Chapter 4

DROSOPHILA bHLH-PAS DEVELOPMENTAL REGULATORY PROTEINS

Stephen T. Crews
The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

1. INTRODUCTION

Drosophila bHLH-PAS proteins play important roles in development and physiology. They can be divided into three groups: (i) Tgo and its bHLH-PAS dimerization partners, (ii) circadian rhythms, and (iii) hormone function. The Tgo dimerization group carries-out many developmental roles, and includes: Dysfusion (Dys), Similar (Sima), Single-minded (Sim), Spineless (Ss), and Trachealess (Trh). They are the subject of this chapter (the other *Drosophila* PAS genes are reviewed in the chapters by Hogenesch and Kay, Montell, and Wilson). Members of the Tgo dimerization group share a number of common features. Foremost, they form DNA-binding heterodimers with Tgo. They are well-conserved between nematodes, insects, and mammals. Each carries-out multiple developmental roles, and some have roles as master regulators of tissue formation. The tissues, cell types, and biological processes whose development and function are influenced by bHLH-PAS proteins are diverse. There is little redundancy or overlap of function, although, there is one example of a bHLH-PAS protein regulating levels of another.

2. TANGO: THE DIMERIZATION PARTNER

2.1 Structure and biochemistry

Tgo (1, 2) is the *Drosophila* ortholog of the Aryl hydrocarbon nuclear receptor (Arnt), the first vertebrate bHLH-PAS protein identified (3), and its close relative, Arnt2. The evolving realization that *Drosophila* had orthologs of mammalian Ahr and HIF-1α, which both use Arnt as a dimerization partner; the ability of *Drosophila* Sim to dimerize with Human Arnt; and the identification of a CNS midline enhancer element, whose sequence was consistent with binding to a Sim:Arnt heterodimer (4), strongly suggested that a fly *Arnt* existed. Use of human *Arnt* probes to screen *Drosophila* clone libraries (1, 2) were employed to identify *tgo*. The sequence structure of Tgo resembles Arnt, and both can dimerize with either insect or mammalian bHLH-PAS proteins (1), but there are differences. Vertebrate Arnt has a functional N-terminal nuclear localization sequence absent in Tgo (5), and Tgo has a Paired repeat near its C-terminus (1). The Paired repeat consists of alternating His-Pro residues and is found on a number of interesting transcription factors, including Bicoid and Paired (6). While its function in Tgo is unknown, interestingly, three *tgo* mutants have stop codons just before the Paired repeat.

Tgo forms DNA-binding heterodimers with its partners, and there is no evidence that it can function as a homodimer. However, if it does function as a homodimer or monomer, then it must be in an unconventional manner unrelated to DNA binding, since Tgo resides in the cytoplasm in the absence of a bHLH-PAS dimerization partner, and in the nucleus in the presence of a partner (7). Both cell culture transfection and in vivo experiments have indicated that Sim:Tgo, Trh:Tgo, and Sima:Tgo preferentially bind an ACGTG core sequence that is referred to as a CNS midline element (CME) (1, 2, 4, 8), and Ss:Tgo binds GCGTG (9). Tgo binds the GTG half-site and the partner binds the other AC or GC half-site. The recognition sequence of Dys:Tgo is unknown. The transient transfection results also indicate that each heterodimer (Dys:Tgo has not been tested) functions as a transcriptional activator.

2.2 Genetics

Given the striking phenotypes of *sim*, *trh*, and *ss* mutants, it was surprising that mutants in *tgo* had not been identified before it was cloned. Nevertheless, two approaches yielded *tgo* mutants. One utilized reverse genetics (1). In this scheme, a P-element transposon was mobilized into the

tgo gene creating a lethal mutation. Four EMS *tgo* mutants were then isolated that failed to complement the P-element mutant strain. Another approach screened for dominant enhancers of a weak *ss* mutant phenotype (9). In this latter method, it was anticipated that mutations in genes that function in the same developmental pathways as *ss* would enhance the weak *ss* phenotype. Three alleles of *tgo* were identified in this manner. These results also provided genetic evidence that Ss interacts with Tgo in vivo.

Mutations of *tgo* are embryonic lethal, and phenotypic analysis of mutant embryos revealed CNS midline and tracheal defects (1), consistent with a role as a partner of Sim in controlling CNS midline cell development and Trh in controlling tracheal development. However, all *tgo* mutants analyzed showed weaker embryonic defects than observed in *sim* and *trh* null mutants. Since the evidence is strong that Tgo is a required partner for Sim and Trh, and the *Drosophila* genome contains no other *tgo*-like genes, the weak phenotypes are likely due to the presence of maternally-contributed *tgo* RNA and the possibility that none of the *tgo* mutants are null.

Mutant strains of *ss* produce viable adult flies and show defects in adult bristle, antennal, and leg morphology (10). These tissues are all derived from imaginal discs that develop during larval growth and metamorphosis. Since *tgo* mutants are embryonic lethal, testing *tgo* mutants for *ss* defects required generating mosaic *tgo* flies (9). In this manner, embryos heterozygous for *tgo* are allowed to develop, the homozygous *tgo* mutant cells are induced during postembryonic development, and flies assayed for adult morphological defects. The three *tgo* mutants analyzed showed *ss*-like phenotypes, and animals with tgo^5 mutant tissue showed defects nearly as severe as *ss* null mutants. In summary, the genetic analysis of *tgo* is consistent with it being a dimerization partner for Sim, Ss, and Trh.

2.3 bHLH-PAS protein interactions and Tgo subcellular localization

Both *tgo* RNA and protein are found in all embryonic cells (1, 2, 7). In most cells, Tgo protein is present in the cytoplasm and excluded from the nucleus (7). However, in a number of cells Tgo protein is localized to nuclei. These cells include the CNS midline, trachea, salivary duct, sensory cells, and larval antennal primordia – all sites of Sim, Ss and Trh protein localization. This led to the idea that Tgo is cytoplasmic if no bHLH-PAS partner protein is present, but in the presence of a partner protein the two dimerize and translocate into the nucleus (Figure 1). This was confirmed by ectopically expressing *sim*, *ss*, or *trh* and showing that both the partner protein and Tgo colocalized to nuclei in those cells (7, 9). Since nuclear localization was observed in all cell types and developmental times

investigated, this indicated that dimerization and nuclear localization were not under developmental regulation. However, this does not imply that other factors are unimportant in dimerization and localization.

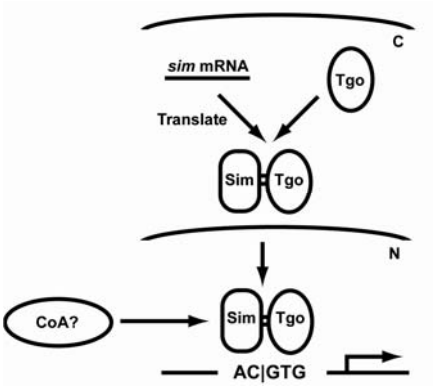

Figure 1. Model for Tgo interactions with partner bHLH-PAS proteins. Tgo is localized to the cytoplasm (C) in the absence of a partner bHLH-PAS protein. When a bHLH-PAS gene, such as *sim*, is expressed in the cell, the mRNA is translated into protein, dimerizes with Tgo, and the complex translocates into the nucleus (N), where it binds DNA and activates transcription. Sim:Tgo binds an ACGTG-containing binding site. Target gene expression likely also requires interactions between Sim:Tgo and transcriptional coactivators (CoA), although regulation by corepressors remains a possibility. This model holds for all partners of Tgo, although nuclear translocation of Trh:Tgo requires phosphorylation, and the appearance of Sima protein is due to hypoxia-induced inhibition of protein degradation rather than transcriptional control.

3. SINGLE-MINDED – MASTER REGULATOR OF CNS MIDLINE CELL DEVELOPMENT

The *sim* gene functions as a master regulator of CNS midline cell development. The lethal gene designated as *l(3)S8* was first reported in 1964 (11, 12). Later it was shown that *l(3)S8* mutations had a severely disorganized embryonic CNS (13). The gene was renamed "*single-minded*" because the two longitudinal axonal connectives that run along the length of the wild-type *Drosophila* CNS were now fused into a single connective in the mutant (Figure 2A, B). Further analysis indicated that *sim* mutant embryos were missing the cells that lie along the midline of the CNS (13). *sim* mutants were also isolated based on an absence of the cuticular ventral midline denticles (14), a defect later shown to be due to an absence of a midline-to-epidermis signaling pathway (15, 16). The region around *sim* was genetically and molecularly well-characterized. This facilitated the

cloning of *sim*, and its identification was based on its prominent expression in the CNS midline cells (13) (Figure 2C). The *sim* gene is also expressed in a number of cells besides the midline cells.

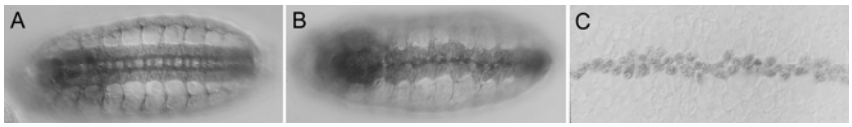

Figure 2. sim phenotypes and expression. Anterior is to the left in all panels. (A) Wild-type *Drosophila* embryo stained with an antibody that reacts with all nerve cells and axons. The axon scaffold consists of two longitudinal connectives running along the A/P axis and two commissures/ganglion that cross the midline. (B) *sim* mutant embryo showing a collapsed axon phenotype with the appearance of a single axon bundle running along the A/P axis, instead of the two characteristic longitudinal connectives. (C) Wild type embryo stained with anti-Sim showing the appearance of Sim in CNS midline cell nuclei. Magnification is higher in (C) than in (A, B).

3.1 *Drosophila* CNS midline cells

3.1.1 Development of the CNS midline cells

The *Drosophila* embryonic CNS consists of a brain and ventral nerve cord. The nerve cord is comprised of segmentally-repeated ganglia. Each ganglion has ~1000 neurons and glia. The ganglion is bilaterally symmetrical, and has a distinct set of cells at the midline. The mature embryonic *Drosophila* midline cells consist of ~15 neurons and 2-3 glia (13, 17, 18) (Figure 3). The midline neurons are: (i) 2 midline precursor 1 (MP1) interneurons, (ii) 2 MP3 interneurons, and (iii) ~10 progeny of the median neuroblast that include interneurons and neurosecretory motoneurons. The midline glia enwrap the two commissural axon bundles that cross the midline, and also participate in signaling pathways that control a variety of developmental processes.

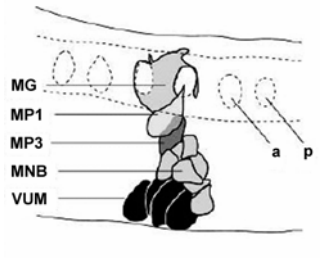

Figure 3. Mature *Drosophila* CNS midline cells. The different CNS midline cells of a mature embryo are labeled to the left (see text). The anterior commissure (a) and posterior commissures (p) that cross the midline are shown enwrapped by the midline glia (MG). The midline neurons lie below the MG. Adapted from Bossing and Technau (17).

The mature CNS midline cells are derived from a small set of midline precursor cells, which are often referred to as the mesectoderm (reviewed in (19, 20). The mesectoderm consists of two single cell-wide strips of cells that run along the anterior/posterior axis of the blastoderm embryo adjacent to the mesoderm. As gastrulation takes place, the two mesectodermal stripes join together at the ventral midline. Initially, there are ~8 midline precursor cells/segment. These cells undergo a synchronous cell division to give rise to 16 precursors/segment. This is followed by changes in cellular morphology in which the midline nuclei migrate inward and the cell maintains a cytoplasmic extension to the embryo surface. The midline precursors go on to divide and differentiate into mature midline neurons and glia, and the specific fates of individual midline cells are determined, at least in part, by the functions of the segment polarity genes, including *engrailed*, *hedgehog*, *patched*, and *wingless* (21). Within the midline glial lineage, apoptosis acts to establish the final number of mature glia in each segment (22). Given their simplicity and extensive study, the *Drosophila* CNS midline cells have the potential to be one of the premier neurogenomic systems for studying how neural precursor cells develop into a diverse set of motoneurons, interneurons and glial cell types. Over 200 genes have been identified that are expressed or function in the CNS midline cells.

3.1.2 Control of development and differentiation by CNS midline cell signaling

The midline cells constitute an important signaling center during *Drosophila* embryogenesis that influences the differentiation and migration of a number of neighboring cell types. CNS midline-dependent cellular processes include the: (i) formation and patterning of the underlying ventral epidermis and salivary tissue, (ii) proliferation of brain neuroblasts and

formation of axonal connections between brain and ventral nerve cord, (iii) control of midline axon crossing and organization, (iv) muscle cell migration, and (v) formation of nearby mesodermal and CNS cells. These processes are mediated in most cases by morphogens secreted by the midline cells.

Spitz signaling control of cell fate and proliferation. One such morphogen is Spitz, a TGF-α-like protein that acts through the *Drosophila* Epidermal growth factor receptor (Egfr). Spitz is secreted by the midline cells and patterns the underlying ventral epidermis in a graded manner (15, 16). The Spitz signaling pathway is also involved in directing the ventral-most salivary primordia to become duct cells (see 4.6.1.2 below) (23), and for the formation of the mesodermally-derived dorsal median cells, an unusual set of cells that lie above the CNS (24, 25). Recent work has shown that Sim and Spitz are required for neuroblast proliferation and subsequent formation of the midbrain (26). What is particularly noteworthy about this observation is that this signal emanates from Sim+ cells in the foregut, not the CNS. This is anatomically consistent, since the Sim+ foregut cells constitute a "midline" within the midbrain primordia (i.e. the foregut passes through the embryonic brain), and *sim* blastoderm expression in the ectodermal foregut cells is an anterior extension of *sim* expression in the mesectoderm. Nevertheless, it is striking that this function is carried-out in two distinct tissue types. This foregut-derived signaling pathway also likely mediates the formation of axonal connections between the embryonic brain and nerve cord (27). It has been hypothesized that the ancestral brain of arthropods was located dorsal to the foregut and connected to the ventral nerve cord by axonal connectives without an intervening midbrain. Page (26) has proposed that an evolutionary extension of Sim/Spitz signaling from the CNS midline cells into the foregut cells promoted the formation of the now existent midbrain.

Midline control of axonogenesis. The CNS midline cells also play an important role in formation of the commissural axon bundles that connect the lateral halves of the CNS. Most nerve cells send their axons across the midline via commissural bundles where they join distinct longitudinal pathways that run along the longitudinal axis of the CNS (28). The Netrins are secreted by the midline and attract axons to the midline. Conversely, the Slit protein is also secreted by the midline and acts to both repel axons that do not cross and to inhibit recrossing of commissural axons. In addition, Slit acts as a morphogen to direct axons into distinct longitudinal bundles.

Slit and FGF control of cell migration and fate. There is a subpopulation of somatic muscle cells that originate above the CNS. Midline-derived Slit acts as a repellent to guide these muscle cell precursors outwards from the interior of the embryo to the body wall (29, 30). Both Slit

and Fibroblast Growth Factor (FGF) pathways are involved in directing the fates and differentiation of some non-midline nervous system cells. *slit* is required for the formation of specific neurons from the ganglion mother cell (GMC) precursors (31). Normally, GMC-1 of the RP2/sib lineage divides asymmetrically to generate two distinct neurons: RP2 and RP2sib. In *sim* or *slit* mutants, GMC-1 fails to divide asymmetrically bout instead generates two identical RP2 neurons. The insect midline also has an FGF-like activity that influences the differentiation of serotonergic neurons upon crossing the midline (32). Finally, analysis of *sim* mutants indicated that 15% of the lateral CNS neurons were absent (33), and it was proposed that this effect is due to the influence of multiple midline-derived signaling pathways, consistent with the other reports described in this section. These results all point to the midline as an important developmental signaling center. Interestingly, this function is a role shared with the vertebrate floor plate, a group of specialized neuroepithelial cells, which resides at the ventral midline of the developing spinal cord.

3.2 Sim midline genetics

Examination of *sim* mutant embryos revealed that the collapsed CNS phenotype is due to a severe disruption in CNS midline cell development (13). Gastrulation is normal and the mesectodermal cells come together at the ventral midline. However, the midline precursor cells never properly form, and all subsequent developmental events fail to take place, including cell division, cell shape changes, and differentiation into midline neurons and glia (34). Ultimately, the cells die (22). Most, if not all, genes that are expressed in the CNS midline precursor cells require *sim* to initiate or maintain their expression (35). Conversely, genes that are expressed in the adjacent, lateral CNS cells, but not the midline cells, are expressed in the midline in *sim* mutants (36-38). Misexpression of *sim* throughout the neuroectoderm is able to convert the entire CNS to midline cells (34). This indicates that *sim* acts as a genetic switch to activate midline transcription and to repress midline expression of genes normally expressed in the adjacent lateral CNS. Presumably, the combination of these two functions directs a neuroectodermal cell to become CNS midline and not lateral CNS. The data are consistent with the idea that the *sim* gene acts as a master regulator of CNS midline cell development.

3.3 *sim* midline expression

The *sim* gene is prominently expressed in the CNS midline cells throughout development (13, 39). Initial expression is in midline precursors

(mesectoderm) at the cellular blastoderm stage just before gastrulation. It remains on in all midline precursors until they begin to differentiate into neurons and glia. At that time, *sim* is expressed at high levels in glia, but low levels in neurons. The biphasic expression of *sim*, initially in midline precursors and later in midline glia, is reflected in the organization of the gene. The *sim* gene consists of 8 exons spanning over 20.5 kb (40), and contains two promoters, an early promoter (P_E) and a late promoter (P_L) (34). P_E governs initial expression of *sim* in the mesectoderm and subsequent midline precursors. P_L directs *sim* transcription later in midline precursors and then in the midline glia. *sim* remains on in midline cells of the larval ventral nerve cord (41). Functionally, there are three major modes of *sim* expression: (i) initial activation in the mesectoderm, (ii) maintenance of expression in the midline precursors, and (iii) strong expression in midline glia and weak expression in midline neurons.

3.4 Flipping the switch: activating *sim* expression in the midline cells

The mesectoderm lies between the mesoderm and neuroectoderm along the blastoderm embryonic dorsal/ventral (D/V) axis. The key event in dictating whether a cell will become mesectoderm is activation of *sim* gene expression. Thus, specification of the CNS midline cells is essentially an issue of how D/V patterning genes activate *sim* expression in the two single-cell wide mesectodermal stripes. Since this D/V patterning process requires the specification of only a single cell diameter, it is not surprising that a complex set of developmental processes are required including: morphogenetic gradients, transcriptional activation and repression, combinatorial actions of transcription factors, and cell signaling (Figure 4A). The results of genetics and biochemical experiments indicate that the Dorsal and Twist transcription factors act on the *sim* promoter in conjunction with a bHLH (Daughterless:Scute) heterodimer and activate *sim* transcription broadly in the ventral region of the blastoderm (19, 40). The ventral boundary of *sim* expression is established by the Snail zinc finger protein, which is expressed at high levels in the mesoderm but is off in the mesectoderm. Sna represses *sim* transcription, thereby establishing a sharp ventral *sim* expression boundary between the mesoderm and mesectoderm (19, 42). The dorsal boundary of *sim* expression is established via the Suppressor of Hairless [Su(H)] transcription factor. Su(H) represses *sim* transcription in the neuroectoderm forming the dorsal *sim* expression on-off boundary (43). In addition, Notch signaling, positioned by Sna promotion of Notch-Delta endocytosis (44, 45), converts Su(H) from a repressor in the mesectoderm to a direct activator of *sim*. The result is that *sim* is expressed

at high levels in the mesectoderm, but not at all in adjacent mesodermal or neuroectodermal cells. All of these transcription factors act directly on the *sim* gene, and *sim* P_E has a dense and complex arrangement of transcription factor binding sites (Figure 4B).

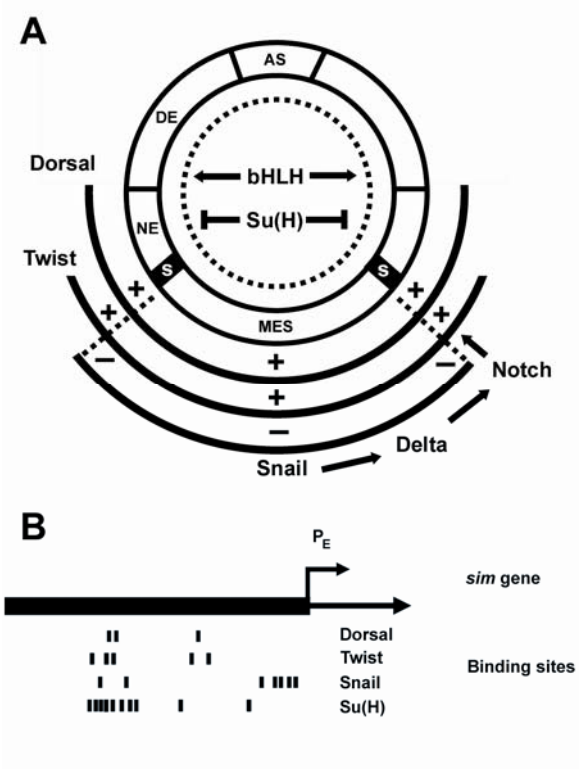

Figure 4. Activation of *sim* expression along the D/V axis of the blastoderm embryo. See text for details. (A) Schematic of cross-section of a blastoderm embryo showing the distribution of transcription factors that regulate initial sim expression. Ventral is at bottom. Filled box with white "s" represents mesectoderm – the site of initial *sim* expression. AS – amnioserosa; DE – dorsal ectoderm; NE – neuroectoderm; MES – mesoderm; "+" and arrows indicate positive regulation, and (-) and blocked lines indicate negative regulation of the *sim* gene. Dotted lines indicate that bHLH and Su(H) are expressed throughout the blastoderm. (B) The arrangement of transcription factor binding sites in the *sim* early regulatory region is shown below a schematic of the *sim* gene. Shown is 3.7 kb of DNA that flanks the *sim* early promoter (P_E).

3.5 Structure and biochemistry

Drosophila Sim belongs to a subfamily of bHLH-PAS proteins highly conserved between insects and vertebrates. Mammals have two *Sim* genes,

Sim1 and *Sim2* (see chapter by Fan). *Drosophila* Sim has four major regions, which, from N-terminus to C-terminus, are: (i) bHLH DNA domain, (ii) PAS-1 and PAS-2 domains, (iii) Ala-Ala-Gln repeats, and (iv) homopolymeric stretches (Figure 5). The basic region mediates DNA binding in combination with the Tgo basic region. Biochemically, Sim heterodimerizes with Tgo (1), and this interaction is dependent on HLH and PAS-1 domains (Nystrom and Crews, unpubl.). Sim does not homodimerize (1). Together, Sim:Tgo binds to CMEs, which contain a core ACGTG recognition sequence (1, 4). The PAS domain-containing region binds to the chaperone Hsp90 (46), and also interact with unidentified cofactors required for transcriptional specificity (8, 47). Closely following the PAS domains is a stretch of 10 Ala-Ala-Gln repeats. Despite its striking sequence structure, deletion of the Ala-Ala-Gln region does not affect the ability of Sim to function in transcriptional activation in vivo (48). The C-terminal region has at least three discrete transcriptional activation regions consisting of stretches rich in asparagine, glutamine, glycine, histidine, and serine (49).

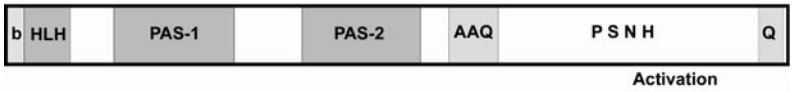

Figure 5. Sequence structure of the Sim protein. Shown are the bHLH, PAS-1, and PAS-2 domains, along with regions rich in various amino acids. These include a stretch of Ala-Ala-Gln (AAQ), Pro, Ser, His, Asn (PSNH), and Gln (Q). There are multiple transcriptional activation sequences in the C-terminal section of the protein.

3.6 Midline precursor gene expression – Sim activation, autoregulation, and repression

Activation. The critical role of *sim* in CNS midline cell development is indicated by the finding that genes that are normally expressed in the CNS midline cells require *sim* function for either the initiation or maintenance of their midline expression. The cis-regulatory regions of three genes, *Toll*, *rhomboid*, and *breathless*, have been studied in vivo using germline transformation and in vitro mutagenesis techniques. Analysis of the *Toll* gene identified 4 putative Sim:Tgo binding sites (CMEs) in an 0.9 kb fragment that drove high levels of midline precursor transcription (4). The CMEs were within a 662 bp region. Mutation of all 4 CMEs abolished midline transcription. The role of the CME was further demonstrated by multimerizing a 20 bp fragment containing *Toll* CME-4, and showing it could drive strong midline precursor transcription in vivo. Similar analysis

of *rhomboid* revealed two CMEs within 60 bp that are required for midline precursor transcription (50), and study of *breathless* revealed three essential CMEs within 144 bp (2). It is likely that every gene expressed in the midline precursor cells is directly regulated by Sim:Tgo heterodimers, and their regulatory regions contain multiple, clustered Sim:Tgo binding sites.

However, binding of Sim:Tgo is not sufficient for activation of midline precursor transcription. Ectopic expression of *sim* throughout the ectoderm reveals that activation of midline precursor-expressed genes can occur only in the ventral ectoderm, but not dorsal ectoderm (8, 34, 48). This occurs despite the observation that Sim:Tgo can enter nuclei and bind DNA in the dorsal ectoderm (7). Additional analysis of the *rhomboid* gene indicated that additional cis-regulatory sequences were required in addition to the CMEs for midline transcription (8). Most likely, midline precursor gene expression normally involves Sim:Tgo as well as either a ventral coactivator or a dorsal corepressor. Transgenic swap experiments have revealed that the unidentified factor(s) functions through the Sim PAS domains (8).

Autoregulation. Genetic and molecular studies demonstrated that *sim* is required for its own continued expression in midline precursors. Expression of a transgene containing the *sim* P_E regulatory region driving *lacZ* expression showed a more rapid reduction in levels in a *sim* mutant compared to wild-type (34). In addition, levels of *sim* RNA declined more rapidly in *sim* mutant embryos (34). Analysis of the *sim* P_E regulatory region revealed 4 Sim:Tgo binding sites within 560 bp, and mutation of these sites abolished *sim*-dependent midline precursor expression (4). This indicates that autoregulation is due to direct action of Sim:Tgo on the *sim* regulatory region. Midline precursor expression is also driven from the *sim* P_L (34). In *sim* mutant embryos, P_L-mediated expression is absent. Thus, DV patterning proteins initially activate *sim* in the mesectodermal cells, and then Sim maintains its own expression via a positive feedback loop acting on both promoters. Additional factors must be responsible for P_L transcription in midline glia and the absence of midline glial P_E expression.

Repression. Mesectodermal cells are initially fated to become *ventral nerve cord defective* (*vnd*)-positive neuroectodermal cells, but the expression of *sim* switches them to a midline fate [*vnd* is a key regulator of ventral neuroectodermal neural specification in much the way *sim* controls midline development (51, 52)]. *sim* represses midline expression of genes, including *tartan*, *wingless*, and *vnd*, that are expressed in the ventral neuroectoderm (36-38). The mechanism of *sim* repression was revealed in studies on the *vnd* gene (48). Three lines of evidence demonstrated that Sim is a pure transcriptional activator that represses indirectly. (i) Deletion of Sim:Tgo binding sites from the *vnd* regulatory region does not affect midline repression when tested in vivo. (ii) Analysis of mutant forms of Sim, tested

using an in vivo misexpression-repression assay, indicated that mutations that abolished the ability of Sim to activate transcription (e.g. removal of the DNA binding or activation domains) also abolished repression. (iii) Substitution of the Sim activation domain with the VP16 activation domain restored the ability of Sim to both activate and repress transcription in vivo. These results indicate that Sim represses midline transcription indirectly by activating transcription of distinct repressive factors. Complete understanding of the role of transcriptional repression in CNS midline cell development will require identification of these repression factors.

3.7 Midline glial regulation

The Sim protein is prominently expressed in midline glia, both in the embryo and larva. Since distinct, but overlapping, sets of genes are expressed in midline precursors and midline glia, does Sim:Tgo directly control midline glial expression, and do the two modes of regulation require the use of different regulatory cofactors? Midline glial expression is derived from the *sim* late promoter, P_L (34). The *slit* gene, which plays important roles in midline-directed axon guidance, is expressed in the midline glia, and was studied as a representative midline glial-expressed gene. Analysis of fragments of *slit* genomic DNA by germline transformation discovered a 380 bp fragment of *slit* from intron 1 that drove *lacZ* in the midline glia (53). This fragment has a single CME, and mutation of the CME results in a loss of midline glial expression (4). This indicates that Sim:Tgo directly regulates *slit* expression.

Two genes, *fish-hook* (*fish*; also called *Dichaete*) and *drifter* (*dfr*), both encode transcription factors expressed in midline glia. *fish* encodes a Sox HMG domain transcription factor and *dfr* encodes a POU-homeobox transcription factor. Mutations in either gene result in weak phenotypes in which expression of midline glial markers is relatively normal (47, 54). Although midline glia are present, they fail to migrate properly. However, when *fish dfr* double mutants were analyzed, midline glia fail to form and gene expression, including *slit-lacZ,* is greatly reduced. Biochemical experiments revealed that Fish binds Sim via the Sim PAS domain and Fish Sox domain. Dfr binds Fish via the Dfr POU domain (47). These results indicate that midline glial expression is dependent on a transcription factor complex that includes Tgo:Sim:Fish:Dfr. Neither Fish nor Dfr have been implicated in controlling midline precursor cell transcription. Thus, the ability of Sim:Tgo to regulate gene expression in both CNS midline precursor cells and midline glia is dependent, at least in part, on interactions of Sim:Tgo with different coregulatory proteins and distinct cis-regulatory sequences.

Sim is able to control the transcription of different gene sets in midline cells and other diverse cell types (also see below). Current results suggest that transcriptional specificity arises from the interaction of Sim:Tgo with different coregulators. The identification of these coregulatory proteins and biochemical study of how they interact with Sim (particularly the role of the PAS domain) are important areas of future research. In addition, it is of great interest to determine how Sim, which dictates general CNS midline identity, and the segment polarity proteins, which promote different midline neural and glial cell fates, interact at the molecular level to promote specific transcription patterns in individual midline cells.

3.8 Non-midline functions of Sim

3.8.1 Postembryonic brain function – control of locomotion

There are two sites of *sim* expression in the larval brain: (i) the lamina and medulla of the optic lobes, and (ii) clusters of neurons in the central complex (41). The optic lobes mediate processing of visual information that is received from the retina. In the 3rd instar larva, *sim* is expressed in most or all of the neurons of the lamina and many medullary neurons. *sim* expression is not in the optic lobe proliferative zones, but precedes neural differentiation. Analysis of flies heterozygous for a *sim* temperature sensitive mutant allele ($sim^{J1\text{-}47}$) and a *sim* null mutant (sim^{H9}) reared at the permissive temperature (17°C) revealed defects in axonal organization (41). At the level of the inner optic chiasm in which the medullary axons connect with the lobula and lobular plate (additional sites of higher order visual processing), there were axonal fibers entering the lobula from the medulla via abnormal paths. These results suggest that *sim* may be controlling aspects of medullary neuron axon guidance.

The larval brain central complex expression of *sim* is in three paired clusters of neurons that lie on either side of the midline. One function of the central complex is the coordination of movement (55). Analysis of $sim^{J1\text{-}47}/sim^{H9}$ flies showed adult behavioral and morphological brain defects consistent with the central complex expression (41). When wild-type flies were tested in a behavioral paradigm consisting of a circular stage and two opposing visual cues, they walked in straight lines back and forth between the cues. However, $sim^{J1\text{-}47}/sim^{H9}$ flies did not walk in straight lines, but walked in circles. An individual fly could turn left or turn right, but not both, nor walk straight. Although basic locomotion appeared normal, its coordination was defective. Male courtship behavior was also affected. Examination of the adult brain revealed that interhemispheric axonal

connections were defective; the neuropil was thinner in *sim* mutants than in wild-type and there was disorganization at sites of axonal crossing. Thus, mutant defects in the Sim-positive cells may affect interhemispheric communication, leading to a split-brain fly that cannot properly coordinate its movement. Future questions concern the development and function of the Sim-positive central complex cells, and the developmental roles of *sim* in the central complex and optic lobes. Does *sim* control neurogenesis in the central brain as it does in the CNS midline cells, or does it influence other neurodevelopmental processes, such as axon guidance?

3.8.2 Genital structures and sterility

During embryonic development, the midline expression of *sim* extends past the presumptive CNS into segments A9-10, and terminates at the proctodeal (anal) opening (35, 41). The staining in segments A9 and A10 overlaps with the sites of the genital disc primordia. These are the presumptive imaginal structures that will give rise to the adult genitalia. Analysis of *sim* mutant embryos indicates that the genital discs form, but are misplaced (41). The misplacement could be due to either a defective genital disc or a consequence of improper condensation of the CNS. The proctodeum is also abnormal. Analysis of the cuticle indicates that the anal slit, which constitutes the proctodeal opening, is absent and the adjacent anal pads have fused. Thus, *sim* contributes to the formation of the genital disc and midline structures associated with the proctodeum.

$sim^{J1\text{-}47}$ mutants kept at the permissive temperature or $sim^{H9}/Df(3R)ry^{75}$ mutants survive into adulthood, but exhibit male and female sterility (41). A small percentage of $sim^{J1\text{-}47}$ mutant adults lack the genitalia and anus. Males lack the clasper and penis, and females lack the vulva. Both sexes lack an anus and the anal plates, and the flies are closed at the posterior end. Male and female gonads were only rudiments, and unattached to the gut via the internal genital structures. Thus, defects in the *sim* mutant genital discs result in severe defects in the genital structures and sterility. While gut structure appeared normal in newly emerged adults, since the hindgut was not connected to an anal opening, it became swollen. This resulted in the premature death of the flies.

3.8.3 Muscle precursors

There is a small cluster of 3-5 *sim*-expressing cells/hemisegment that arise just above the CNS (30). These cells migrate laterally to the body wall, where they differentiate into ventral oblique somatic muscles. The proper migration of these mesodermal cells requires Slit repulsive signals derived

from the CNS midline cells (29, 30). The expression of *sim* in these cells is transient, occurring initially as the cells first appear as pre-migratory muscle precursors before migration, and largely disappearing before they differentiate as muscles. Despite the expression of *sim* at a critical time in the development of these cells, genetic analysis of *sim* mutants specifically lacking expression in the muscle precursor cells did not reveal any obvious abnormalities in somatic musculature (30). This suggests that, unlike its major role in midline neurogenesis, *sim* does not play a major role in myogenesis.

4. DROSOPHILA TRACHEALESS – REGULATOR OF TRACHEAL AND SALIVARY DUCT DEVELOPMENT

The *Drosophila trh* gene was first discovered in the Nobel Prize-winning genetic screen of cuticle phenotypes by Nusslein-Volhard and Wieschaus as a mutant embryo devoid of trachea and a defective with a defective filzkörper (56). Two groups, one interested in tracheal development and the other in salivary gland development, identified the *trh* gene by virtue of P-element enhancer trap insertions that showed expression in both cell types (57, 58). Subsequent cloning and expression analysis of the *trh* gene revealed that it is expressed in the embryonic trachea, salivary duct, and subset of CNS cells. *trh* plays important developmental roles in the trachea and salivary duct. The CNS function is unknown, and it is expressed after their development into mature cell types (57) (Ward and Crews, unpubl.), suggesting a function in axonogenesis, synaptic connectivity, or neural function. Identification of *trh* in other insects and arthropods has revealed other additional sites of expression and potential functions, including silk gland development in the silk moth, and osmoregulation in brine shrimp.

4.1 Tracheal development and *trh* genetics

Since insects do not have an oxygen-carrying circulatory system, they depend on a diffuse, multi-branched trachea to deliver oxygen. Tracheal development (59) begins with the formation of segmentally repeated placodes. These cells invaginate and branch. Tracheal branches from different segments fuse to form the tubular, air-filled trachea that is closed except at the anterior and posterior spiracular openings. The posterior spiracle is connected to the trachea by the filzkörper, an elaborate structure which acts as a tracheal air filter.

The tracheal placode fails to invaginate in *trh* mutants, and the cells do not differentiate and form trachea (57, 58, 60). This is accompanied by an absence of expression (57, 58) of a number of tracheal-expressed genes, including *trh*, which undergoes positive autoregulation (58). One important tracheal-expressed gene whose expression is dependent on *trh* is *breathless* (*btl*), which is required for tracheal migration (2). Ectopic expression of *trh* resulted in additional tracheal placodes forming at two distinct sites in the dorsal ectoderm (58). These results suggested that *trh* is a master regulator of tracheal development. However, the situation is more complex. While *trh* clearly plays an important role in tracheal development, there are tracheal-expressed genes that are not dependent on *trh* function (61). In addition, while the filzkörper fail to elongate in *trh* mutants, they are able to secrete cuticle, suggesting that the cells have the correct identity, but cannot form tubes. Since *trh* is required for formation of trachea and salivary duct, both tubular cell types, it has been suggested that *trh* comprises a regulatory cassette that functions in tubule formation (57). Since known tracheal target genes of *trh* include *btl* and *rho*, which both participate in tracheal migration and invagination, the role of *trh* may be more in morphogenesis than tracheal precursor cell fate.

4.2 Trh biochemistry and tracheal expression

The sequence of Trh is similar to the other bHLH-PAS partners of Tgo: a bHLH domain near the N-terminus followed by PAS-1 and PAS-2, and ending with a large unconserved C-terminal region that functions as a transcriptional activation domain (57, 58). There are two splice variants that affect the sequence of PAS-1 and the spacer between PAS-1 and PAS-2 (58) - the significance of these variant proteins is unknown. Trh:Tgo binds to CMEs with ACGTG core sequences, and activates transcription (1, 2, 8). It is not surprising that Trh:Tgo recognizes the same binding site as Sim:Tgo and Sima:Tgo, since the Trh basic region has 9/13 aa identity with Sim and 11/13 aa identify with Sima. In vivo analysis has been carried-out on two Trh:Tgo target genes, *btl* and *rho* (2, 50). Both have multiple CMEs in close proximity that are required for tracheal expression. Both genes are expressed in the CNS midline cells, and the same CMEs are required for both tracheal expression by Trh and midline expression by Sim.

There is prominent *trh* expression in all embryonic tracheal cells and the posterior spiracle (57, 58). *trh* expression appears just as the tracheal primordia form, and remains on throughout embryogenesis and larval development in most tracheal cells. One exception is tracheal fusion cells, in which Trh levels decline due to negative regulation by Dys (62). It will be important to determine what function *trh* plays during late embryonic and

postembryonic development, and whether Sima (see below) may control Trh levels and function under hypoxic conditions, akin to Dys regulation of Trh.

4.3 Trh nuclear localization and phosphorylation

Examination of Trh subcellular localization shows it to be an exclusively nuclear protein (7, 58). Misexpression of *trh* results in ectopic Trh nuclear localization regardless of the cell type it is expressed (7). While this suggests that Trh dimerization and nuclear localization are unregulated, additional factors are required. Nuclear localization of Trh is dependent on phosphorylation by the *Dakt1* protein kinase B (PKB) (63). Trh is phosphorylated by Dakt1 at S665, which lies in the C-terminal region thought to be involved in transcriptional activation and nuclear localization (and also outside of the bHLH-and PAS domains). Mutants of *Dakt1* result in reduced levels of *trh* and *btl*, consistent with a role in regulating Trh function. How Dakt1 influences Trh nuclear localization is unknown, but could be due to effects on: (i) nuclear import by allowing access to the Trh nuclear localization sequence, (ii) export and removal of an inhibitor to nuclear export, or (iii) ability to dimerize with Tgo. The zygotic expression of *Dakt1* includes tracheal cells and sites in which ectopic trachea are formed when *trh* is expressed throughout the dorsal ectoderm. *Dakt1* corresponds to the vertebrate *akt* oncogene, which is regulated by phosphatidylinositol signaling. This suggests that Trh nuclear localization and function is dependent on a signaling pathway, although its biological roles remain unknown. Dakt1 may also be a component of the accessory factors proposed to regulate Trh function post-transcriptionally in the dorsal and dorsolateral ectoderm (8).

4.4 Transcriptional specificity: Drifter-Trh interactions

Ectopic expression of *trh* throughout the ectoderm results in ectopic tracheal pits and gene expression (58). However, the additional tracheal tissue forms at only two sites within the dorsal ectoderm. This suggests that additional factors are required for *trh* function and at least one of them is spatially restricted to the sites of ectopic *trh*-induced tracheal cells. Another aspect of transcriptional specificity concerns Sim and Trh. Even though Sim:Tgo binds the same DNA sequences as Trh:Tgo it is unable to induce tracheal gene expression when ectopically expressed (8). Transgenic domain swap experiments revealed that a Trh protein with Sim PAS domains behaved like Sim and unlike Trh, indicating that transcriptional specificity resided within the PAS domains (8). Two factors, Dakt1 and Drifter (Dfr) have emerged as cofactors for Trh function. One component of Trh tissue

specificity is phosphorylation by Dakt1, since it is required for the formation of ectopic trachea when *trh* is misexpressed. The zygotic transcription of *Dakt1* is spatially restricted to tracheal regions. However, the *Dakt1* maternal component is broadly expressed early in embryogenesis (63), and. the S665 PKB phosphorylation site also lies well outside of the PAS domains, indicating that other spatially-restricted factors are required for Trh function.

The Drifter (Dfr) POU-homeobox gene influences tracheal development and is expressed early in tracheal formation, similar to *trh* expression (61, 64). Dfr is also expressed at the sites in the dorsal ectoderm where ectopic trachea form when *trh* is misexpressed (50). Misexpression of both *trh* and *dfr* revealed the presence of ectopic trachea at sites in the ectoderm and head, beyond those observed for ectopic expression of only *trh*. These results demonstrate that Dfr is a coactivator of Trh in vivo. Biochemical experiments revealed that the Dfr POU domain directly binds to the Trh PAS domain (50) (Figure 6), consistent with the requirement of the Trh PAS domain for transcriptional specificity. Thus, Trh function requires the presence of both the Dfr coactivator and phosphorylation for proper function, and the coactivator clearly restricts Trh target gene activation. It is interesting that Dfr functions as a coactivator with both Trh and Sim, but biochemically behaves differently: Dfr binds directly to Trh via its PAS domain, but indirectly with Sim, which requires an additional cofactor, Fish, that binds to both Sim and Dfr (47) (Figure 6).

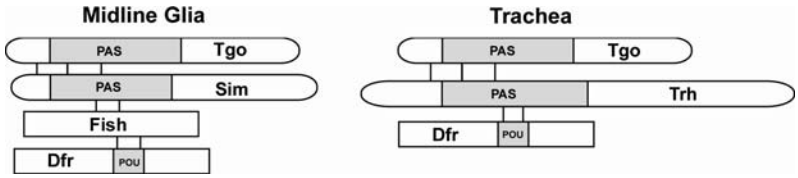

Figure 6. Dfr employs multiple modes of biochemical interactions with bHLH-PAS proteins. In the midline glia, Dfr interacts with Sim indirectly by binding to Fish via the Dfr POU domain, which binds to Sim via the Sim PAS domain. In the trachea, the Dfr POU domain binds directly to Trh via the Trh PAS domain.

4.5 Initiation of *trh* transcription

The expression of *trh* and corresponding formation of tracheal placodes occurs at precise positions in the dorsal ectoderm. It is convenient to think of the tracheal placodes as coordinates along the D/V and A/P axes, specified by axis patterning genes. Along the D/V axis, the dorsal extent of *trh* expression is established by repression by the Decapentaplegic (Dpp)

TGF-β signaling pathway (57, 58), and the ventral extent is set by repression by the epidermal growth factor (EGF) pathway (16). The cues governing A/P positioning are not well known although Wingless has been implicated (58, 65). The tracheal placodes are relegated to 10 thoracic and abdominal segments and are not found in the terminal segments. This is due to terminal region repression by the *spalt* gene (66). Initial expression of *dfr* is independent of *trh* and likely governed directly by the same genes that regulate *trh* expression. This is consistent with their co-equal roles in tracheal development. After *trh* and *dfr* are activated in tracheal placode cells, they combine to autoregulate their own expression, since the initial patterning cues fade out. Trh and Dfr also combine to activate transcription of some tracheal target genes, and they may individually combine with other proteins to regulate expression of other target genes (61). Further insight into the respective developmental roles of *trh* and *dfr* will emerge as the identities of additional target genes for each transcription factor are discovered.

4.6 Non-tracheal roles of *trh*

4.6.1 Drosophila salivary duct development

4.6.1.1 Expression and genetics

The *Drosophila trh* gene plays a prominent role in controlling embryonic salivary duct development. The formation of the salivary tissue has emerged as an excellent system for studying tissue formation (67, 68). The two salivary glands are connected to the pharynx by ducts. Each gland is joined to its own duct, and these two ducts merge at their anterior ends to form a common duct. The salivary primordia consist of precursors to both duct and gland cells. *trh* is initially expressed in the entire primordia, but is later restricted to the salivary duct cells (57, 58). Despite the expression of *trh* in the primordia, only the duct cells are affected in *trh* mutant embryos (57, 58, 69). The *trh* mutant salivary glands appear relatively normal, but are closed at the site where they would normally join with the ducts. In contrast, the salivary duct cells fail to invaginate, and remain on the surface of the embryo. Trh regulates expression of several genes in the duct cells, including *eyegone*, which is required for formation of the individual ducts (70). However, *trh* may be carrying-out only a subset of duct cell developmental functions, such as morphogenesis, since some duct-expressed genes are not dependent on *trh* (67). Trh and Tgo colocalize to nuclei in both salivary primordia and duct. However, the Trh coactivator, Dfr, is

absent in the salivary duct, so that Trh:Tgo presumably interacts with other coregulatory proteins to control duct cell transcription.

4.6.1.2 Control of *trh* salivary duct expression

Since *trh* plays an important role in salivary duct development, understanding the factors that control its expression provide insight into how regulatory proteins dictate cell fates. There are also interesting evolutionary implications since similar factors may control the formation of tissues in other species (see below). Initial expression of *trh* is in the ventral ectoderm of parasegment 2 (corresponds to the posterior maxillary segment plus anterior labial segment) (23, 57), which is the domain for the salivary primordia. *trh* expression is then restricted to the duct primordia, which lies ventral to the gland cells. Thus, regulation of *trh* involves three issues: (i) how D/V patterning restricts expression to the ventral ectoderm, (ii) how homeotic genes, which control parasegmental identity, restrict expression to parasegment 2, and (iii) how *trh* expression is restricted to the duct cells, and is absent from the gland cells.

The *Sex combs reduced* (*Scr*) homeotic gene controls parasegment 2 identity, and *Scr* is required, along with the *extradenticle* and *homothorax* transcription factor genes, to restrict *trh* expression to only parasegment 2 (23, 57). Thus, mutants of *Scr* result in a loss of *trh* expression and ubiquitous expression of *Scr* results in additional *trh* expression in more anterior segments (57). The *Dpp* gene functions along with *brinker* in the dorsal ectoderm to repress *trh* expression, thus restricting *trh* expression to the salivary primordia within the ventral ectoderm (23, 57). Within the salivary primordia, the ventral-most cells become duct cells, and the more dorsal cells form the glands. *trh* expression is present in the duct cells and is absent from the gland cells. The transcriptional repressor, Fork head (Fkh) functions to repress *trh* in the gland cells (69). *fkh* expression is complementary to *trh*: it is off in the duct cells, but present in the gland cells. Thus, a key issue is how is *fkh* expression restricted to the gland cells, but not duct cells. This is due to midline-directed Sim/Spitz signaling.

The ventral-most midline cells in parasegment 2 express *sim*, which likely controls *spitz*/*EGFR* signaling (23). Spitz acts as a morphogen emanating from the midline that is high in concentration ventrally and low dorsally. Mutants in *sim* or other *spitz* class genes result in a loss of duct cell fate and a corresponding expansion of gland cell fate (23, 69). One of the consequences of Spitz signaling is repression of *fkh* in the duct cell primordia. Since Fkh represses *trh* expression, this results in restriction of *trh* to the duct cells.

4.6.2 *Bombyx trh* and silk gland development

The silkworm, *Bombyx mori*, is well known for its ability to produce silk. It has been proposed that the silk-producing gland is a modified salivary gland, since it is derived from the labial segment, as is the salivary gland. The silk gland consists of three distinct regions: the anterior silk gland (ASG), middle silk gland (MSG), and posterior silk gland (PSG). The ASG is a duct for the silk proteins secreted by the MSG and PSG. This functional specialization of the silk gland resembles that of the salivary gland in which an anterior salivary duct connects the two salivary glands to the foregut.

Bombyx possesses a *trh* gene (*Bm-trh*) (71) that is highly related to *Drosophila trh*. Embryonic expression of *Bm-trh* was observed in the silk gland, trachea, and supracolonic trachea (homolog of the *Drosophila* posterior spiracle). Tracheal and supracolonic tracheal expression begins in the placodes and continues throughout development. This is analogous to *trh* expression in *Drosophila* and suggests a similar function as a regulator of tracheal development. Initial *Bm-trh* expression is present in the primordia of the entire silk gland, but then becomes restricted to the ASG, the silk gland duct. This also resembles *trh* expression in the *Drosophila* salivary primordia, in which *trh* is initially expressed throughout the primordia, and is then restricted to the salivary duct.

The *Bombyx silk gland factor-3* (*SGF-3*) gene (72) encodes a POU-homeobox protein highly related to *Drosophila* Dfr. *SGF-3* is expressed in the developing trachea and silk gland. Initially, *SGF-3* is expressed throughout the entire silk gland primordia, but is later restricted to the ASG and part of the MSG. This expression pattern overlaps with *Bm-trh*, which is also expressed in the trachea, silk gland primordia, and then ASG. As in *Drosophila* and *Artemia* (see below), Bm-Trh and SGF-3 may interact as a regulatory cassette to control transcription and development. The one phylogenetic difference is that *Drosophila* Trh and Dfr do not interact to control salivary duct development.

Further molecular analysis of silk gland gene expression, revealed additional similarities between silk gland and salivary gland development. *Drosophila trh* is initially expressed in cells that will give rise to both salivary duct and glands, but is later expressed in only the duct cells. The *Drosophila fkh* gene is expressed in the gland cells and required for their development (73). *fkh* represses *trh* in the gland cells, thus restricting *trh* expression to the duct cells. The *Bombyx fkh* gene (*Salivary Gland Factor-1*; *SGF-1*) is expressed only in the MSG and PSG, but not ASG (74). *Bm-trh* initially is expressed throughout the silk gland primordia, and is later restricted to the ASG (71). Thus, *Bombyx* Fkh may be acting to repress *Bm-*

trh and *SGF-3* in the MSG and PSG, similar to its role in the *Drosophila* salivary primordia.

4.6.3 *Artemia trh* and osmoregulation

The branchiopod crustacean brine shrimp, *Artemia franciscana*, resides in salt pools as cysts. Upon hydration, the cysts hatch and larvae develop and live in the hyperosmotic salt ponds. Nauplius larvae, the first emerging larval forms, possess a specialized organ, the salt gland, which regulates osmolarity. Later in development, the salt gland is resorbed, and osmoregulation is carried-out by the thoracic epipods that reside on the appendages (75). Given its lifestyle, osmoregulation is a critical element of brine shrimp physiology.

Artemia possess a *trh* gene (*Af-trh*) highly related to insect *trh* (76). *In situ* hybridization of *Artemia* larvae with an *Af-trh* probe revealed expression in the salt gland of the nauplius and the thoracic epipods of older larvae – both sites of osmoregulation (Figure 7). Thus, another function of *trh* may be the development and function of crustacean osmoregulatory organs. An *Artemia drifter* gene, *APH-1*, was identified and its expression pattern analyzed (77). Interestingly, it is also expressed in the salt gland (epipod expression was not reported). This suggests that Af-Trh and APH-1 interact to control transcription in *Artemia* similar to their role in insect tracheal development, and represent an evolutionarily-conserved regulatory protein cassette.

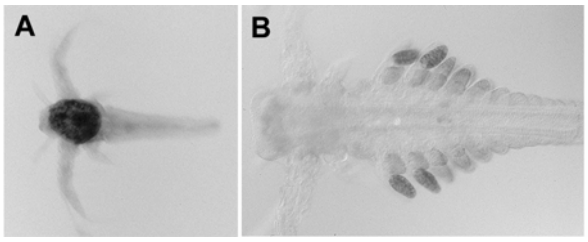

Figure7. Artemia trh is expressed in sites of osmoregulation. (A) Expression of *Af-trh* in the naupliar salt gland. (B) Later expression of *Af-trh* in the larval epipods. Adpated from (76).

Drosophila osmoregulation occurs in a subset of cells in the hindgut (78), a site in which *trh* expression has not been reported. Thus, this aspect of *trh* function may not be conserved between crustaceans and insects. *Artemia* do not possess identifiable respiratory organs, and respiration is likely due to diffusion over the entire body surface, including the epipods. It remains an interesting, but open question, whether the role of *trh* in respiratory system development is conserved in the crustacean subphylum.

5. DYSFUSION: REGULATOR OF TRACHEAL FUSION

5.1 Dys structure and expression

The Dys bHLH-PAS protein was identified using a bioinformatic screen of the *Drosophila* genome (62). It is a member of well-conserved subgroup of bHLH-PAS proteins conserved in *C. elegans* (C15C8.2) (see chapter by Powell-Coffman) and mammals (NXF). Dys is a nuclear protein and overlaps with nuclear Tgo, strongly suggesting that Dys and Tgo heterodimerize. Expression of *dys* is found in a variety of cell types including: (i) tracheal fusion cells, (ii) epidermal leading edge, (iii) foregut atrium, (iv) nervous system, (v) hindgut, and (vi) anal pad. Work on the *dys* gene has focused on its role in tracheal fusion cells.

The insect trachea are derived from a series of segmentally-reiterated placodes. Signals from adjacent cells induce the tracheal cells to migrate in multiple directions (59). In most cases, these branches meet and fuse. There are four branch types that fuse, including the: (i) dorsal branch, (ii) dorsal trunk, (iii) lateral trunk, and (iv) ganglionic branch. Each migrating branch has a tip cell that guides branching and then mediates fusion (79). In this manner, the mature, diffuse tracheal system is created. The fusion process is complex, involving recognition, adhesion, cytoskeletal rearrangement, formation of adherens junctions, and formation of a continuous tubule. The fusion cell has a distinct tracheal cell identity indicated by a unique pattern of gene expression and cellular function.

The *dys* gene is expressed in all tracheal fusion cells, but no other tracheal cells. The *escargot* (*esg*) gene encodes a zinc finger transcription factor that is also expressed in tracheal fusion cells. *esg* expression occurs prior to fusion, and plays an important role in fusion cell fate and function (80, 81). *dys* is expressed in fusion cells after the appearance of *esg*, but also before fusion occurs (62). Analysis of *dys* expression in *esg* mutant embryos showed that *esg* regulates *dys* expression in all branches, except the dorsal trunk. The expression of *dys* in fusion cells suggested a role in the fusion cell process.

5.2 Dys function and Dys-Trh interactions

The function of *dys* was examined using RNAi (62). Embryos injected with *dys* RNAi had a high percentage of tracheal fusion defects. The dorsal branch, lateral trunk, and ganglionic branch showed an absence of fusion.

The dorsal trunk was unaffected. These results were similar to those for *esg*, which showed defects in all branches except the dorsal trunk.

Initially, the Trh protein exists at uniformly high levels in all tracheal cells (7). As Dys protein levels rise in fusion cells during wild-type tracheal development, the levels of Trh decline in fusion cells (62). The decline of Trh is dependent on *dys*, since Trh levels remain high when *dys* expression is absent due to injection of *dys* RNAi. The mechanism in which Dys downregulates Trh is unknown. One model is that Dys outcompetes Trh for their common dimerization partner, Tgo, and the reduction of Trh:Tgo heterodimers causes a corresponding reduction in *trh* transcription, since *trh* is autoregulatory.

The developmental significance of the reduction in Trh by Dys is unknown, but may facilitate the transition of the tip cell from a migratory role to a fusion role. Since there are a number of examples in vertebrates and invertebrates in which bHLH-PAS genes overlap in expression (82, 83), the study of Dys-Trh interactions may prove to be an important model. There are a number of important issues to be addressed. What are the identities of Dys target genes and what are their roles in tracheal fusion and related processes. Does *dys* carry-out a subset of *esg* functions? Does Dys activate or repress transcription? What is the mechanism by which Dys reduces Trh levels, and what is the biological significance of this reduction? Does Dys regulate levels of other bHLH-PAS proteins?

6. SIMILAR: HYPOXIA REGULATION

The *similar* (*sima*) gene was identified by low stringency hybridization with a *sim* probe (84). When the human *hypoxia inducible factor-1α* (*HIF-1α*) (see chapter by Semenza) gene was cloned and sequenced (85), it became apparent that Sima was most related to HIF-1α, and represented a potential candidate for an insect HIF. Subsequent work (86, 87) has confirmed that Sima is a HIF-1α ortholog, and functions in regulating the response to hypoxia, much the way HIF-1α regulates the mammalian response to hypoxia.

6.1 *Drosophila* hypoxia responsiveness

Insight into the *Drosophila* physiological, developmental, and behavioral responses to low oxygen levels have appeared in several studies. Use of a transgenic fly strain containing a hypoxia-sensitive reporter indicated that the transcriptional response to hypoxia occurs throughout development (86). Responsiveness increases during embryogenesis with a peak in late-stage

embryos, and decreases afterwards. However, it remains relatively high during larval development. This is particularly relevant since larvae crawl through rotting food where oxygen may be limiting. Consistent with this, hypoxia alters the feeding behavior of larvae through a nitric oxide/cyclic GMP pathway, such that oxygen-deprived larvae stop feeding and begin moving (presumably with the desire to find an environment with higher oxygen levels) (88). Induction of the reporter peaked at oxygen concentrations around 3% (normoxia is 20%). Induction was absent at 0% oxygen (anoxia), consistent with work showing that anoxia results in cell cycle and metabolic arrest.

During hypoxic conditions, the trachea is the main cell type showing high hypoxia-sensitive reporter induction (86). The significance of this is unclear. One proposal is that the trachea may act as a sensor for oxygen and convey that information to the nervous system, which influences larval behavior. However, the main effects of hypoxia have been observed in non-tracheal cells, as they seek a source of oxygen (59). Hypoxic larval cells secrete Branchless (Bnl; a fly fibroblast growth factor-like protein) and induce new tracheal branches to migrate towards the Bnl-producing cells (89). Another mechanism involves the extension of cytoplasmic projections from hypoxic non-tracheal cells that attach to tracheoles and pull them towards the oxygen-starved cells (59, 90). Consistent with the ability of non-tracheal cells to respond to hypoxia, induction of the hypoxia-dependent reporter can be seen outside of the trachea when conditions of hypoxia are relatively severe (86). It will be important to analyze *sima* mutants to understand its role in hypoxia regulation and additional, unforeseen, biological processes. Does *sima* regulate Bnl-mediated tracheal branching and the ability of non-tracheal cells to extend cytoplasmic processes and grab existing tracheoles? What is the relevance of the strong induction of hypoxia-dependent gene regulation in the trachea? Does *sima* mediate long-term changes in morphology, physiology, and behavior?

6.2 Structure and expression

The Sima protein is one of the longest *Drosophila* bHLH-PAS proteins (1505 aa). Sima heterodimerizes with Tgo (1) and is able to bind and activate transcription from an ACGTG CME-like DNA sequence (Estes and Crews, unpubl.) (8, 86). As with other bHLH-PAS proteins, the bHLH domain is near the N-terminus followed by PAS-1 and PAS-2 domains. The C-terminal half of Sima has a large number of homopolymeric stretches, including 13 stretches of poly[glutamine]. These homopolymeric stretches likely contribute to the transcriptional activation function that resides in the C-terminal region. Within the C-terminal region is also an evolutionarily-

conserved oxygen-dependent degradation domain (ODDD), which regulates the levels of Sima protein with respect to oxygen concentration (86). In addition, the ODDD contains sequences required for cytoplasmic localization under normoxia (see below).

RNA levels of *sima* appear uniform in most, if not all, cells of the embryo under normoxia (84), and the levels change little under hypoxia (86). Sima protein levels are low at normal oxygen levels, and the protein is localized to the cytoplasm (86). Protein levels dramatically increase under low oxygen tension, and the protein is localized to nuclei (86). As with mammalian HIF-1α protein, this suggests that Sima protein levels and function are predominantly regulated at the level of protein degradation.

6.3 Biochemistry of hypoxia regulation

The regulation of HIF-1α stability under varying oxygen conditions has been well-studied in vertebrates (see chapter by Semenza), and the same basic model applies to Sima. Under normoxia, HIF-1α is hydroxylated on two prolines in the ODDD by HIF prolyl hydroxylases. These hydroxylated compounds bind von Hippel Lindau (VHL) factor, which mediates degradation by the ubiquitin/proteasome. Under hypoxia, the hydroxylase is inhibited, VHL fails to bind, and the protein is no longer degraded. While the biochemical details have not been described, *Drosophila* proteins related to those involved in hypoxia control of HIF-1α have been identified. These include *Drosophila* von Hippel-Lindau (D-VHL) factor (91) and HIF prolyl hydroxylase (Hph; CG1114) (86).

D-VHL RNAi experiments resulted in tracheal defects, such as excessive branching, looping, and breakage (91), its role in hypoxia and mediating Sima degradation has not been studied. Its expression is predominantly tracheal. The *Drosophila Hph* gene is ubiquitously expressed in the embryo. *Hph* RNAi and genetic experiments carried-out under normoxia resulted in upregulation of Sima protein and its localization to the nucleus (86), accompanied by activation of the hypoxia-sensitive reporter. Expression of *Hph* was upregulated under hypoxia, similar to affects seen with mammalian prolyl hydroxylases. These results indicate that prolyl 4-hydroxylase functions in hypoxia regulation. Mutational analysis of Sima is consistent with Hph acting directly on Sima. The ODDD contains two prolines within sequences conserved between Sima and HIF-1α. Deletion of the ODDD resulted in nuclear accumulation of Sima under normoxia and activation of the hypoxia-sensitive reporter (86). This confirms that the ODDD controls Sima degradation and cytoplasmic retention under normoxia, and that it is regulated by prolyl 4-hydroxylase. While, the role of D-VHL and proteins of the ubiquitin degradation pathway have not been analyzed for their roles

in hypoxia regulation, the evidence is strong that the regulatory pathway is highly related to that utilized in mammals and *C. elegans*. Nevertheless, it will be interesting to see whether there exist features of hypoxia regulation idiosyncratic to *Drosophila*.

6.4 Transcriptional specificity

Sima:Tgo is able to bind and activate transcription from a multimerized ACGTG (CME) core sequence, just as Sim:Tgo and Trh:Tgo. Since each protein complex regulates different developmental and physiological processes, this raises the issue of transcriptional specificity and how the different complexes control distinct gene batteries. Insight into this issue has emerged from analysis of two related, but distinct, in vivo gene reporters that were tested for their response to hypoxia (86). When a pentamer of the murine *erythropoietin* hypoxia response element (HRE), which contains HIF-1α:Arnt (and presumably Sima:Tgo) CME binding sites, was tested in vivo, it failed to be activated under hypoxia. When a fragment of the mammalian *lactate dehydrogenase A* (*LDH-A*) gene was tested, it was activated under hypoxia. The *LDH-A* fragment has two HREs and a cyclic AMP responsive element (CRE). These results suggest that a CRE-binding protein acts as a coactivator with Sima:Tgo for hypoxia induction. Since this reporter is not expressed at significant levels in CNS midline cells or trachea, this regulatory region is specific for Sima:Tgo activation , but not for Sim:Tgo or Trh:Tgo. Further progress will emerge when target genes of Sima:Tgo are identified and analyzed.

7. *SPINELESS*: APPENDAGE IDENTITY

Mutations in the *ss* gene were first reported in 1923 by Bridges and Morgan (92). Null mutants of *ss* are viable, and possess a number of interesting phenotypes that include: (i) reduction in bristle size, (ii) deletion of much of the tarsal segment of the leg, and (iii) transformation of the distal antenna to distal leg (Figure 8). The latter two phenotypes have established *ss* as a key gene in controlling appendage identity. The *ss* gene was cloned based on the identification of a P-element transposon insertion that had a weak *ss* phenotype (10). Molecular analysis revealed that *Drosophila ss* gene is orthologous to the mammalian *aryl hydrocarbon receptor* (*ahr*; dioxin receptor) and *C. elegans ahr-1* (93), indicating that these genes encode a highly-conserved subfamily of bHLH-PAS proteins. Unlike *ahr*, which is best known for its role in toxin metabolism (see chapter by Yao et

al.), the known roles of *ss* are developmental. Direct tests have shown that *Drosophila* Ss does not bind dioxin (94).

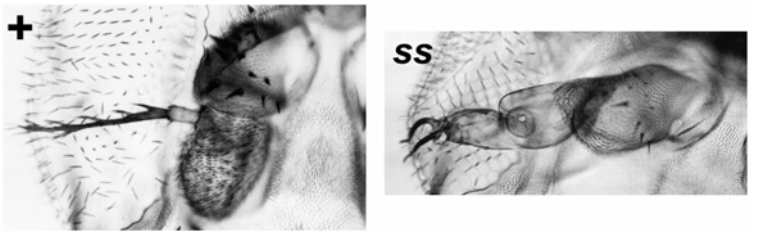

Figure 8. ss mutant shows a transformation of antenna to leg. Wild-type adult (+) showing antenna including the distal arista attached to segment A3. In *ss* null mutant, the arista is transformed to distal leg. Adapted from (9).

7.1 Antennal and leg anatomy and development

The adult antenna and leg are derived from imaginal discs that are specified during embryogenesis and develop during the larval and pupal stages. The mature antenna is an olfactory organ that consists of 6 segments (proximal to distal): A1-5, and the arista (A6). Olfactory neurons reside on A3. Below the antenna are the maxillary palps, which are also olfactory organs. The adult leg is derived from the leg imaginal disc and consists of the coxa, trochanter, femur, tibia, and 5 tarsal segments (proximal to distal). The last tarsal segment terminates in a claw.

7.2 Biochemistry

Like Sim and Trh, Ss heterodimerizes with Tgo, likely migrates to the nucleus, binds DNA and activates transcription (9). Ss:Tgo prefers GCGTG binding sites, as do Ahr:Arnt heterodimers. This differs from the ACGTG preference of Sim:Tgo, Trh:Tgo, and Sima:Tgo, and provides a partial explanation for how Ss:Tgo regulates different genes and biological processes distinct from the other bHLH-PAS proteins. When *ss* is ectopically expressed, Tgo localizes in the nucleus at the ectopic sites, suggesting that a Ss:Tgo nuclear complex is formed at those sites (9). Since no differences have been observed with respect to cell type, this suggests that nuclear entry is not dependent on localized or tissue-specific ligand binding, as is required for vertebrate Ahr. However, this observation does not rule out that some Ss function may still require ligand binding or modification.

7.3 Expression

ss is generally expressed at sites of *ss* genetic function (10). Embryonic sites include: (i) the larval antennal sense organ primordia, (ii) the gnathal segments (maxillary, mandibular, and labial), (iii) in patches within the thoracic segments that correspond to the presumptive legs, and (iv) sensory cells. All of these sites colocalized with nuclear Tgo, except the leg primordia. Corresponding to the sites of embryonic expression are two larval *ss* phenotypes: a deformed antennal sensory organ and mislocalized dorsomedial papilla of the maxillary sense organ.

Postembryonically, *ss* is expressed in the leg, antennal, wing, genital, and eye discs and all bristle precursor cells (10). *ss* mutants delete the distal region of leg segment 1 and segments 2-4, resulting in a fusion of 1 and 5. Leg disc expression of *ss* likely corresponds to those tarsal cells deleted in *ss* mutants. Expression in the antennal disc is predominantly in antennal segments A2, A3 and the aristae, consistent with the *ss* mutant defects in A2, A3 and the aristae, and the absence of *ss* defects in antennal segments A1 (10, 95). During pupariation, *ss* is expressed in bristle sensory organ precursors in most discs, which also correlates with the *ss* bristle defects. Later it can be seen that *ss* expression is restricted to the bristle cells, but not to the socket cells. The postembryonic expression of *ss* correlates nicely with the known *ss* defects, although expression in the genital, labial, and eye discs does not correlate with any known *ss* phenotype. The *ss* gene is dynamically expressed in a variety of cell types; critical to its function is understanding its regulation and the nature of the target genes it regulates.

7.4 *ss* genetics and regulation

7.4.1 Antennal development

ss null mutants result in a transformation of the antennal arista into distal leg (see (10). The antennal third segment is converted to smooth cuticle of unspecified identity. In addition misexpression of *ss* in imaginal discs is able to transform maxillary palps, rostral membrane (which lies between the antenna and maxillary palp), and distal leg to antennal-like structures (10). These results indicate that *ss* plays an important role in specifying the identity of the distal antenna. Since ectopic expression of *ss* only converts certain cell types to antennal fates, Ss presumably functions with cell-type specific coregulators to carry-out its functions.

The *Distalless* (*Dll*) and *homothorax* (*hth*) homeobox genes regulate both appendage identity and proximodistal positional information (95). Mutants

in both genes result in a transformation of antenna to leg. Mutations in *Dll* also result in distal deletions of both leg and antenna, and *hth* mutants result in proximal deletions of both appendages. Genetic studies indicate that both *Dll* and *hth* activate *ss* expression in the antenna and leg (10, 95). Both *Dll* and *hth* are spatially restricted in the developing appendages, and their expression and activities are likely required for the spatial distribution of *ss* in the antennal and leg discs. While *ss* plays an important role in antennal identify, it is not a master regulator of antennal gene expression, since numerous genes involved in antennal differentiation and function are not regulated by *ss (95)*. Two roles for *ss* are repression of tarsal and tarsal claw in the antenna, and the formation of olfactory sense organs in A3. Thus, *ss* is a downstream target of Dll and Hth and carries-out a subset of their functions.

7.4.2 Leg development

ss null mutants result in a loss of the distal part of tarsal segment 1 (T1) and T2-4. When *ss* is misexpressed in distal and proximal leg segments, it results in a deletion of the medial femur and tibia segments and a transformation of the claw into arista (10). Regulation of *ss* in the leg resembles antennal *ss* regulation, since *Dll* controls ss expression in both appendages (10). One function of *ss* is to regulate the expression of *bric-a-brac* (*bab*). *bab* is required for the formation of several tarsal joints (96). Ss positively regulates *bab* in the leg (10), and *dachshund* is involved in repressing *bab* (97), resulting in stripes of *bab* expression at the tarsal joints.

7.5 Evolutionary considerations

Despite seemingly different biological roles for *ss* and *Ahr*, it has been proposed that an ancestral *ss* gene was involved in chemosensory function (9). This is consistent with the expression of *C. elegans Ahr-1* in sensory neurons (see chapter by Powell-Coffman), *Drosophila ss* in formation of sensory organs, and mammalian Ahr in binding (sensing) toxic molecules. Future work may reveal additional similarities. It cannot be excluded that the nematode and insect proteins also function as ligand-binding receptors and may have physiological roles. Similarly, further work on *Ahr* may reveal conserved developmental functions.

8. SUMMARY

The *Drosophila* bHLH-PAS proteins that partner with Tgo share a number of properties. It is convenient to group Dys, Sim, Ss, and Trh together, since their regulation and mode of action differs from Sima. Each member of this "DSST" subgroup is specifically expressed in a number of cell types during embryonic and postembryonic development. There is little redundancy or overlap among these proteins, and this lack of redundancy likely contributes to the numerous mutant phenotypes associated with these genes. Several DSST genes play fundamental roles in tissue-specific development. The *sim* gene can reasonably be called a master regulator of CNS midline cell development, since it is required for all midline cell developmental events (directly or indirectly), and the initiation or maintenance of expression of every midline-expressed gene assayed. The *ss* and *trh* genes also play important roles in the formation of tissue types, since like *sim*, their ectopic expression results in formation of their corresponding cell types (antenna, trachea, and CNS midline cells) at ectopic sites. However, even in the antenna and trachea, ss and trh, respectively, do not control all aspects of development and transcription. For all three genes, the misexpression experiments that result in additional sites of tissue generation are not ubiquitous but found only at specific locations. This reinforces that idea that these proteins require additional factors for function. These may be coregulators that directly interact with or modify the bHLH-PAS proteins or additional factors that act independently and carry-out functions complementary to those of the bHLH-PAS proteins.

The ability of each bHLH-PAS gene to regulate different target genes in different cell types, as well as the fact that multiple bHLH-PAS proteins dimerize with Tgo and bind the same DNA sequences raises issues of transcriptional specificity. While other mechanisms may exist, current evidence suggests that specificity arises from interactions between bHLH-PAS proteins and different coregulatory proteins (Trh-Dfr, Sim-Fish-Dfr). The PAS domain mediates these interactions, but more biochemical work needs to be carried-out to better define the PAS domain interactions, as well as understand the entire range of PAS domain functions. Why are there two PAS domains/protein, and why is the sequence so evolutionarily diverse? Even though regulatory proteins, such as Sim and Trh, regulate different cell types (CNS midline and trachea) there exist many genes that are expressed in both cell types and utilize the same binding sites for both modes of expression. Is there an evolutionary significance to this, such that an ancestral organism possessed a cell type or tissue that is the predecessor to both existing *Drosophila* cell types? Does this extend to other cell types regulated by bHLH-PAS proteins?

The *Drosophila* bHLH-PAS partners of Tgo are expressed in a diverse array of cell types, although there are similarities. Three of the five (Dys, Sima, and Trh) play roles in tracheal development. Is this fortuitous or purposeful? Evidence for the latter derives from the observation that Dys downregulates Trh in tracheal fusion cells. Since many prokaryotic and eukaryotic PAS proteins function in sensory roles and the PAS domain often mediates ligand binding, it is tempting to think that the *Drosophila* developmental bHLH-PAS proteins may also respond to developmental signals. Evidence to data has not turned up much evidence for this, although Trh function is dependent on phosphorylation. Instead the spatial and temporal control over function derives primarily from the transcriptional regulation of these genes. The *sim* gene is a large gene with multiple promoters and a sophisticated array of cis-control elements that direct its highly specific expression. Although only *sim* has been studied in detail, the *dys*, *ss*, and *trh* genes are also large and likely to contain a large number of cis-control elements. Nevertheless, the developmental, physiological, and environmental regulation of function via cofactor binding (protein or small molecule), or modification via the PAS domain or other protein region, should be considered. As an example, phosphorylation of Trh function by PKB was unexpected. The PAS domain proteins remain central to understanding the regulation of developmental and physiological processes in *Drosophila*.

ACKNOWLEDGEMENTS

I would like to thank Debbie Andrew, Ian Duncan, Dan Lau, and John Nambu for thoughtful comments on the manuscript.

REFERENCES

1. 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.
2. 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.
3. 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.

4. Wharton, J., K. A., 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.
5. 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.
6. Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene *paired* and the Drosophila PRD gene set as part of a gene network. *Cell* 47:735-746.
7. Ward, M. P., J. T. Mosher, and S. T. Crews. 1998. Regulation of *Drosophila* bHLH-PAS protein cellular localization during embryogenesis. *Development* 125:1599-1608.
8. 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.
9. Emmons, R. B., D. Duncan, P. A. Estes, P. Kiefel, J. T. Mosher, M. Sonnenfeld, M. P. Ward, I. Duncan, and S. T. Crews. 1999. The Spineless-Aristapedia and Tango bHLH-PAS proteins interact to control antennal and tarsal development in *Drosophila*. *Development* 126:3937-3945.
10. Duncan, D. M., E. A. Burgess, and I. Duncan. 1998. Control of distal antennal identity and tarsal development in *Drosophila* by *spineless-aristapedia*, a homolog of the mammalian dioxin receptor. *Genes Dev.* 12:1290-1303.
11. Hilliker, A. J., S. H. Clark, W. M. Gelbart, and A. Chovnick. 1981. Cytogenetic analysis of the rosy micro-region, polytene chromosome interval 87D2-4; 87E12-F1, of *Drosophila melanogaster*. *Drosophila Inform. Serv.* 56:65-72.
12. Schalet, A., R. P. Kernaghan, and A. Chovnick. 1964. Structural and phenotypic definition of the rosy cistron in Drosophila melanogaster. *Genetics* 50:1261-1268.
13. 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.
14. 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.
15. 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.
16. Raz, E., and B. Z. Shilo. 1993. Establishment of ventral cell fates in the Drosophila embryonic ectoderm requires DER, the EGF receptor homolog. *Genes Dev.* 10:1937-1948.
17. 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.
18. Schmid, A., A. Chiba, and C. Q. Doe. 1999. Clonal analysis of Drosophila embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development* 126:4653-4689.
19. Crews, S. T. 1998. Control of cell lineage-specific development and transcription by bHLH-PAS proteins. *Genes Dev.* 12:607-620.

20. Jacobs, J. R. 2000. The midline glia of Drosophila: a molecular genetic model for the developmental functions of glia. *Prog. Neurobiol.* 62:475-508.
21. Hummel, T., K. Schimmelpfeng, and C. Klambt. 1999. Commissure formation in the embryonic CNS of *Drosophiila* I. Identification of the required gene functions. *Dev. Biol.* 209:381-398.
22. Sonnenfeld, M. J., and J. R. Jacobs. 1994. Mesectodermal cell fate analysis in *Drosophila* midline mutants. *Mech. Dev.* 46:3-13.
23. Panzer, S., D. Weigel, and S. K. Beckendorf. 1992. Organogenesis in Drosophila melanogaster: embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* 114:49-57.
24. Luer, K., J. Urban, C. Klambt, and G. M. Technau. 1997. Induction of identified mesodermal cells by CNS midline progenitors in *Drosophila*. *Development* 124:2681-2690.
25. Zhou, L., H. Xiao, and J. R. Nambu. 1997. CNS midline to mesoderm signaling in Drosophila. *Mech. Dev.* 67:59-68.
26. Page, D. T. 2003. A function for EGF receptor signaling in expanding the developing brain in Drosophila. *Curr. Biol.* 13:474-82.
27. Therianos, S., S. Leuzinger, F. Hirth, C. S. Goodman, and H. Reichert. 1995. Embryonic development of the Drosophila brain: formation of commissural and descending pathways. *Development* 121:3849-60.
28. Dickson, B. J. 2002. Molecular mechanisms of axon guidance. *Science* 298:1959-64.
29. Kidd, T., C. Russell, C. S. Goodman, and G. Tear. 1998. Dosage sensitive and complementary functions of roundabout and commissureless control axon crossing of the CNS midline. *Neuron* 20:25-33.
30. 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.
31. Mehta, B., and K. M. Bhat. 2001. Slit signaling promotes the terminal asymmetric division of neural precursor cells in the Drosophila CNS. *Development* 128:3161-8.
32. Condron, B. G. 1999. Serotonergic neurons transiently require a midline-derived FGF signal. *Neuron* 24:531-40.
33. Menne, T. V., K. Luer, G. M. Technau, and C. Klambt. 1997. CNS midline cells in *Drosophila* induce the differentiation of lateral neural cells. *Development* 124:4949-4958.
34. 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 which acts as a master regulator of CNS midline development. *Cell* 67:1157-1167.
35. 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.
36. 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.
37. 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.

38. 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.
39. 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.
40. Kasai, Y., S. Stahl, and S. Crews. 1998. Specification of the *Drosophila* CNS midline cell lineage: direct control of *single-minded* transcription by dorsal/ventral patterning genes. *Gene Expression* 7:171-189.
41. Pielage, J., G. Steffes, D. C. Lau, B. A. Parente, S. T. Crews, R. Strauss, and C. Klambt. 2002. Novel behavioral and developmental defects associated with Drosophila single-minded. *Dev. Biol.* 249:283-99.
42. 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. USA* 89:3414-3418.
43. Morel, V., and F. Schweisguth. 2000. Repression by suppressor of hairless and activation by Notch are required to define a single row of single-minded expressing cells in the Drosophila embryo. *Genes Dev.* 14:377-88.
44. Cowden, J., and M. Levine. 2002. The Snail repressor positions Notch signaling in the Drosophila embryo. *Development* 129:1785-93.
45. Morel, V., R. Le Borgne, and F. Schweisguth. 2003. Snail is required for Delta endocytosis and Notch-dependent activation of single-minded expression. *Dev. Genes Evol.* 213:65-72.
46. McGuire, J., P. Coumailleau, M. L. Whitelaw, J. A. Gustafsson, and L. Poellinger. 1995. The basic helix-loop-helix/PAS factor Sim is associated with hsp90. Implications for regulation by interaction with partner factors. *J. Biol. Chem.* 270:31353-7.
47. Ma, Y., K. Certel, Y. Gao, E. Niemitz, J. Mosher, A. Mukherjee, M. Mutsuddi, N. Huseinovic, S. T. Crews, W. A. Johnson, et al. 2000. Functional interactions between Drosophila bHLH/PAS, Sox, and POU transcription factors regulate CNS midline expression of the slit gene. *J. Neurosci.* 20:4596-605.
48. Estes, P., J. Mosher, and S. T. Crews. 2001. *Drosophila single-minded* represses gene transcription by activating the expression of repressive factors. *Dev. Biol.* 232:157-175.
49. 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.
50. Zelzer, E., and B. Shilo. 2000. Interaction between the bHLH-PAS protein Trachealess and the POU-domain protein Drifter, specifies tracheal cell fates. *Mech. Dev.* 19:163-173.
51. McDonald, J. A., S. Holbrook, T. Isshiki, J. Weiss, C. Q. Doe, and D. M. Mellerick. 1998. Dorsoventral patterning in the Drosophila central nervous system: the vnd homeobox gene specifies ventral column identity. *Genes Dev.* 12:3603-3612.
52. Chu, H., C. Parras, K. White, and F. Jimenez. 1998. Formation and specification of ventral neuroblasts is controlled by vnd in Drosophila neurogenesis. *Genes Dev.* 12:3613-3624.

53. Wharton, J., K. A., and S. T. Crews. 1993. CNS midline enhancers of the *Drosophila slit* and *Toll* genes. *Mech. Dev.* 40:141-154.
54. Sanchez-Soriano, N., and S. Russell. 1998. The Drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline. *Development* 125:3989-3996.
55. Strauss, R. 2002. The central complex and the genetic dissection of locomotor behaviour. *Curr. Opin. Neurobiol.* 12:633-8.
56. Jurgens, G., E. Wieschaus, C. Nusslein-Volhard, and H. Kluding. 1984. Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster. *Roux Arch. Dev. Biol.* 193:283-295.
57. Isaac, D. D., and D. J. Andrew. 1996. Tubulogenesis in *Drosophila*: a requirement for the *trachealess* gene product. *Genes Dev.* 10:103-117.
58. 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.
59. Manning, G., and M. A. Krasnow. 1993. Development of the *Drosophila* tracheal system, p. 609-685. In M. Bate and A. Martinez Arias (ed.), The Development of *Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
60. Younossi-Hartenstein, A., and V. Hartenstein. 1993. The role of the tracheae and musculature during pathfinding of Drosophila embryonic sensory axons. *Dev. Biol.* 158:430-47.
61. Boube, M., M. Llimargas, and J. Casanova. 2000. Cross-regulatory interactions among tracheal genes support a co-operative model for the induction of tracheal fates in the Drosophila embryo. *Mech. Dev.* 91:271-8.
62. Jiang, L., and S. T. Crews. 2003. The *Drosophila dysfusion* bHLH-PAS gene controls tracheal fusion and levels of the Trachealess bHLH-PAS protein. *Mol. Cell. Biol.* In press.
63. Jin, J., N. Anthopoulos, B. Wetsch, R. C. Binari, D. D. Isaac, D. J. Andrew, J. R. Woodgett, and A. S. Manoukian. 2001. Regulation of Drosophila tracheal system development by protein kinase B. *Dev Cell* 1:817-27.
64. Anderson, M. G., G. L. Perkins, P. Chittick, R. J. Shrigley, and W. A. Johnson. 1995. *drifter*, a *Drosophila* POU-domain transcription factor, is required for correct differentiation and migration of tracheal cells and midline glia. *Genes Dev.* 9:123-137.
65. de Celis, J. F., M. Llimargas, and J. Casanova. 1995. Ventral veinless, the gene encoding the Cf1a transcription factor, links positional information and cell differentiation during embryonic and imaginal development in Drosophila melanogaster. *Development* 121:3405-16.
66. Kuhnlein, R. P., and R. Schuh. 1996. Dual function of the region-specific homeotic gene spalt during Drosophila tracheal system development. *Development* 122:2215-23.
67. Andrew, D. J., K. D. Henderson, and P. Seshaiah. 2000. Salivary gland development in Drosophila melanogaster. *Mech. Dev.* 92:5-17.
68. Bradley, P. L., A. S. Haberman, and D. J. Andrew. 2001. Organ formation in Drosophila: specification and morphogenesis of the salivary gland. *Bioessays* 23:901-11.

69. Kuo, Y. M., N. Jones, B. Zhou, S. Panzer, V. Larson, and S. K. Beckendorf. 1996. Salivary duct determination in *Drosophila*: roles of the EGF receptor signaling pathway and the transcription factors fork head and trachealess. *Development* 122:1909-17.
70. Jones, N. A., Y. M. Kuo, Y. H. Sun, and S. K. Beckendorf. 1998. The Drosophila Pax gene eye gone is required for embryonic salivary duct development. *Development* 125:4163-74.
71. Matsunami, K., H. Kokubo, K. Ohno, P. Xu, K. Ueno, and Y. Suzuki. 1999. Embryonic silk gland development in Bombyx: molecular cloning and expression of the Bombyx trachealess gene. *Dev. Genes Evol.* 209:507-14.
72. Matsunami, K., H. Kokubo, K. Ohno, and Y. Suzuki. 1998. Expression pattern analysis of SGF-3/POU-M1 in relation to sericin-1 gene expression in the silk gland. *Dev. Growth Differ.* 40:591-7.
73. Weigel, D., H. Bellen, G. Jurgens, and H. Jackle. 1989. Primordium-specific requirement of the homeotic gene *fork head* in the developing gut of the *Drosophila* embryo. *Wilhelm Roux's Arch. Dev. Biol.* 198:201-201.
74. Kokubo, H., S. Takiya, V. Mach, and Y. Suzuki. 1996. Spatial and temporal expression pattern of *Bombyx fork head*/SGF-1 gene in embryogenesis. *Dev. Genes Evol.* 206:80-85.
75. Martin, J. W. 1992. Branchiopoda, Microscopic Anatomy of the Invertebrates, Volume 9: Crustacea. Wiley-Liss Inc.
76. Mitchell, B., and S. T. Crews. 2002. Expression of the *Artemia trachealess* gene in the salt gland and epipod. *Evol. Dev.* 4:1-10.
77. Chavez, M., C. Landry, S. Loret, M. Muller, J. Figueroa, B. Peers, F. Rentier-Delrue, G. G. Rousseau, M. Krauskopf, and J. A. Martial. 1999. APH-1, a POU homeobox gene expressed in the salt gland of the crustacean Artemia franciscana. *Mech. Dev.* 87:207-12.
78. Lengyel, J. A., and D. D. Iwaki. 2002. It takes guts: the Drosophila hindgut as a model system for organogenesis. *Dev. Biol.* 243:1-19.
79. Samakovlis, C., N. Hacohen, G. Manning, D. Sutherland, K. Guillemin, and M. A. Krasnow. 1996. Development of the *Drosophila* tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. *Development* 122:1395-1407.
80. Samakovlis, C., G. Manning, P. Steneberg, N. Hacohen, R. Cantera, and M. A. Krasnow. 1996. Genetic control of epithelial tube fusion during *Drosophila* tracheal development. *Development* 122:3531-6.
81. Tanaka-Matakatsu, M., T. Uemura, H. Oda, M. Takeichi, and S. Hayashi. 1996. Cadherin-mediated cell adhesion and cell motility in *Drosophila* trachea regulated by the transcription factor Escargot. *Development* 122:3697-705.
82. Gradin, K., J. McGuire, R. H. Wenger, I. Kvietikova, M. L. Whitelaw, R. Toftgard, L. Tora, M. Gassmann, and L. Poellinger. 1996. Functional interference between hypoxia and Dioxin signal transduction pathways: competition for recruitment of the Arnt transcription factor. *Mol. Cell. Biol.* 16:5221-5231.
83. Woods, S. L., and M. L. Whitelaw. 2002. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element. Cross-talk between basic

helix-loop-helix/per-Arnt-Sim homology transcription factors. *J. Biol. Chem.* 277:10236-43.

84. Nambu, P. A., and J. R. Nambu. 1996. The Drosophila fish-hook gene encodes a HMG domain protein essential for segmentation and CNS development. *Development* 122:3467-75.
85. Wang, G. L., B.-H. Jiang, 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. USA* 92:5510-5514.
86. Lavista-Llanos, S., L. Centanin, M. Irisarri, D. M. Russo, J. M. Gleadle, S. N. Bocca, M. Muzzopappa, P. J. Ratcliffe, and P. Wappner. 2002. Control of the hypoxic response in Drosophila melanogaster by the basic helix-loop-helix PAS protein similar. *Mol. Cell. Biol.* 22:6842-53.
87. Bacon, N. C. M., P. Wappner, J. F. O'Rourke, S. M. Bartlett, B. Shilo, C. W. Pugh, and P. J. Ratcliffe. 1998. Regulation of the *Drosophila* bHLH-PAS protein Sima by hypoxia: functional evidence for homology with mammalian HIF-1α. *Bioch. Biophys. Res. Comm.* 249:811-816.
88. Wingrove, J. A., and P. H. O'Farrell. 1999. Nitric oxide contributes to behavioral, cellular, and developmental responses to low oxygen in Drosophila. *Cell* 98:105-14.
89. Jarecki, J., E. Johnson, and M. A. Krasnow. 1999. Oxygen regulation of airway branching in Drosophila is mediated by branchless FGF. *Cell* 99:211-20.
90. Wigglesworth, V. B. 1977. Structural changes in the epidermal cells of Rhodnius during tracheole capture. *J. Cell Sci.* 26:161-74.
91. Adryan, B., H. J. Decker, T. S. Papas, and T. Hsu. 2000. Tracheal development and the von Hippel-Lindau tumor suppressor homolog in Drosophila. *Oncogene* 19:2803-11.
92. Bridges, C., and T. H. Morgan. 1923. The third-chromosome group of mutant characters of Drosophila melanogaster. *Carnegie Inst. Wash. Publ.*:1-251.
93. Powell-Coffman, J. A., C. A. Bradfield, and W. B. Wood. 1998. *Caenorhabditis elegans* orthologs of the aryl hydrocarbon receptor and its heterodimerization partner the aryl hydrocarbon receptor nuclear translocator. *Proc. Nat. Acad. Sci. USA* 95:2844-2849.
94. Butler, R. A., M. L. Kelley, W. H. Powell, M. E. Hahn, and R. J. Van Beneden. 2001. An aryl hydrocarbon receptor (AHR) homologue from the soft-shell clam, Mya arenaria: evidence that invertebrate AHR homologues lack 2,3,7,8-tetrachlorodibenzo-p-dioxin and beta-naphthoflavone binding. *Gene* 278:223-34.
95. Dong, P. D., J. S. Dicks, and G. Panganiban. 2002. Distal-less and homothorax regulate multiple targets to pattern the Drosophila antenna. *Development* 129:1967-74.
96. Godt, D., J. L. Couderc, S. E. Cramton, and F. A. Laski. 1993. Pattern formation in the limbs of Drosophila: bric a brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. *Development* 119:799-812.
97. Chu, J., P. D. Dong, and G. Panganiban. 2002. Limb type-specific regulation of bric a brac contributes to morphological diversity. *Development* 129:695-704.