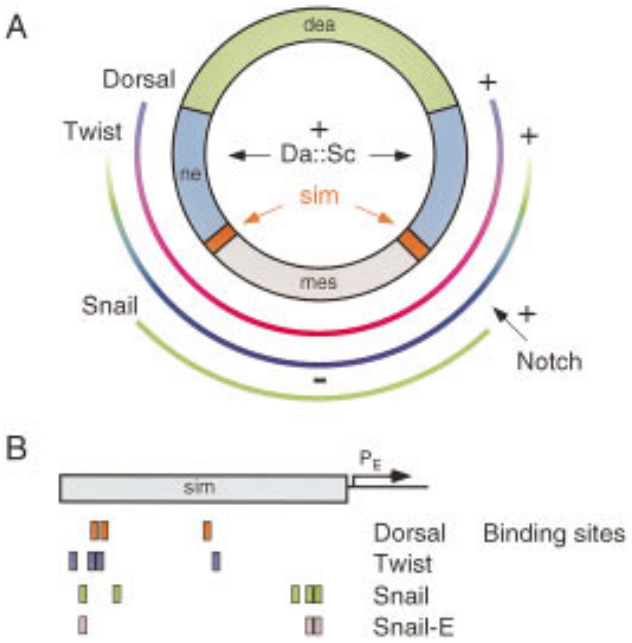

Figure 5. Control of initial *sim* transcription by dorsal–ventral patterning genes. (*A*) Schematic cross section of a blastoderm embryo with dorsal at *top*. Regions depicted are dorsal ectoderm plus amnioserosa (dea; green), ventrolateral neuroectoderm (blue; ne), mesectoderm (red), and mesoderm (mes; lavendar). The *sim* gene is specifically activated in the mesectodermal cells. Positive influences (+) that activate *sim* transcription ventrally result from the combined actions of (1) Dorsal protein, shown as a gradient with highest concentrations ventrally, (2) Twi protein, also shown as a gradient with highest concentrations ventrally, (3) Da::Sc heterodimers, which are found throughout the embryo, and (4) Notch signaling pathway. Sna protein is found in the mesoderm and represses (−) *sim* transcription in those cells. (*B*) The 2.8-kb *sim* regulatory DNA that is sufficient for initial *sim* transcription in mesectodermal cells is indicated by the blue box. The arrow denotes P_E and the direction of transcription. Below are shown the locations of binding sites for Dorsal, Twi, and Sna. The lavendar boxes at the bottom show the subset of Sna binding sites that contain E-boxes that bind Da::Sc. Mutation of selected Dorsal, Twi, and Sna–E sequences results in an absence of midline transcription indicating that these proteins function in vivo to directly regulate *sim* transcription.

Crews, in prep.). The *sim* gene contains two promoters, one (P_E) that controls early midline transcription and another (P_L) that controls late midline transcription (Nambu et al. 1991; Muralidhar et al. 1993). Comparing the sequence of the *sim* P_E regulatory region between two different *Drosophila* species revealed a series of conserved sequence elements (Wharton et al. 1994), which include predicted binding sites for Dorsal, Twi, Da::Sc, and Sna. These factors bind to the *sim* P_E regulatory region in vitro (Fig. 5B; Kasai et al. 1991; Y. Kasai, M. Sonnenfeld, J. Lewis, S. Stahl, and S. Crews, in prep.). Mutagenesis and analysis by germ-line transformation indicate that the binding sites are used in vivo: elimination of Dorsal, Twi, and Da::Sc binding sites results in an absence of initial mesectodermal transcription. These observations suggest a model in which the *sim* early regulatory region employs binding sites for the cooperatively-acting Dorsal, Twi, and Da::Sc transcription factors. Presumably, the Dorsal nuclear gradient is insufficient to establish on-off transcription with single cell resolution, and thus, additional proteins are required. Sna sets the ventral boundary of initial *sim* transcription by repressing *sim* in the adjacent mesoderm. As several Da::Sc binding sites are embedded within a subset of Sna binding sites (Kasai et al. 1992), one attractive model of Sna repression is that it directly competes with Da::Sc binding sites in the mesoderm (Ip et al. 1992). Is there a similar repressor designed to limit the dorsal boundary? This remains possible. Alternatively, the dorsal boundary may be attributable strictly to the steep concentration gradients of Dorsal and Twi, allowing activation of *sim* transcription in the mesectoderm but not in more dorsal cells. In summary, the *sim* gene is designed to respond directly to multiple positively and negatively-acting regulatory proteins that are expressed in the early embryo. These proteins direct *sim* spatial expression, and also dictate that *sim* is first expressed at gastrulation when cell lineage specification is established.

sim controls CNS midline cell specification

Sim protein specifically accumulates in mesectodermal cell nuclei during gastrulation (Fig. 6A) at the time when ectodermal cells are acquiring their fates (Crews et al. 1988). *sim* is not expressed in other ectodermal cells. Sim protein is expressed in the midline cells throughout neurogenesis and is present in the differentiated midline neurons and glia (Crews et al. 1988). Null mutants of *sim* have a complete absence of midline cell development: midline precursor cells fail to divide, and subsequently do not undergo their characteristic cell shape changes or differentiate into neurons and glia (Thomas et al. 1988; Nambu et al. 1991). Sim exerts its affect by controlling target gene transcription (Nambu et al. 1990). Midline expression of >20 genes is abolished in *sim* mutant embryos (Table 1), and probably all midline transcription is directly or indirectly dependent on *sim* function. The master regulatory role of *sim* is reinforced by experiments in which *sim* is ectopically expressed using a heat shock–*sim* transgene (Nambu et al. 1991). If *sim* is induced in neuroectodermal cells as they are adopting their fates, they are transformed from lateral CNS into CNS midline cells. This is accompanied by ectopic gene transcription of midline-expressed genes.

Midline cells in *sim* mutant embryos take on a lateral neuroectodermal cell fate and misexpress genes that correspond to this lineage (Chang et al. 1993; Mellerick and Nirenberg 1995; Xiao et al. 1996). Thus, it appears that the default state of all neuroectodermal cells is lateral CNS. When *sim* is turned on in these cells, it activates midline transcription and represses lateral CNS transcription. The combination of these two activities results in CNS midline cell development.

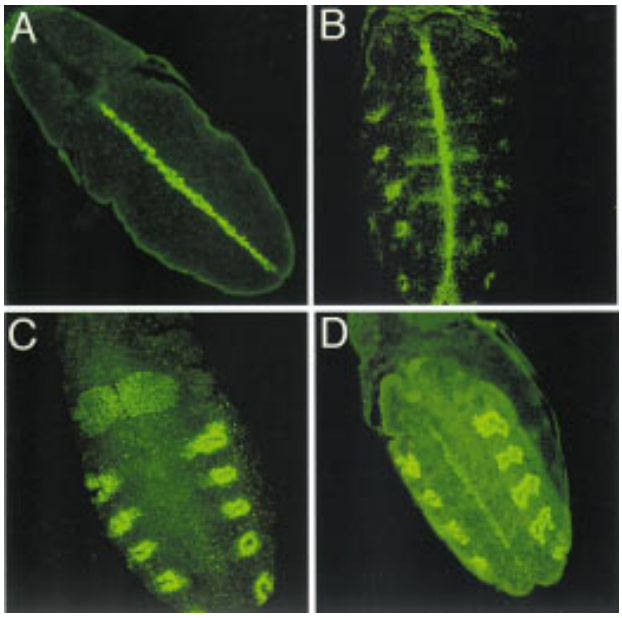

Figure 6. Embryonic localization of bHLH–PAS proteins and their target gene transcription. Confocal images of fluorescently stained whole-mount *Drosophila* wild-type embryos are shown. Anterior is to the *left*. (*A*) Dorsal view of a stage 10 embryo stained with anti-Sim antibodies showing nuclear accumulation in the CNS midline cells. (*B*) Dorsal view of a stage 11 embryo containing the P[4×CME]–*lacZ* transgene stained with anti-β-galactosidase antibodies. Expression is observed in the CNS midline and tracheal cells, indicating that the CME acts as a binding site for Sim::Tgo and Trh::Tgo heterodimers in vivo. The embryo has been overexposed to reveal the relatively weak tracheal staining. (*C*) Ventral view of a stage 11 embryo stained with anti-Trh antibodies showing nuclear accumulation in the tracheal pits and salivary gland placode. (*D*) Dorsal view of a stage 11 embryo stained with anti-Tgo antibodies showing cytoplasmic localization in all embryonic cells except nuclear localization in the CNS midline and tracheal pits.

Sim controls target gene transcription through a midline enhancer element

Once activated in CNS midline cells, Sim controls midline transcription, and also maintains its own expression by positive autoregulation (Nambu et al. 1991; Muralidhar et al. 1993). Progress in understanding how *sim* controls midline transcription has been achieved by identifying target genes and their midline enhancer elements and identifying the bHLH–PAS dimerization partner of Sim. Numerous genes are expressed in the CNS midline precursor cells soon after the initial appearance of Sim in cell nuclei (Crews et al. 1992) that are likely to be directly regulated by *sim*. In many cases, the genes were cloned without prior knowledge of their midline expression, and subsequently shown to be expressed in the CNS midline cells. However, two enhancer trap screens have identified additional CNS midline-expressed genes (Klämbt et al. 1991; Crews et al. 1992).

Table 1. *Genes whose CNS midline expression is absent in* Sim *mutant embryos*

Genes		Enhancer traps	
abrupt	*breathless*	47F	AA41
center divider	*DER*	AA142	BP28
drifter	*engrailed*	T13	X55
midline fasciclin	*orthodenticle*		
period	*rhomboid*		
Single-minded	*slit*		
spitz	*Toll*		

Nambu et al. (1990, 1991); Klämbt et al. (1991); Kim and Crews (1993); Sonnenfeld and Jacobs (1994); Ohshiro and Saigo (1997); S.T. Crews (unpubl.).

Four *sim* target genes, *breathless* (*btl*), *sim*, *slit*, and *Toll* (*Tl*), have been characterized in detail at the molecular level (Wharton and Crews 1993; Wharton et al. 1994; Ohshiro and Saigo 1997). Each of these genes represents a distinct mode of midline regulation: (1) *Tl* is expressed in midline precursor cells; (2) *sim* is an autoregulatory target; (3) *slit* is expressed in differentiated midline glial cells; and (4) *btl* is expressed in both midline and tracheal cells. Each regulatory region was assayed for the ability to drive *lacZ* in the midline cells. Deletional analysis and site-directed mutagenesis identified a CNS midline enhancer element (CME), with a core ACGTG sequence (Wharton et al. 1994). This element is found in multiple copies in *btl*, *sim*, and *Tl* and as a single copy in *slit*. The CME is required for midline transcription in all four genes (Wharton et al. 1994; Ohshiro and Saigo 1997) and, when multimerized, *Tl* site 4 is sufficient to drive transcription from a heterologous promoter in the CNS midline precursor cells and differentiated midline neurons and glia (Fig. 6B) (Wharton et al. 1994; Sonnenfeld et al. 1997).

Sim::Tango heterodimers activate CNS midline transcription

Theoretical (Wharton et al. 1994) and experimental (Swanson et al. 1995) considerations predicted that the ACGTG core sequence was a binding site for heterodimers between Sim and a *Drosophila* Arnt-like protein. The *Drosophila tango* (*tgo*) gene was cloned and shown to be highly related to mammalian Arnt (Ohshiro and Saigo 1997; Sonnenfeld et al. 1997). Several lines of evidence indicate that Sim forms heterodimers with Tgo to bind the CME in vivo and activate CNS midline gene transcription (Sonnenfeld et al. 1997). Both proteins are found in CNS midline cells during embryonic development. Sim and Tgo can form dimers and activate transcription of a multimerized CME when cotransfected into *Drosophila* cell culture. Mutations in both *sim* and *tgo* affect CNS midline transcription and development, including transcription of the multimerized CME transgene. The phenotype of *tgo* mutants is less severe than null *sim* or *trh* mutations, although this is most likely because of a maternal contribution of *tgo* and the use of hypomorphic *tgo* alleles. Gene dosage experiments reveal that loss of a single copy of *sim* enhances the CNS

midline phenotype of *tgo* mutations resulting in a severe ''*sim*-like'' collapsed CNS phenotype. Activation of transcription in vivo by Sim::Tgo heterodimers is likely to be direct since both proteins possess potent transcriptional activation domains (Franks and Crews 1994; Sonnenfeld et al. 1997).

Sim autoregulation and early and late phase transcription

Once *sim* is localized to cell nuclei, it activates transcription of target genes in CNS midline precursor cells. In addition, *sim* transcription is controlled by a positively acting autoregulatory feedback loop (Nambu et al. 1991). This may be a mechanism in which the midline lineage-conferring properties of *sim* are maintained throughout embryonic development. Autoregulation has an interesting twist in that both *sim* promoters, P_E and P_L, are autoregulated. P_E continues to be transcribed in the midline because of Sim::Tgo function (Nambu et al. 1991). Eventually, this transcription is extinguished by an unknown mechanism as the midline precursors differentiate into neurons and glia. P_L is also activated by Sim autoregulation (Nambu et al. 1991; Muralidhar et al. 1993). This promoter drives transcription in the midline precursor cells, later in the differentiated midline cells, and in a subset of muscle precursor cells (Lewis and Crews 1994). Although it is clear that the maintenance of *sim* transcription is dependent on *sim*, the developmental significance of later *sim* transcription is unknown, as genetic studies that completely eliminate only late *sim* function have not been carried out. However, the midline glial enhancer of the *slit* gene contains a CME that is required for midline transcription suggesting that Sim::Tgo or related proteins function in late phases of midline transcription (Wharton et al. 1994).

Mammalian *sim* and Down Syndrome

Two mammalian *sim* orthologs (*Sim1* and *Sim2*) have been discovered that share a number of functional similarities with *Drosophila sim* (for review, see Michaud and Fan 1997). Both *Sim1* and *Sim2* proteins dimerize with mammalian Arnt and bind the CME in vitro (Ema et al. 1997a; Probst et al. 1997). The mammalian and *Drosophila* genes are both expressed in the developing CNS and mesoderm (Lewis and Crews 1994; Dahmane et al. 1995; Fan et al. 1996; Ema et al. 1997a). Within the CNS, *Drosophila sim* expression is restricted to the CNS midline cells. The floor plate of the vertebrate spinal cord is thought to be the analogous cell type. Neither *Sim1* nor *Sim2* is expressed in the floor plate, but both are expressed in the ventral diencephalon and *Sim1* is present in the spinal cord cells adjacent to the floor plate. The *Sim* genes are expressed early in brain development suggesting that they might play roles in neurogenesis analogous to *Drosophila sim*.

The roles of *Sim1* and *Sim2* in embryonic development should be soon forthcoming since both genes have been knocked out in mice (Michaud and Fan 1997). One other distinction is that cell culture transfection experiments suggest that both mammalian *Sim* proteins function as transcriptional repressors (Ema et al. 1997a; Moffett et al. 1997). While both cell culture and in vivo experiments have established the ability of *Drosophila* Sim to activate transcription (Franks and Crews 1994), genetic experiments have established that *sim* can also repress transcription (Chang et al. 1993; Mellerick and Nirenberg 1995; Xiao et al. 1996). It will be interesting to see if *Drosophila sim*-mediated midline repression is mechanistically similar to mammalian *Sim* repression.

The most intriguing aspect of mammalian *Sim* is that *Sim2* maps to Chromosome 21 in the region responsible for Down Syndrome (DS) (Chen et al. 1995; Dahmane et al. 1995; Muenke et al. 1995; Chrast et al. 1997). Given the important role of *sim* in *Drosophila* development and the expression of *Sim2* in cell types that are affected in DS individuals, it was proposed that *Sim2* may play a causative role in DS. This remains speculative, however, as evidence is lacking and other candidate DS genes exist. However, the existence of mouse models of DS (Reeves et al. 1995) and systematic approaches to uncover the genetic basis for DS (Lamb and Gearhart 1995) will hopefully provide answers to this question. Since DS is a trisomy of chromosome 21 (and *Sim2*), if *Sim2* does play a role in DS, one possibility is that surplus *Sim2* protein may act by excessively binding Arnt, leaving Arnt unable to interact with other bHLH–PAS proteins that are critical for proper development or cellular function.

Trachealess::Tgo heterodimers control formation of the trachea and salivary ducts

The trachealess (*trh*) gene encodes a bHLH–PAS protein that is specifically expressed in the developing trachea plus posterior spiracle, salivary gland placode, and salivary ducts (Isaac and Andrew 1996; Wilk et al. 1996). Trh protein is found in the developing tracheal pits and the later-formed tracheal tubules (Fig. 6C; see also Wilk et al. 1996). Genetic analysis indicates that *trh* is required for the formation of trachea and controls the transcription of genes involved in this process (Younossi–Hartenstein and Hartenstein 1993; Isaac and Andrew 1996; Wilk et al. 1996). Ectopic expression of *trh* results in formation of ectopic trachea (Wilk et al. 1996). *trh* is also required for formation of the posterior spiracle and salivary duct (Isaac and Andrew 1996). These results indicate that *trh* function is required for the specification and invagination of the trachea, and probably acts similarly in the development of the posterior spiracle and salivary duct.

The functional similarity between Sim and Trh is remarkable; both are lineage-specific regulators, autoregulatory, and bHLH–PAS transcriptional activators. Work described below indicates that Sim and Trh control transcription in a similar fashion by binding the same DNA sequence element using Tgo as a dimerization partner (Ohshiro and Saigo 1997; Sonnenfeld et al. 1997; Zelzer et al. 1997). This is particularly interesting since it is commonly observed that many genes expressed in the

CNS midline cells are also expressed in trachea (Manning and Krasnow 1993).

Numerous experiments indicate that Tgo is a dimerization partner for Trh, and that together they bind the CME and activate tracheal transcription in vivo. Tgo can form dimers with Trh assayed by two-hybrid analysis and co-immunoprecipitation (Sonnenfeld et al. 1997), Trh::Tgo binds CME-containing DNA in vitro (Ohshiro and Saigo 1997), and Trh::Tgo can activate transcription from CME-bearing promoters in *Drosophila* cell culture (Sonnenfeld et al. 1997). In addition, *tgo* mutants show tracheal defects, and double mutant analysis reveals genetic interactions between *trh* and *tgo*, further indicating in vivo associations (Sonnenfeld et al. 1997). Analysis of embryos harboring a multimerized *Tl* site 4 CME shows that this transgene is expressed not only in the CNS midline cells, but in the trachea and salivary duct, those cells in which *sim* and *trh* function (Fig. 6B) (Sonnenfeld et al. 1997; Zelzer et al. 1997). Consistent with the idea that both *sim* and *trh* act through the CME, mutants in *sim* specifically abolish CME midline expression, whereas mutants in *trh* abolish CME tracheal and salivary duct expression (Sonnenfeld et al. 1997). Additional evidence that Trh::Tgo functions through the CME comes from work on *btl* (Ohshiro and Saigo 1997), which is expressed in both trachea and CNS midline cells. The *btl* gene contains three CMEs upstream of the promoter and mutational analysis indicates that they are required for both CNS midline and tracheal expression. The *tgo* gene is also expressed at elevated levels in tracheal cells (Ohshiro and Saigo 1997; Sonnenfeld et al. 1997), raising the possibility that *tgo* is autoregulated by Trh::Tgo heterodimers.

Control of bHLH–PAS protein nuclear localization: Sim and Trh direct the nuclear accumulation of their respective Sim::Tgo and Trh::Tgo heterodimers

The close biochemical relationship between Sim, Trh and the ligand-binding Ahr has engendered speculation that nuclear localization of Sim and Trh are controlled by small molecule binding. This view is reinforced by the binding of Hsp90 to Sim in vitro (McGuire et al. 1996; Probst et al. 1997), suggesting that Sim may be held in a ligand-responsive state as proposed for the Ahr–Hsp90 complex. Examination of the subcellular localization of Sim, Trh, and Tgo in wild-type, mutant, and transgenic *Drosophila* embryos has provided insight into whether their nuclear localization is regulated by ligand.

Sim protein is first detected as the mesectodermal cells move towards the midline at gastrulation, and imediately accumulates in cell nuclei (Fig. 6A; see also Crews et al. 1988). Sim protein remains predominantly nuclear in the CNS midline cells throughout embryonic development. When *sim* is ectopically expressed in the embryo, Sim protein also rapidly accumulates in nuclei (Ward and Crews 1998). Thus, the localization of Sim during normal development does not provide positive evidence for regulated nuclear localization. If nuclear localization of Sim is dependent on binding to an unknown ligand, then the ligand is not spatially or temporally localized (and thus, not developmentally significant). In a similar fashion, Trh protein appears in tracheal pit cell nuclei soon after it can be detected, and remains in tracheal cell nuclei throughout embryogenesis (Fig. 6C; see also Wilk et al. 1996). Ectopic expression experiments also reveal that Trh is localized to nuclei in all embryonic cells assayed (Ward and Crews 1998).

In contrast, Tgo protein localization is more dynamic. Tgo protein is found in all embryonic cells (Sonnenfeld et al. 1997), but its nuclear localization correlates with sites of function (Ward and Crews 1998). Tgo is localized to the cytoplasm in many cells, but is nuclear in the CNS midline cells and trachea (Fig. 6D). Ectopic expression experiments show that Tgo accumulates in cell nuclei in all cells in which Sim and Trh are expressed.

These results suggest a model (Fig. 7), in which Tgo is retained in the cytoplasm in the absence of a dimerizing bHLH–PAS protein. When Sim or Trh protein appears, it forms dimers with Tgo, and the Sim::Tgo or Trh::Tgo dimer complex enters the nucleus. Although it remains possible that Sim or Trh binds and responds to a ligand during either embryogenesis or postembryonically, these results suggest that it is more likely that their nuclear localization is not ligand responsive. The role of Hsp90 in binding to Sim may be to facilitate dimerization of Sim to Tgo, as has been postulated for other bHLH proteins (Shue and Kohtz 1994), rather than promote ligand interactions. Thus, ligand-mediated control of nuclear transport may not be a feature common to all bHLH–PAS proteins and specificity of Sim::Tgo and Trh::Tgo function is instead dependent on expression of *sim* in mesectodermal cells and *trh* in ectodermal cells by dorsal/ventral and anterior/posterior patterning proteins. The situation may be similar in mammals since Arnt subcellular localization varies spatially and temporally during embryonic development (Abbott and Probst 1995). However, differences exist, since mammalian Arnt possesses nuclear localization sequences absent in *Drosophila* Tgo (Eguchi et al. 1997), and Arnt is localized to cultured cell nuclei in the absence of any obvious bHLH–PAS protein partner (Pollenz et al. 1994; Eguchi et al. 1997).

bHLH–PAS proteins regulate hypoxia responsiveness

The cellular response to oxygen deprivation can trigger a number of important physiological responses that are controlled, in part, at the level of transcription. In mammals, depending on the cell type, these can include induction of glycolytic pathway enzymes, erythopoiesis, and angiogenesis. The breakthrough concerning how these responses are controlled was the identification of the key regulatory protein, HIF (Wang et al. 1995). HIF was shown to consist of two subunits, HIF-1α, a Sim-related bHLH–PAS protein, and HIF-1β, which is Arnt. Interestingly, the binding site for HIF, the hypoxia response element (HRE), contains a core ACGTG sequence, identical to the CME (Firth et al. 1994; Semenza et al. 1994). New mammalian bHLH–PAS proteins, such

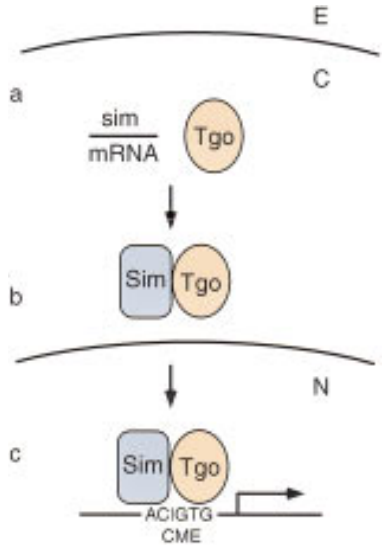

Figure 7. Ligand-independent regulation of bHLH–PAS protein function by Sim::Tgo. (*a*) Tgo resides in the cytoplasm of embryonic cells. The *sim* gene is transcribed in CNS midline cells. (*b*) Sim protein appears, and forms dimers with Tgo in the cytoplasm. (*c*) Sim::Tgo dimers enter cell nuclei, bind the CME, and activate transcription. Sim binds to the AC half-site and Tgo the GTG half-site. Trh is postulated to work in a similar fashion.

as endothelial PAS domain protein 1 (EPAS1), (Ema et al. 1997b; Hogenesch et al. 1997; Tian et al. 1997) that are related to HIF-1α and also form dimers with Arnt, are also likely to play roles in controlling the physiological response to oxygen levels.

Insects also respond to oxygen deprivation by transcriptional up-regulation of glycolytic pathway genes (Nagao et al. 1996). Biochemical studies in *Drosophila* cell culture have identified a hypoxia-inducible factor that can bind a HRE (Nagao et al. 1996). Although biochemical identification of the protein factors is lacking, Tgo may be a constituent of the HRE binding activity. If so, one candidate for Tgo's partner is the Sima bHLH–PAS protein (Nambu et al. 1996). Sima is most related to HIF-1α, is ubiquitously-expressed in the embryo as is HIF-1α, and can form stable dimers with Tgo (Nambu et al. 1996; Sonnenfeld et al. 1997). Another relevant aspect of insect respiratory physiology concerns the regulation of tracheal terminal branching. Studies in the blood-sucking insect, *Rhodnius*, demonstrated that branching is dependent on the oxygen levels of the surrounding tissues (Wigglesworth 1954). It is tempting to speculate that a HIF-like activity operating in the cells adjacent to the tracheoles may regulate terminal branching (Guillemin and Krasnow 1997).

Coregulators influence the specificity of bHLH–PAS target gene transcription

There exist a number of genes expressed in both CNS midline cells and trachea that are likely targets of both Sim and Trh. However, many genes regulated by Sim and Trh are expressed in only one of the two cell types. As Sim::Tgo and Trh::Tgo (as well as HIF-1α::Arnt) heterodimers bind the same core ACGTG sequence in vivo to regulate target gene transcription, additional regulatory elements and factors are required to generate transcriptional specificity. Experiments using the multimerized *Tl* site 4 CME in vivo and in cell culture suggest this element is sufficient for transcription in midline precursors, midline neurons and glia, trachea, and SL2 cells (Sonnenfeld et al. 1997; Zelzer et al. 1997). However, tracheal expression is weak compared to midline transcription (Fig. 6B). Misexpression studies suggest that Trh requires an additional factor restricted to dorsal ectoderm for activation of the CME, whereas Sim can activate CME transcription throughout the ectoderm and other cell types (Zelzer et al. 1997; Ward and Crews 1998). This restriction does not function at the level of Sim, Trh, and Tgo nuclear localization, since ectopic expression experiments indicate that the proteins are localized to nuclei in all cell types examined.

Additional factors besides the CME are necessary for transcriptional activation by Trh. Mutational analysis of the *btl* gene has shown that the CME and adjacent sequences are both required for tracheal transcription (Ohshiro and Saigo 1997). The *rhomboid* gene is expressed in both CNS midline and tracheal cells, and an 0.7-kb fragment containing 4 CMEs is expressed in both tissues (Ip et al. 1992; S.T. Crews, unpubl.). However, an 0.3-kb subfragment containing 2 CMEs is strongly expressed in the midline, but is greatly reduced in the trachea (Ip et al. 1992), indicating the presence of elements required for tracheal expression distinct from those required for midline expression. Ectopic expression studies also suggest the existence of tracheal-specific elements distinct from the CME (Zelzer et al. 1997). In a model consistent with existing data (Zelzer et al. 1997), it is proposed that midline-specific target genes cannot be activated by Trh::Tgo because they lack tracheal-specific control elements in addition to the CME. Presumably, tracheal-specific target genes cannot be activated by Sim::Tgo in the midline because of the existence of positive or negative control elements in addition to the CME.

Elegant work on bHLH proteins that control myogenesis and neurogenesis have demonstrated that specific basic region residues are necessary for transcriptional specificity. In the case of the vertebrate myogenic bHLH proteins, including MyoD and Myogenin, it has been shown that two adjacent basic region residues are required for muscle-specific transcription (Davis et al. 1990; Brennan et al. 1991, Davis and Weintraub 1992). Biochemical experiments have shown that these residues are required for interaction of the MEF2 coregulatory MADS-box protein with the MyoD::E12 bHLH heterodimer (Molkentin et al. 1995). Mutational and chimeric protein studies of the *Drosophila* neurogenic bHLH proteins Atonal and Scute also suggest that residues within their basic regions are involved in the ability of these proteins to activate transcription in different classes of nerve cells (Chien et al. 1996). Similar experi-

ments carried out with Sim and Trh indicate that transcriptional specificity resides not within the basic region, but within the PAS domain (Zelzer et al. 1997). It was proposed that the PAS domain mediates interactions with the additional factors hypothesized to impart tissue specificity.

Genetic experiments indicate that *sim* can act as a midline repressor, as well as activator. The biochemical mechanism of midline repression by *sim* is unknown. It could involve binding of Sim::Tgo heterodimers to CMEs on target genes, in which case, the presence of adjacent corepressor sites would dictate repression instead of activation. This mechanism is analogous to how Dorsal can both activate transcription ventrally in the blastoderm embryo and repress ventral transcription in combination with sites of corepression (Jiang et al. 1993; Kirov et al. 1993; Huang et al. 1995). An alternative mechanism postulates Sim disrupting a positively acting transcription complex. If so, it is unlikely to be inhibiting a bHLH–PAS::Tgo heterodimer since cellular studies show Tgo to be cytoplasmic (and presumably transcriptionally inert) in the lateral CNS. Answers to questions of transcriptional specificity and mode of action await a concerted analysis using germline transformation, biochemistry, and genetic approaches.

Evolutionary conservation and functional diversity of the bHLH–PAS regulatory cassette

bHLH–PAS proteins mediate a wide variety of biological processes, which raises two issues. (1) Does the PAS domain carry out related biochemical functions in these disparate developmental and physiological events? (2) Is there a common origin to these biological events? The PAS domain clearly represents a polyfunctional interaction domain. Its large size allows a variety of interactions facilitating complex regulation of protein function. There are common functions to PAS domains: most bHLH–PAS proteins require the PAS domain for interaction with Arnt and Hsp90. In contrast, the relative lack of sequence conservation within the PAS domain suggests that different PAS proteins can mediate distinct molecular interactions. For example, Per interacts through its PAS domain with Timeless (Tim), a non-PAS protein (Gekakis et al. 1995). Yet, other bHLH–PAS proteins do not interact with Tim (G. Nystrom and S.T. Crews, unpubl.). Ahr interacts with halogenated aromatic hydrocarbons such as TCDD (dioxin) through its PAS domain to control nuclear localization, yet Sim, Per, and Arnt do not bind TCDD (Swanson and Bradfield 1993; Coumailleau et al. 1995). In all cases, the bHLH–PAS protein PAS domain mediates protein–protein interactions. These can be regulated by interactions with other molecules or unregulated.

Is there a functional connection between the different developmental and physiological events governed by bHLH–PAS proteins? There are some interesting similarities. Two of the basic biological processes that bHLH–PAS proteins participate in are biological rhythms and response to oxygen levels. Controlling gene expression in response to the circadian light/dark cycle is widespread across phylogeny. The identification of related PAS proteins implicated in rhythms between insects and mammals (King et al. 1997; Z.S. Sun et al. 1997; Tei et al. 1997) indicates that the mechanism of circadian regulation is evolutionarily well conserved. More surprising is the discovery that fungal PAS proteins mediate light-controlled rhythmic behavior (Linden and Macino 1997), suggesting an even stronger association between the PAS domain and regulation of rhythms. Conceivably, this could be a primordial function of bHLH–PAS proteins.

Physiological regulation of oxygen responsiveness is another basic organismal function necessary since the origins of the oxygen-rich environment 1.4 billion years ago (Bunn and Poyton 1996). bHLH–PAS proteins control both developmental and physiological aspects of oxygen delivery. HIF and related proteins control the response to oxygen levels, including finer aspects of vascular branching. Development of the respiratory system including the formation of tracheal tubules is controlled by *trh*. These genes could be specializations of a more primitive respiratory system regulatory protein. More speculative is the possible relationship between CNS midline and tracheal cell development. Although the CNS and trachea have different functions, numerous genes are utilized in the development of both lineages, and both tissues are ectodermal derivatives in which bHLH–PAS::Tgo heterodimers are activated in undifferentiated cells to form their respective tissues. It is possible that the CNS midline cells, which comprise a tissue distinct from the lateral CNS, and the trachea may have a common evolutionary origin. Although uncovering ancestral relationships from extant creatures can be problematic, hopefully functional analysis of PAS proteins from different organisms may shed light on these issues.

Comparative analysis of bHLH–PAS gene functions (Figs. 2 and 7) indicates that they constitute an evolutionarily-conserved regulatory gene system (Sonnenfeld et al. 1997). Organisms as diverse as *Caenorhabditis elegans, Drosophila,* and mammals have Arnt proteins that can form heterodimers with a variety of Sim/Ahr-related bHLH–PAS proteins. These heterodimers bind a sequence element related to the XRE (core GCGTG) or CME (core ACGTG). From this original regulatory cassette has emerged gene combinations that control developmental processes including neurogenesis and tracheal formation, and physiological processes including toxin metabolism, the response to oxygen deprivation and probably circadian rhythms (Fig. 8). Analysis of Ahr, Sim, and Trh have established that there are at least two fundamental modes of bHLH–PAS protein function. Ahr is a broadly distributed protein whose function is controlled by ligand dependent nuclear transport. In contrast, the Sim and Trh developmental regulators are not controlled at the level of nuclear transport but achieve specificity of function by virtue of their restricted expression. It will be interesting to see whether other bHLH–PAS proteins fit these modes or are regulated in novel ways.

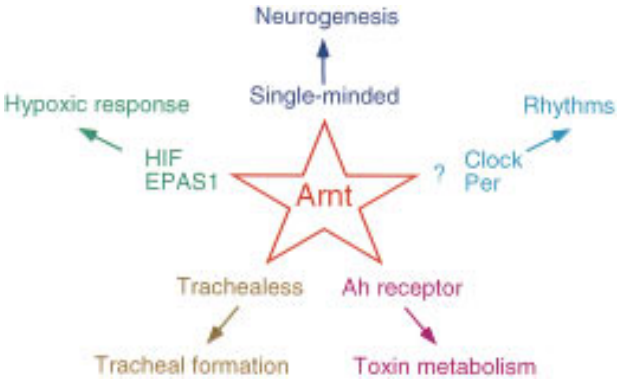

Figure 8. Array of Arnt interactions. The diagram shows the different proteins that Arnt/Tgo interacts with and the developmental and physiological processes these protein complexes control. The in vivo relationship between Arnt, Clock, and Per remains speculative.

Acknowledgments

I thank the scientists in my laboratory, Wei Chen, Robert Franks, Song Hu, Yumi Kasai, Sang Hee Kim, Josette Lewis, Beverly Matthews, Jack Mosher, John Nambu, Jay Nystrom, Margaret Sonnenfeld, Stephanie Stahl, Mary Ward, and Keith Wharton, who have worked on bHLH–PAS proteins and midline development for many stimulating discussions and collaborations. My initial interest in this area began with a productive and enjoyable collaboration with John Thomas and Corey Goodman. I also thank Chen-Ming Fan, Oliver Hankinson, Michael Levine, Lorenz Poellinger, Michael Rosbash, and Greg Semenza for useful discussions, Mark Peifer for critically reading the manuscript, and Mary Ward for generating the images shown in Figure 6. Work from my laboratory was supported by the National Institute of Child Health and Human Development, the National Science Foundation, and the Lucille P. Markey Charitable Trust.

Note added in proof

Analysis of the cell culture-derived Ahr D mutation indicates that the mutation is in the PAS domain and that the mutant protein is capable of ligand binding and dimerization to Arnt, but binds DNA weakly. This suggests that the PAS domain may directly influence DNA binding (W. Sun et al. 1997).

References

Abbott, B.D. and M.R. Probst. 1995. Developmental expression of two members of a new class of transcription factors: II. Expression of aryl hydrocarbon receptor nuclear translocator in the C57BL/6N mouse embryo. *Dev. Dynam.* **204:** 144–155.

Bate, C.M. and E.B. Grunewald. 1981. Embryogenesis of an insect nervous system II: A second class of neuron precursor cells and the origin of the intersegmental connectives. *J. Embryol. Exp. Morphol.* **61:** 317–330.

Bossing, T. and G.M. Technau. 1994. The fate of the CNS midline progenitors in *Drosophila* as revealed by a new method for single cell labelling. *Development* **120:** 1895–1906.

Brennan, T.J., T. Chakraborty, and E.N. Olson. 1991. Mutagenesis of the myogenin basic region identified an ancient protein motif critical for activation of myogenesis. *Proc. Natl. Acad. Sci.* **88:** 5675–5679.

Bunn, F.H. and R.O. Poyton. 1996. Oxygen sensing and molecular adaptation to hypoxia. *Physiol. Rev.* **3:** 839–885.

Burbach, K.M., A. Poland, and C.A. Bradfield. 1992. Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. *Proc. Natl. Acad. Sci.* **89:** 8185–8189.

Campos-Ortega, J.A. and V. Hartenstein. 1985. *The embryonic development of* Drosophila melanogaster. Springer Verlag, Berlin, Heidelberg, Germany.

Chang, Z., B.D. Price, S. Bockheim, M.J. Boedigheimer, R. Smith, and A. Laughon. 1993. Molecular and genetic characterization of the *Drosophila tartan* gene. *Dev. Biol.* **160:** 315–332.

Chen, H., R. Chrast, C. Rossier, A. Gos, S.E. Antonarakis, J. Kudoh, A. Yamaki, N. Shindoh, H. Maeda, S. Minoshima, and N. Shimizu. 1995. Single-minded and Down syndrome. *Nature Genet.* **10:** 9–10.

Chien, C-T., C-D. Hsiao, L.Y. Jan, and Y.N. Jan. 1996. Neuronal type information encoded in the basic-helix-loop-helix domain of proneural genes. *Proc. Natl. Acad. Sci.* **93:** 13239–13244.

Chrast, R., H.S. Scott, H. Chen, J. Kudoh, C. Roisser, S. Minoshima, Y. Wang, N. Shimizu, and S.E. Antonarakis. 1997. Cloning of two human homologs of the *Drosophila single-minded* gene SIM1 on Chromosome 6q and SIM2 on 21q within the Down Syndrome chromosomal region. *Genome Res.* **7:** 615–624.

Coumailleau, P., L. Poellinger, J.-A. Gustafsson, and M.L. Whitelaw. 1995. Definition of a minimal domain of the dioxin receptor that is associated with hsp90 and maintains wild type ligand binding affinity and specificity. *J. Biol. Chem.* **270:** 25291–25300.

Crews, S., R. Franks, S. Hu, B. Matthews, and J. Nambu. 1992. The Drosophila *single-minded* gene and the molecular genetics of CNS midline development. *J. Exp. Zool.* **261:** 234–244.

Crews, S.T., J.B. Thomas, and C.S. Goodman. 1988. The Drosophila *single-minded* gene encodes a nuclear protein with sequence similarity to the *per* gene product. *Cell* **52:** 143–151.

Dahmane, N., G. Charron, C. Lopes, M.L. Yaspo, C. Maunoury, L. Decorte, P.M. Sinet, B. Bloch, and J.M. Delabar. 1995. Down syndrome-critical region contains a gene homologous to *Drosophila sim* expressed during rat and human central nervous system development. *Proc. Natl. Acad. Sci.* **92:** 9191–9195.

Davis, R.L. and H. Weintraub. 1992. Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. *Science* **256:** 1027–1030.

Davis, R.L., P.F. Cheng, A.B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60:** 733–746.

Dolwick, K.M., H.I. Swanson, and C.A. Bradfield. 1993. In vitro analysis of Ah receptor domains involved in ligand-activated DNA recognition. *Proc. Natl. Acad. Sci.* **90:** 8566–8570.

Eguchi, H., T. Ikuta, T. Tachibana, Y. Yoneda, and K. Kawajiri. 1997. A nuclear localization signal of human aryl hydrocarbon receptor nuclear translocator/hypoxia-inducible factor 1β is a novel bipartite type recognized by the two components of nuclear pore-targeting complex. *J. Biol. Chem.* **272:** 17640–17647.

Ema, M., M. Morita, S. Ikawa, M. Tanaka, Y. Matsuda, O. Gotoh, Y. Saijoh, H. Fujii, H. Hamada, and Y. Fujii-Kuriyama. 1997a. Two new members of murine *Sim* gene family

are transcriptional repressors and show different expression patterns during mouse embryogenesis. *Mol. Cell. Biol.* **16:** 5865–5875.

Ema, M., K. Sogawa, N. Watanabe, Y. Chujoh, N. Matsushita, O. Gotoh, Y. Funae, and Y. Fujii-Kuriyama. 1992. cDNA cloning and structure of mouse putative Ah receptor. *Biochem. Biophys. Res. Commun.* **184:** 246–253.

Ema, M., S. Taya, N. Yokotani, K. Sogawa, Y. Matsuda, and Y. Fujii-Kuriyama. 1997b. A novel bHLH–PAS factor with close sequence similarity to hypoxia-inducible factor 1α regulates the *VEGF* expression and is potentially involved in lung and vascular development. *Proc. Natl. Acad. Sci.* **94:** 4273–4278.

Fan, C.M., E. Kuwana, A. Bulfone, C.F. Fletcher, N.G. Copeland, N.A. Jenkins, S. Crews, S. Martinez, L. Puelles, J.L.R. Rubenstein, and M. Tessier-Lavigne. 1996. Expression patterns of two murine homologs of *Drosophila single-minded* suggest possible roles in embryonic patterning and in the pathogenesis of Down syndrome. *Cell Mol. Neurobiol.* **7:** 1–16.

Firth, J.D., B.L. Ebert, C.W. Pugh, and P.J. Ratcliffe. 1994. Oxygen-regulated control elements in the phosphoglycerate kinase 1 and lactate dehydrogenase A genes: Similarities with the erythropoietin 3′ enhancer. *Proc. Natl. Acad. Sci.* **91:** 6496–6500.

Franks, R.G. and S.T. Crews. 1994. Transcriptional activation domains of the Single-minded bHLH protein are required for CNS midline cell development. *Mech. Dev.* **45:** 269–277.

Gekakis, N., L. Saez, A.M. Delahaye-Brown, M. Myers, A. Sehgal, M.W. Young, and C.J. Weitz. 1995. Isolation of *timeless* by PER protein interaction: Defective interaction between *timeless* protein and long-period mutant PER^L. *Science* **270:** 811–815.

Golembo, M., E. Rax, and B.Z. Shilo. 1996. The *Drosophila* embryonic midline is the site of Spitz processing, and induces activation of the EGF receptor in the ventral ectoderm. *Development* **122:** 3363–3370.

Goodman, C.S. 1982. Embryonic development of identified neurons in the Grasshopper. In *Neuronal development* (ed. N.C. Spitzer), Plenum Publishing Corporation, New York, NY.

———. 1996. Mechanisms and molecules that control growth cone guidance. *Annu. Rev. Neurosci.* **19:** 341–377.

Goodman, C.S. and C.Q. Doe. 1993. Embryonic development of the *Drosophila* central nervous system. In *The development of* Drosophila melanogaster. (ed. M. Bate and A. Martinez-Arias), pp. 1131–1206. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Guillemin, K. and M.A. Krasnow. 1997. The Hypoxic Response: Huffing and HIFing. *Cell* **89:** 9–12.

Hankinson, O. 1995. The Aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol.* **35:** 307–340.

Harris, R., L.M. Sabatelli, and M.A. Seeger. 1996. Guidance cues at the Drosophila CNS midline: Identification and characterization of two Drosophila Netrin/UNC6 homologs. *Neuron* **17:** 217–228.

Hoffman, E.C., H. Reyes, F. Chu, F. Sander, L.H. Conley, B.A. Brooks, and O. Hankinson. 1991. Cloning of a subunit of the DNA-binding form of the Ah (dioxin) receptor. *Science* **252:** 954–958.

Hogenesch, J.B., W.K. Chan, V.H. Kackiw, R.C. Brown, Y.-Z. Gu, M. Pray-Grant, G.H. Perdew, and C.A. Bradfield. 1997. Characterization of a subset of the basic-helix-loop-helix-PAS superfamily that interacts with components of the dioxin signaling pathway. *J. Biol. Chem.* **272:** 8581–8593.

Huang, J.-D., T. Dubnicoff, G.-J. Liaw, Y. Bai, S. Valentine, J. Shirokawa, J.A. Lengyel, and A.J. Courey. 1995. Binding sites for transcription factor NTF-1/Elf-1 contribute to the ventral repression of *decapentaplegic*. *Genes & Dev.* **9:** 3177–3189.

Huang, Z.J., I. Edery, and M. Rosbash. 1993. PAS is a dimerization domain common to *Drosophila* period and several transcription factors. *Nature* **364:** 259–262.

Ip, Y.T., R.E. Park, D. Kosman, E. Bier, and M. Levine. 1992. The *dorsal* gradient morphogen regulates stripes of *rhomboid* expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes & Dev.* **6:** 1728–1739.

Isaac, D.D. and D.J. Andrew. 1996. Tubulogenesis in *Drosophila:* A requirement for the *trachealess* gene product. *Genes & Dev.* **10:** 103–117.

Jan, Y.N. and L.Y. Jan. 1993. HLH proteins, fly neurogensis, and vertebrate myogenesis. *Cell* **75:** 827–830.

Jacobs, J.R. and C.S. Goodman. 1989. Embryonic development of axon pathways in the Drosophila CNS. I. A glial scaffold appears before the first growth cones. *J. Neurosci.* **9:** 2402–2411.

Jain, S., K.M. Dolwick, J.V. Schmidt, and C.A. Bradfield. 1994. Potent transactivation domains of the Ah receptor and the Ah receptor nuclear translocator map to their carboxyl termini. *J. Biol. Chem.* **269:** 31518–31524.

Jiang, J., H. Cai, Q. Zhou, and M. Levine. 1993. Conversion of a *dorsal*-dependent silencer into an enhancer: Evidence for *dorsal* corepressors. *EMBO J.* **12:** 3201–3209.

Jiang, J. and M. Levine. 1993. Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the Drosal gradient morphogen. *Cell* **72:** 741–752.

Kasai, Y., J.R. Nambu, P.M. Lieberman, and S.T. Crews. 1992. Dorsal-ventral patterning in *Drosophila:* DNA binding of snail protein to the *single-minded* gene. *Proc. Natl. Acad. Sci.* **89:** 3414–3418.

Kim, S.H. and S.T. Crews. 1993. Influence of *Drosophila* ventral epidermal development by the CNS midline cells and *spitz* class genes. *Development* **118:** 893–901.

King, D.P., Y. Zhao, A.M. Sangoram, L.D. Wilsbacher, M. Tanaka, M.P. Antoch, T.D.L. Steeves, M.H. Vitaterna, J.M. Kornhauser, P.L. Lowrey, F.W. Turek, and J.S. Takahashi. 1997. Positional cloning of the mouse circadian *Clock* gene. *Cell* **89:** 641–653.

Kirov, N., L. Zhelnin, J. Shah, and C. Rushlow. 1993. Conversion of a silencer into an enhancer: Evidence for a co-repressor in *dorsal*-mediated repression in *Drosophila. EMBO J.* **12:** 3193–3199.

Klämbt, C., J.R. Jacobs, and C.S. Goodman. 1991. The midline of the Drosophila central nervous system: A model for the genetic analysis of cell fate, cell migration, and growth cone guidance. *Cell* **64:** 801–815.

Kolodziej, P.A., L.C. Timpe, K.J. Mitchell, S.R. Fried, C.S. Goodman, L.Y. Jan, and Y.N. Jan. 1996. *frazzled* encodes a Drosophila member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. *Cell* **87:** 197–204.

Kosman, D., Y.T. Ip, M. Levine, and K. Arora. 1991. Establishment of the mesoderm-neuroectoderm boundary in the Drosophila embryo. *Science* **254:** 118–122.

Lamb, B.T. and J.D. Gearhart. 1995. YAC transgenics and the study of genetics and human disease. *Curr. Opin. Genet. Dev.* **5:** 342–348.

Leptin, M. 1991. *twist* and *snail* as positive and negative regulators during *Drosophila* mesoderm development. *Genes & Dev.* **5:** 1568–1576.

Leptin, M. and B. Grunewald. 1990. Cell shape changes during gastrulation in *Drosophila. Development* **110:** 73–84.

Lewis, J.O. 1994. ''The *single-minded* gene of Drosophila and its role in development of the embryonic central nervous sys-

tem and larval muscles.'' Ph.D. Thesis, University of California, Los Angeles, CA.

Lewis, J.O. and S.T. Crews. 1994. Genetic analysis of the *Drosophila single-minded* gene reveals a CNS influence on muscle patterning. *Mech. Dev.* **48:** 81–91.

Li, H., L. Dong, and J.P. Whitlock. 1994. Transcriptional activation function of the mouse Ah receptor nuclear translocator. *J. Biol. Chem.* **269:** 28098–28105.

Linden, H. and G. Macino. 1997. White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. *EMBO J.* **16:** 98–109.

Littlewood, T.D. and G.I. Evan. 1995. Transcription factors 2: helix-loop-helix. *Protein Profile* **2:** 621–702.

Lüer, K., J. Urban, C. Klämbt, and G.M. Technau. 1997. Induction of identified mesodermal cells by CNS midline progenitors in *Drosophila*. *Development* **124:** 2681–2690.

Ma, Q. and J.J.P. Whitlock. 1997. A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2, 3, 7, 8-Tetrachlorodibenzo-p-dioxin. *J. Biol. Chem.* **272:** 8878–8884.

Manning, G. and M.A. Krasnow. 1993. Development of the *Drosophila* tracheal system. In *The development of* Drosophila melanogaster. (ed. M. Bate and A. Martinez Arias), pp. 609–685. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Martin-Bermudo, M.D., A. Carmena, and F. Jimenez. 1995. Neurogenic genes control gene expression at the transcriptional level in early neurogenesis and in mesectoderm specification. *Development* **121:** 219–224.

Mayer, U. and C. Nüsslein-Volhard. 1988. A Group of genes required for pattern formation in the ventral ectoderm of the *Drosophila* embryo. *Genes* & *Dev.* **2:** 1496–1511.

McGuire, J., P. Coumailleau, M.L. Whitelaw, J.-A. Gustafsson, and L. Poellinger. 1996. The basic helix-loop-helix/PAS factor Sim is associated with hsp90. *J. Biol. Chem.* **52:** 31353–31357.

Mellerick, D.M. and M. Nirenberg. 1995. Dorsal-ventral patterning genes restrict NK-2 homeobox gene expression to the ventral half of the central nervous system of *Drosophila* embryos. *Dev. Biol.* **171:** 306–316.

Menne, T.V. and C. Klämbt. 1994. The formation of commissures in the *Drosophila* CNS depends on the midline cells and on the *Notch* gene. *Development* **120:** 123–133.

Menne, T.V., K. Lüer, G.M. Technau, and C. Klämbt. 1997. CNS midline cells in *Drosophila* induce the differentiation of lateral neural cells. *Development* **124:** 4949–4958.

Michaud, J. and C.M. Fan. 1997. *Single-minded*–two genes, three chromosomes. *Genome Res.* **7:** 569–571.

Mitchell, K.J., J.L. Doyle, T. Serafini, T. Kennedy, M. Tessier-Lavigne, C.S. Goodman, and B.J. Dickson. 1996. Genetic analysis of *Netrin* genes in Drosophila: Netrins control guidance of commissural axons and peripheral motor axons. *Neuron* **17:** 203–215.

Moffett, P., M. Reece, and J. Pelletier. 1997. The murine *Sim-2* gene product inhibits transcription by active repression and functional interference. *Mol. Cell. Biol.* **17:** 4933–4947.

Molkentin, J.D., B.L. Black, J.F. Martin, and E.N. Olson. 1995. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83:** 1125–1136.

Muenke, M., L.J. Bone, H.F. Mitchell, I. Hart, K. Walton, K. Hall-Johnson, E.F. Ippel, J. Dietz-Band, K. Kvaloy, C.-F. Fan, M. Tessier-Lavigne, and D. Patterson. 1995. Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of SIM2 as a candidate gene for holoprosencephaly, and mapping of SIM2 to a region of chromosome 21 important for Down Syndrome. *Am. J. Hum. Genet.* **57:** 1074–1079.

Muralidhar, M.G., C.A. Callahan, and J.B. Thomas. 1993. *Single-minded* regulation of genes in the embryonic midline of the *Drosophila* central nervous system. *Mech. Dev.* **41:** 129–138.

Nagao, M., B.L. Ebert, P.J. Ratcliffe, and C.W. Pugh. 1996. *Drosophila melanogaster* SL2 cells contain a hypoxically inducible DNA binding complex which recognises mammalian HIF-1 binding sites. *FEBS Lett.* **387:** 161–166.

Nambu, J.R., R.G. Franks, S. Hu, and S.T. Crews. 1990. The *single-minded* gene of Drosophila is required for the expression of genes important for the development of CNS midline cells. *Cell* **63:** 63–75.

Nambu, J.R., J.L. Lewis, K.A. Wharton, and S.T. Crews. 1991. The Drosophila *single-minded* gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* **67:** 1157–1167.

Nambu, J.R., J.O. Lewis, and S.T. Crews. 1993. The development and function of the *Drosophila* CNS midline cells. *Comp. Biochem. Physiol.* **104A:** 399–409.

Nambu, J.R., W. Chen, S. Hu, and S.T. Crews. 1996. The *Drosophila melanogaster similar bHLH-PAS* gene encodes a protein related to human Hypoxia-inducible factor 1α and *Drosophila* Single-minded. *Gene* **172:** 249–254.

Ohshiro, T. and K. Saigo. 1997. Transcriptional regulation of *breathless* FGF receptor gene by binding of TRACHEALESS/dARNT heterodimers to three central midline elements in *Drosophila* developing trachea. *Development* **124:** 3975–3986.

Pollenz, R.S., C.A. Sattler, and A. Poland. 1994. The aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein show distinct subcellular localizations in Hepa 1c1c7 cells by immunofluorescence microscopy. *Mol. Pharmacol.* **45:** 428–438.

Probst, M.R., C.-M. Fan, M. Tessier-Lavigne, and O. Hankinson. 1997. Two murine homologs of the *Drosophila* Single-minded protein that interact with the mouse aryl hydrocarbon receptor nuclear translocator protein. *J. Biol. Chem.* **272:** 4451–4457.

Rao, Y., H. Vaessin, L.Y. Jan, and Y.-N. Jan. 1991. Neuroectoderm in *Drosophila* embryos is dependent on the mesoderm for positioning but not for formation. *Genes* & *Dev.* **5:** 1577–1588.

Raz, E. and B.-Z. Shilo. 1992. Dissection of the *faint little ball* (*flb*) phenotype: Determination of the development of the *Drosophila* central nervous system by early interactions in the ectoderm. *Development* **114:** 113–123.

Reeves, R.H., N.G. Irving, T.H. Moran, A. Wohn, C. Kitt, S.S. Sisodia, C. Schmidt, R.T. Bronson, and M.T. Davisson. 1995. A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nature Genet.* **11:** 177–184.

Roth, S., D. Stein, and C. Nüsslein-Volhard. 1989. A gradient of nuclear localization of the *dorsal* protein determines dorsoventral pattern in the Drosophila embryo. *Cell* **59:** 1189–1202.

Rowlands, J.C. and J.-A. Gustafsson. 1997. Aryl hydrocarbon receptor-mediated signal transduction. *Crit. Rev. Toxicol.* **27:** 109–134.

Rusch, J. and M. Levine. 1996. Threshold responses to the dorsal regulatory gradient and the subdivision of primary tissue territories in the Drosophila embryo. *Curr. Opin. Genet. Dev.* **6:** 416–423.

Rushlow, C., K. Han, J.L. Manley, and M. Levine. 1989. The graded distribution of the *dorsal* morphogen is initiated by

selective nuclear transport in Drosophila. *Cell* **59:** 1165–1177.

Rushlow, C.A. and K. Arora. 1990. Dorsal-ventral polarity and pattern formation in Drosophila. *Semin. Cell Biol.* **1:** 137–149.

Seeger, M.A., G. Tear, D. Ferres-Marco, and C.S. Goodman. 1993. Mutations affecting growth cone guidance in Drosophila: Genes necessary for guidance towards or away from the midline. *Neuron* **10:** 409–426.

Semenza, G.L., P.H. Roth, H.-M. Fang, and G.L. Wang. 1994. Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *J. Biol. Chem.* **269:** 23757–23763.

Shue, G. and D.S. Kohtz. 1994. Structural and functional aspects of basic helix-loop-helix folding by Heat-shock Protein 90. *J. Biol. Chem.* **269:** 2707–2711.

Sonnenfeld, M., M. Ward, G. Nystrom, J. Mosher, S. Stahl, and S. Crews. 1997. The *Drosophila tango* gene encodes a bHLH–PAS protein that is orthologous to mammalian Arnt and controls CNS midline and tracheal development. *Development* **124:** 4583–4594.

Sonnenfeld, M.J. and J.R. Jacobs. 1994. Mesectodermal cell fate analysis in *Drosophila* midline mutants. *Mech. Dev.* **46:** 3–13.

Steward, R. 1989. Relocalization of the *dorsal* protein from the cytoplasm correlates with its function. *Cell* **59:** 1179–1188.

Sun, W., J. Zhang, and O. Hankinson. 1997. A mutation in the aryl hydrocarbon receptor (AHR) in a cultured mammalian cell line identifies a novel region of AHR that affects DNA binding. *J. Biol. Chem.* **50:** 31845–31854.

Sun, Z.S., U. Albrecht, O. Zhuchenko, J. Bailey, G. Eichele, and C.C. Lee. 1997. *RIGUI,* a putative mammalian ortholog of the Drosophila *period* gene. *Cell* **90:** 1003–1011.

Swanson, H.I. and C.A. Bradfield. 1993. The AH-receptor: Genetics, structure and function. *Pharmacogenetics* **3:** 213–230.

Swanson, H.I., W.K. Chan, and C.A. Bradfield. 1995. DNA binding specificities and pairing rules of the Ah receptor, ARNT, and SIM proteins. *J. Biol. Chem.* **280:** 26292–26302.

Tear, G., R. Harris, S. Sutaria, K. Kilomanski, C.S. Goodman, and M. Seeger. 1996. *commissureless* controls growth cone guidance across the CNS midline in Drosophila and encodes a novel membrane protein. *Neuron* **16:** 501–514.

Tei, H., H. Okamura, Y. Shigeyoshi, C. Fukuhara, R. Ozawa, M. Horise, and Y. Sakaki. 1997. Circadian oscillation of a mammalian homologue of the *Drosophila period* gene. *Nature* **389:** 512–516.

Thomas, J.B., S.T. Crews, and C.S. Goodman. 1988. Molecular genetics of the *single-minded* locus: A gene involved in the development of the Drosophila nervous system. *Cell* **52:** 133–141.

Tian, H., S.L. McKnight, and D.W. Russell. 1997. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. *Genes* & *Dev.* **11:** 72–82.

Wang, G.L., B.-H. Jaing, E. Rue, and G.L. Semenza. 1995. Hypoxia-inducible factor 1 is a basic helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc. Natl. Acad. Sci.* **92:** 5510–5514.

Ward, M.P. and S.T. Crews. 1998. Regulation of *Drosophila* bHLH–PAS protein cellular localization during embryogenesis. *Development* (in press).

Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lasser. 1991. The *myoD* gene family: Nodal point during specification of the muscle cell lineage. *Science* **251:** 761–766.

Wharton, K.A., Jr. and S.T. Crews. 1993. CNS midline enhancers of the *Drosophila slit* and *Toll* genes. *Mech. Dev.* **40:** 141–154.

Wharton, K.A., Jr., R.G. Franks, Y. Kasai, and S.T. Crews. 1994. Control of CNS midline transcription by asymmetric E-box elements: Similarity to xenobiotic responsive regulation. *Development* **120:** 3563–3569.

Wheeler, W.M. 1893. A contribution to insect morphology. *J. Morphol.* **8:** 1–160.

Whitlock, J.P., Jr., S.T. Okino, L. Dong, H.P. Ko, R. Clarke-Katzenberg, Q. Ma, and H. Li. 1996. Induction of cytochrome P450IA1: A model for analyzing mammalian gene transcription. *FASEB J.* **10:** 809–818.

Wigglesworth, V.B. 1954. Growth and regeneration in the tracheal system of an insect, *Rhodnius prolixus* (Hemiptera). *Q. J. Microsc. Sci.* **95:** 115–137.

Wilk, R., I. Weizman, L. Glazer, and B.-Z. Shilo. 1996. *trachealess* encodes a bHLH-PAS protein and is a master regulator gene in the *Drosophila* tracheal system. *Genes* & *Dev.* **10:** 93–102.

Xiao, H., L.A. Hrdlicka, and J.R. Nambu. 1996. Alternate function of the *single-minded* and *rhomboid* genes in development of the *Drosophila* ventral neuroectoderm. *Mech. Dev.* **58:** 65–74.

Younossi-Hartenstein, A. and V. Hartenstein. 1993. The role of the tracheae and musculature during pathfinding of *Drosophila* embryonic sensory axons. *Dev. Biol.* **158:** 430–437.

Zelzer, E., P. Wappner, and B.-Z. Shilo. 1997. The PAS domain confers target gene specificity of *Drosophila* bHLH/PAS proteins. *Genes* & *Dev.* **11:** 2079–2089.

Zhou, L., H. Xiao, and J.R. Nambu. 1997. CNS midline to mesoderm signaling in *Drosophila. Mech. Dev.* **67:** 59–68.