Control of *Drosophila* head segment identity by the bZIP homeotic gene cnc

Jym Mohler1,*, James W. Mahaffey2, Elizabeth Deutsch1 and Kodela Vani1

1Department of Biological Sciences, Barnard College, New York, NY 10027, USA
2Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA

*Author for correspondence

**SUMMARY**

Mutational analysis of *cap’n’collar* (*cnc*), a bZIP transcription factor closely related to the mammalian erythroid factor NF-E2 (p45), indicates that it acts as a segment-specific selector gene controlling the identity of two cephalic segments. In the mandibular segment, *cnc* has a classical homeotic effect: mandibular structures are missing in *cnc* mutant larvae and replaced with duplicate maxillary structures. We propose that *cnc* functions in combination with the homeotic gene *Deformed* to specify mandibular development. Labral structures are also missing in *cnc* mutant larvae, where a distinct labral primordium is not properly maintained in the developing foregut, as observed by the failure to maintain and elaborate patterns of labral-specific segment polarity gene expression. Instead, the labral primordium fuses with the esophageal primordium to contribute to formation of the esophagus. The role of *cnc* in labral development is reciprocal to the role of homeotic gene *forkhead*, which has an identical function in the maintenance of the esophageal primordium. This role of homeotic selector genes for the segment-specific maintenance of segment polarity gene expression is a unique feature of segmentation in the preoral head region of *Drosophila*.

Key words: *Drosophila*, head segment, *cnc*, homeotic gene

**INTRODUCTION**

In *Drosophila*, segment identity is defined by the continuous expression of homeotic selector genes throughout development in segment-specific patterns. Of these homeotic genes, the best characterized are the homeobox containing genes of the Antennapedia (ANTC) and Bithorax (BXC) complexes (see reviews by Peifer et al., 1987, and Kaufman et al., 1990). While clearly of major importance, the homeobox homeotic genes do not account for all aspects of segment identity in the fly. The ANTC and BXC homeotic genes function only in the trunk and postoral head region of the embryo and are not expressed in the terminal regions or in most of the preoral head. Further, these homeobox homeotic genes are usually expressed in domains larger than a single segment, yet each segment has a distinct morphology. While the subtle morphological differences between abdominal segments can be explained by subtle differences in overlapping expression of BXC genes (Peifer et al., 1987), in much of the head region, the ANTC complex genes are expressed in non-overlapping domains in some cases encompassing two morphologically distinct segments (Kaufman et al., 1990). In particular, a single ANTC gene, *Deformed* (*Dfd*), is expressed in the mandibular and maxillary segments, yet the derivatives of these two segments are quite dissimilar and share no common, reiterant structures (Chadwick and McGinnis, 1987; Mahaffey et al., 1989). Thus, specification of segment identity must also involve other genes acting either independently of or in concert with the genes of ANTC and BXC homeotic gene clusters.

Indeed, mutations in many other genes have been described that cause homeotic transformations in *Drosophila*. The majority of these genes act to maintain ANTC and BXC gene expression patterns through development (the Pc-group and trx-group genes, Jürgens, 1985; Kennison and Tamkun, 1988; Paro, 1990). However, a few such genes function as true segment-specific homeotic genes, acting in parallel to the BXC and ANTC complex genes. These latter genes, encoding transcription factors of diverse types, are expressed either in domains overlapping those of the ANTC and BXC genes (e.g. *teashirt* [*tsh*, Röder et al., 1992] and *spalt* [*sal*, Kühnlein et al., 1994]), or may be expressed exclusively in the terminal regions outside of the domains of ANTC and BXC function (e.g. *forkhead* [*fkh*, Weigel et al., 1989]).

We describe here the effects on embryonic segment identity of the loss of function in another segment-specific homeotic gene, *cap’n’collar* (*cnc*). The *cnc* gene is continuously transcribed throughout *Drosophila* embryogenesis in only the labral and mandibular segments, which give rise to the dorsal and to the posterior lateral and ventral portions of the pharynx, respectively (Mohler et al., 1991). *cnc* transcription is maintained through embryogenesis independently of the activity of the ANTC homeotic genes, following activation in blastoderm stages by gap genes and maternal factors (Mohler, 1993). The *cnc* gene encodes a bZIP transcription factor (Mohler et al., 1991) with remarkable similarity to the p45 subunit of the NF-E2 transcription factor regulating hemoglobin synthesis in mammals (Andrews et al., 1993a). The DNA-binding domain of *cnc* is also shared with *skinhead* (*skn1*), a maternal deter-
ominant of *C. elegans* required for the specification of the pharynx and intestine of the early embryo (Bowerman et al., 1992). We demonstrate here that cnc is required in *Drosophila* for segment identity of the two pharyngeal segments (mandibular and labral) in which it is expressed.

**MATERIALS AND METHODS**

**hh enhancer trap**
The 16E enhancer trap strain contains a PwB (Wilson et al., 1989) enhancer trap in which β-galactosidase is expressed in *hh* domains. 16E was derived from hh*PZ* (Mohler and Vani, 1992) by disruption of the enhancer trap (Tower et al., 1993) selecting for homozygous viable transpositions that retained β-galactosidase expression in *hh* domains. Spatial patterns of embryonic and larval β-galactosidase activity in this strain was detected as described by Bellen et al. (1989).

**Domain alignment of cnc and hh expression domains**
The double in-situ/antibody labeling protocol of Azpiazu and Pfeifle (1993) was utilized to compare cnc and hh expression domains in the 16E enhancer trap strain. hh expression domains were detected by HRP-linked immunohistochemistry to β-galactosidase under *hh* control; cnc transcripts were detected by in situ hybridization (Tautz and Pfeifle, 1989) using an antisense RNA DIG-labeled probe to the 1A cnc cDNA (Mohler et al., 1991).

**In situ hybridization**
In situ hybridizations were performed according to protocols of Tautz and Pfeifle (1989). For wg hybridizations, antisense DIG-labeled RNA corresponding to the 3′-terminal PstI fragment of the wg cDNA (Rijsewijk et al., 1987) was used as a probe. For 1.28 hybridizations, the 1.28 probe of Mahaffey et al. (1993) was used as described therein. wg*PZ* hybridizations were detected by in situ hybridization using an antisense RNA DIG-labeled probe to the 1A cnc cDNA (Mohler et al., 1991).

**RESULTS**

**cnc expression is restricted to the labral and mandibular parasegments**

*cnc* expression is spatially restricted to two regions of the *Drosophila* embryo, and is maintained throughout embryonic development (Mohler et al., 1991). *cnc* expression is activated in blastoderm stages (Fig. 1A) in an anterior cap and a posterior stripe. The anterior domain becomes associated with the clypeolabral lobe in late extended germband stages. The posterior domain becomes associated later with the mandibular and hypopharyngeal lobes, both of which are portions of the mandibular segment (Mahaffey et al., 1989). To define clearly the segmental limit of the two anterior expression domains of the *cnc* gene, *cnc* RNA expression patterns were compared to the expression domains of the segment polarity gene, *hedgehog* (*hh*). The segmental nature of the head region is most evident in the extended germband embryo, when segmentally repeated domains of segment polarity gene transcription (wg, en and hh) can be visualized in an S-shaped curve around the anterior midgut invagination (Schmidt-Ott and Technau, 1992, Fig. 2).

In the mandibular region at stage 10 (Fig. 1C), *cnc* is seen within the trapezoidal limited by hh expression posteriorly (mandibular posterior compartment, Md) and laterally (intercalary posterior compartment, Ic) and by the stomodeum anteriorly. Thus, the posterior domain is restricted to the anterior compartment of the mandibular segment (parasegment 0, Fig. 2), and does not extend into the intercalary segment. The previous description of this expression in both the mandibular and intercalary segments (Mohler, 1993) was due to the erroneous assumption that the hypopharyngeal lobe is part of the intercalary segment (as has been suggested by Jürgens and Hartenstein, 1993). Comparison with the expression patterns of segment polarity genes shows the hypopharyngeal lobe resides strictly within the anterior compartment of the mandibular segment.

In the labral region at stage 11 (Fig. 1B), the dorsal limit of *cnc* expression is the cleft separating the clypeolabral lobe from the acron. This cleft is a parasegmental boundary across which *wg* and *en* are expressed (Schmidt-Ott and Technau, 1993), although hh is not. Expression of *cnc* extends ventrally through the dorsal pharyngeal hh domain. This anterior *cnc* expression domain, therefore, corresponds to what has been described as parasegment-4 (PS-4, Schmidt-Ott and Technau, 1993; Fig. 2), although the homology of PS-4 with the other parasegments is unclear (see Jürgens and Hartenstein, 1993). Indeed, the late activation of an additional *wg* domain in the middle of this 'parasegment' (see below) suggests that it is probably not homologous with the other parasegments. For simplicity in this paper, we shall refer to the derivatives of the clypeolabral lobe, corresponding to PS-4 where *cnc* is expressed, as 'labral' structures.

**Isolation of cnc mutations**
The first *cnc* mutation, cnc*PZ*, was obtained from A. Spradling as a transposition of the *PZ* enhancer trap into the 94E3,4 polytene region (Karpen and Spradling, 1992). Southern analysis of genomic DNA containing this mutation localized the insertion of *PZ*-element to a position approximately 300 bp upstream of the 5′-end of the longest *cnc* cDNA (Fig. 3). Embryos homozygous for cnc*PZ* failed to accumulate levels of *cnc* RNA detectable by in situ hybridization. These embryos also failed to express β-galactosidase from the *PZ*-element. Precise excision of this transposon from the 94E3,4 region restored complete viability, indicating that this transposon insertion is the cause of the associated embryonic lethality. Because of this apparent failure of *cnc* expression, *cnc*PZ is a putative null mutation of *cnc*.

Two additional null mutations, *cnc*VL70 and *cnc*VL110, were obtained following imprecise excision of the *PZ*-element from...
cnc homeotic gene

These mutations deleted part or all of the cnc protein coding region (Fig. 3). A fourth cnc null mutation, Df(3R)EB6, which deletes the 3’-portion of the cnc gene as well as several more proximal genes including hh, was obtained following γ-irradiation-induced loss of a w+ transposon (P[w,ry]A) located in the adjacent unk gene (Mohler et al., 1992).

cnc mutants lack mandibular and labral head structures

The three cnc alleles are embryonic lethal mutations and have similar homeotic effects on larval head morphology (Fig. 4B-D). cnc mutant larvae possess head skeletons missing the labrum and dorsal pouch (labral structures, Jürgens et al., 1986), the transverse-ribs and ventral arms (derivatives of the hypopharyngeal lobe, Jürgens et al., 1986), and the lateralgräte

Fig. 1. Relation of cnc and hh expression. (A) Expression of cnc at late blastoderm stage. cnc is expressed in an anterior dorsal cap (5-6 cells wide dorsally narrowing ventrally, between approximately 95-86% egglength at the onset of gastrulation; Mohler, 1993) and a single 2-4 cell wide stripe of expression at approximately 68% egg length. (B,C) Alignment of cnc (blue) and hh expression domains (brown). (B) In an early stage 12 embryo, the dorsal limit of labral cnc expression is the cleft (arrow) dividing the acron (Ac) from the clypeolabral lobe (Cl). Labral cnc expression extends ventrally through the dorsal pharyngeal expression domain of hh, but not further into the foregut. Mandibular cnc expression is anterior to the mandibular expression domain (Md) of hh. (C) In a late stage 10 embryo, mandibular cnc expression is between the intercalary (Ic) and mandibular (Md) hh expression domains. (ES, esophageal hh domain, Mx, maxillary hh domain)

Fig. 2. Schematic of the head of a stage 11 embryo. During stage 10 and early stage 11 of embryogenesis, the cephalic segments snake around the anterior midgut invagination, which formed ventrally. These segments (indicated by parasegment numbers) can be visualized by the periodic expression of posterior compartment genes (en or hh, black). cnc (stippled) is expressed in the anterior compartment of mandibular segment (parasegment 0), which includes the hypopharyngeal lobe and the anterior portion of the mandibular lobe. cnc is also expressed in ‘parasegment-4’, corresponding to the clypeolabral lobe. (Segment abbreviations: FG, foregut; Lr, Labral; Pr, Preantennal; An, Antennal; Ic, Intercalary; Md, Mandibular; Mx, Maxillary.)

Fig. 3. Map of the mutations affecting the cnc transcription unit in polytene band 94E3. The cncPZ insertion maps approximately 300 bp upstream from the end of the longest cnc cDNA. The cncVL70 and cncVL110 delete portions of the cnc locus from the PZ-insertion point into the coding region. Df(3R)EB6 deletes from inside the second exon, downstream of the protein initiation codon, proximally through the rest of cnc locus, the unk and hh loci to 94C2-5.
Fig. 4. Phenotypic effects of head homeotic mutations on the larval head skeleton. (A) Wild-type; (B-D) cnc\textsuperscript{VL110}; (B) lateral; (C) dorsal surface; (D) pharyngeal views. (E) Dfd\textsuperscript{W21}; (F) Dfd\textsuperscript{W21} cnc\textsuperscript{PZ}; (G) lab\textsuperscript{D}; (H) lab\textsuperscript{D} cnc\textsuperscript{PZ}; (I) sal\textsuperscript{A55}; (J) sal\textsuperscript{A55} cnc\textsuperscript{PZ}. Abbreviations: labrum (lr), dorsal bridge (db), hypostomal sclerite (hys), H-piece (H), ventral arms (VA), ventral plate (VP), dorsal arms (DA), T-ribs (T), posterior pharyngeal wall (ppw), antennal sense organ (AnSO), maxillary sense organ (MxSO).
(a derivative of the mandibular lobe, Jürgens et al., 1986). cnc mutant larvae are also missing most of the posterior pharyngeal wall, which our analysis (see below) indicates is derived from the ventral portion of the clypeolabral lobe and is, therefore, of labral origin. cnc mutant larvae also possess duplicate, terminally located bilateral maxillary structures: a secondary mouthhook and an extra cirri row. Both the normal and duplicated mouthhook are missing the mouthhook base.

Some cnc’Z’Z’ individuals show a slightly weaker phenotype, in which one duplicated mouthhook is positioned internally at the base of ventral plate (not shown). These individuals can be interpreted as ones in which the mandibular lobe has migrated properly during head involution before differentiating into maxillary structures. This low penetrant phenotype suggests that cnc’Z’Z’ is a strong, but not quite null, allele. No discernible differences were observed between homozygous cnc mutants and animals heterozygous for a cnc mutant and Df[3R]EB6, which removes the 3’ portion of the cnc gene.

The structures missing in these cnc mutant embryos are labral, hypopharyngeal or mandibular derivatives, corresponding to regions of cnc expression. The fact that structures derived from regions outside the cnc expression domains are essentially unaffected by cnc null mutations indicates that cnc acts in a ‘segment-autonomous’ fashion, consistent with its function as a segment-specific homeotic gene.

**Independence of cnc function and other homeotic selector genes**

Double mutant combinations of cnc and other homeotic mutations affecting the head (Deformed [Dfd], labial [lab] and spalt [sal]) all show additive, independent defects in the head skeleton (Fig. 4E–I). Dfd mutant embryos are missing maxillary structures (mouthhooks, cirri and maxillary sense organs) and retain the hypopharyngeal structures (ventral arms and T-ribs), but the mandibular lateralgräte are misshapen, possibly due to a partial transformation to duplicate ventral arms (Fig. 4E: Merrill et al., 1987; Regulski et al., 1987). Doubly mutant cnc Dfd embryos (Fig. 4F) lack all structures (hypopharyngeal, mandibular and maxillary) missing in the single mutants, including the duplicated mouthhooks and cirri found in the cnc single mutant.

Larvae mutant for lab (Fig. 4G) are missing the ventral plate (presumably an intercalary structure) and the dorsal bridge (originating from at least two domains in the acron region). lab larvae also show a weak ‘dorsal pouch syndrome’ in which the labral structures are poorly involuted (Jürgens et al., 1986, Merrill et al., 1989). Doubly mutant lab cnc larvae (Fig. 4H) lack all the structures missing in the single mutants and retain the duplicated maxillary structures seen in cnc larvae. These double mutant lab cnc larvae have a closed oral apertures, indicating a suppression of the ‘dorsal pouch syndrome’ of lab by cnc.

Larvae mutant for sal (Fig. 4I) lack the hypo-stomal sclerite and H-piece, two labial structures which in sal mutants have been transformed into duplicated prothoracic structures (Jürgens, 1988). sal mutant larvae are also missing the dorsal bridge (an acronal structure) and exhibit a stronger ‘dorsal pouch syndrome’ than seen in lab larvae (compare with Fig. 4G). Doubly mutant sal cnc larvae (Fig. 4J) are missing all the structures missing in the single mutants and retain duplicated maxillary structures of the cnc single mutants. The sal cnc double mutants have a closed oral apertures, demonstrating, as in the case with lab above, the suppression of the sal ‘dorsal pouch syndrome’ by cnc.

In each of the three double mutant combinations of cnc and the other cephalic homeotic genes, the structures missing are the combination of the structures missing in cnc mutant embryos and those of the other homeotic mutant. In the case of lab and sal, the structures affected are distinct from those affected by cnc, suggesting that the defects in the double mutant are due to the combination of two distinct affected domains. Indeed both lab and sal are expressed in adjacent domains, flanking but not overlapping the cnc expression domain. lab expression is restricted to the intercalary segment (Mahaffey et al., 1989) just anterior to the posterior cnc expression domain. In the head, sal is expressed in two domains: in the acron up to stage 9 in a domain slightly posterior to the anterior cnc expression domain and in a band from parasegment 1 through anterior parasegment 3, just posterior to the posterior cnc expression domain (Kühnlein et al., 1994). The phenotypes observed in the double mutant embryos are consistent with the idea that each gene is only required for structures within its expression domain.

In contrast, cnc and Dfd affect overlapping sets of structures. Maxillary structures, such as the cirri and the anterior portion of the mouthhook, require Dfd but not cnc. Hypopharyngeal structures (ventral arms and T-ribs) require cnc expression but not Dfd. The lateralgräte, a derivative of the mandibular lobe, requires both Dfd and cnc expression for normal morphology. Similarly, the posterior portion of the mouthhook requires both Dfd and cnc expression, suggesting it may also be a mandibular structure; however, because it is also absent in Scr mutant larvae, which is suggestive instead of a labial origin (Patel-Cucci et al., 1991), its segmental origin remains uncertain. The overlapping phenotypic effects of Dfd and cnc mutants appear to reflect their overlapping expression in the mandibular segment, with Dfd expression in most of the mandibular segment (except for the hypopharyngeal lobe) and the adjacent maxillary segment (Chadwick and McGinnis, 1987; Mahaffey et al., 1989).

**cnc represses maxillary development in the mandibular segment**

A likely explanation for the duplicated mandibular structures in cnc mutant larvae is that they result from a homeotic transformation of one of the domains in which cnc would normally be expressed. To address this possibility, we examined the expression of a maxillary-specific gene (1.28, Mahaffey et al., 1993) at a stage prior to head involution when the segment primordia could be unambiguously distinguished. In wild-type embryos, cephalic expression of 1.28 is restricted to the posterior lateral portion of the maxillary lobe at stage 12 (Fig. 5A). This cephalic expression is strictly dependent on Dfd expression in the maxillary segment, but does not occur in the mandibular lobe in which Dfd is also expressed (Mahaffey et al., 1993). In cnc mutant embryos, expression of 1.28 occurs in the posterior lateral portions of both the maxillary and mandibular lobes (Fig. 5B). This mandibular expression of 1.28 in cnc mutant embryos represents an unambiguous shift from mandibular to maxillary development in the absence of cnc function and demonstrates that an important role of cnc in the mandibular segment is to repress the Dfd-mediated activation of maxillary-specific genes.
**Fig. 5.** Expression of Dfd-specific gene, 1.28. *In situ* hybridization to 1.28 RNA is generally detected as dots of expression (presumably nuclear, Mahaffey et al., 1993) surrounded by fainter generalized tissue staining in the maxillary lobe above underlying gut expression, which is out of focus in these figures. In wild-type late stage 12 embryos (A), dotted 1.28 RNA expression can be detected in the posterior lateral portion of the maxillary segment (Mx). In homozygous cncPZ embryos, dotted 1.28 expression can be observed in the posterior lateral mandibular (Md) cells as well as in the maxillary cells. No expression is detected in the labial lobe (Lb).

**Fig. 6.** Disruption of wg RNA expression in the foregut by cnc and fkh mutants. At stage 10 in wild-type embryos (A), wg expression is detected at the stomodeum (ES) and in a labral (Lr) domain at the dorsal limit of the clypeolabrum. During late stage 11 (B), an additional wg expression domain (dPH) is induced in the ventral lateral portion of the clypeolabrum, which invaginates during stage 12 to form the dorsal pharyngeal wg domain (C, stage 13). In fkhX78 mutant embryos, the wg domain is normal during late stage 9 (D), but disappears in stage 10 (E). In mid-stage 11 cncX76 mutant embryos (F), the labral wg domain disappears and a reduced dorsal pharyngeal wg domain transiently appears.
Fig. 7. Disruption of hh RNA expression in the foregut by cnc and fkh mutants. In early stage 11 (A) a single hh expression RNA domain is present in the foregut (FG). By stage 12, this domain has in wild-type embryos (B) divided into separate dorsal pharyngeal (dPH) and esophageal (ES) domains. By stage 14, the esophageal hh domain has again separated into anterior (aES) and posterior (pES) domains in wild-type embryos (C). In fkh^{XT6} mutants (D), the division of the foregut domain has not occurred and only a ventrolateral clypeolabral hh domain remains. In cnc mutants this division also does not occur in cnc mutants, but instead only the circumferential esophageal domain remains (E, cnc^{PZ}). In cnc mutants, this remaining esophageal domain later divides normally into anterior and posterior subdomains (F, cnc^{VL110}).

Fig. 8. hh expression domains in the foregut in late embryos and early 1st instar larvae. (A) β-galactosidase activity driven in strain 16E under hh control is found in the dorsal pharyngeal region (dPH), the esophagus (ES) and the inner sleeve of the proventriculus (PV). (B) β-galactosidase activity from the dorsal pharyngeal domain persists to the first instar larva in posterior pharyngeal wall (arrow).
The *cnc* and *fkh* homeotic genes define, respectively, the dorsal pharyngeal and esophageal regions of the foregut

Two homeotic genes, *cnc* and *fkh*, are expressed in the foregut region of the developing embryo: *cnc* in the labral region fated to give rise to the dorsal pharynx and *fkh* in the adjacent esophagus. In the foregut, the segment polarity genes *wingless* (*wg*) and *hedgehog* (*hh*) are expressed in dynamic patterns throughout embryogenesis beginning at blastodermic stages (Baker, 1987, 1988; van den Heuvel et al., 1989; Mohler and Vani, 1992; Lee et al., 1992). To determine the roles of the *cnc* and *fkh* homeotic genes in the foregut, we examined effects of mutations in these genes on the expression of segment polarity genes in this region.

Three distinct *wg* expression domains arise in the foregut region during embryogenesis: esophageal (ES), labral (Lr) and dorsal pharyngeal (dPH, van den Heuvel et al., 1989). The esophageal domain is activated near the anterior end of the cellular blastodermic (stage 7) coincident with the activation of the *wg* segmental stripes, and later (stage 10) forms a ring around the stomodeum (Fig. 6A). The bilaterally paired labral domains are activated slightly later during the quick phase of germ band extension (stage 8) and become located at the dorsal margin of the clypeolabral lobe in stage 11 (Fig. 6A,B). During late stage 11, the dorsal pharyngeal domains of *wg* expression are induced in bilateral anterior ventral regions of the clypeolabral lobe (Fig. 6B). The intensity of the labral domain diminishes after stage 12, in synchrony with the segmental stripes in the trunk, as the expression of the dorsal pharyngeal domain intensifies (Fig. 6C).

In *fkh* mutant embryos, which can be readily identified after gastrulation by the aberrant anterior and posterior midguts, expression of the foregut *wg* domains is normal up to the beginning of stage 10 (Fig. 6D). At this stage, when the esophageal primordium would normally invaginate, *wg* expression in the esophageal domain of *fkh* mutant embryos is shut off (Fig. 6E). The dorsal pharyngeal domains are subsequently activated normally in *fkh* mutants during stage 11 (not shown).

Similarly, in *cnc* mutant embryos, *wg* expression is also activated normally and is maintained throughout the beginning of stage 11. During stage 11 *wg* expression in the labral domain ceases in *cnc* mutants (Fig. 6F). In at least some *cnc* mutant embryos, the dorsal pharyngeal *wg* domain is activated (Fig. 6F), but expression in this domain is curtailed by the end of stage 11.

A similar elaboration of *hh* expression domains also occurs in the development of the foregut. In wild-type embryos, a single *hh* foregut expression domain is established at the cellular blastodermic stage (Mohler and Vani, 1992). This *hh* foregut domain persists through stage 10, when it forms a ring around the stomodeal invagination (Fig. 7A). During stage 11, this foregut *hh* domain normally splits to form paired, dorsal pharyngeal expression domains in the ventral lateral region of the clypeolabrum and a circumferential esophageal expression domain (Fig. 7B). The division of this *hh* foregut domain appears to involve inactivation of *hh* transcription in the intervening cells, because in *hh* enhancer trap strains (with *β*-galactosidase under *hh* control) *β*-galactosidase protein persists in the intervening tissue into stage 13 (data not shown). As the esophaguslengthens, the esophageal *hh* domain also lengthens, gradually fading in the middle, until often two distinct expression domains remain at the anterior and posterior ends of the original *hh* esophageal domain at stage 13 (Fig. 7C).

Strong expression of the *hh* foregut domains persists longer than in the trunk segments, such that the fates of these domains can be detected by perdurance of *β*-galactosidase expression in enhancer trap lines under *hh* control. The dorsal pharyngeal *hh* domain develops into a portion of the posterior pharyngeal wall (Fig. 8B). The posterior esophageal domain forms the inner sleeve of the proventriculus (Fig. 8A).

This elaboration of foregut *hh* expression domains is also altered in *cnc* and *fkh* mutant embryos. In *fkh* mutants, at stage 11, the *hh* foregut expression domain resolves into enlarged ventrolateral paired clypeolabral expression domains and no separate esophageal domain forms (Fig. 7D). In *cnc* mutants, the *hh* foregut domain also fails to bifurcate at stage 11, forming only a single long circumferential esophageal domain (Fig. 7E). This esophageal domain then develops normally, forming separate anterior and posterior esophageal domains (Fig. 7F).

In summary, the *fkh* and *cnc* homeotic genes are required for the maintenance (but not the initiation) of segment polarity gene expression in the foregut. *fkh* is required for the continued expression of the esophageal domains of *wg* and *hh* expression past stage 10. *cnc* is required for the continued expression of the dorsal pharyngeal domains of *wg* and *hh* and the labral domain of *wg* past the middle of stage 11.

**DISCUSSION**

**cnc** is a homeotic gene defining labral and mandibular segment identity

The *cnc* gene is expressed continuously in two cephalic regions throughout *Drosophila* embryogenesis (Mohler et al., 1991): the anterior compartment of the mandibular segment and ‘parasegment-4’ comprising the clypeolabral lobe and the presumed labral segment. *cnc* mutant larvae are missing only mandibular and labral structures derived from regions in which *cnc* is normally expressed. In *cnc* mutants, mandibular tissues are transformed to a maxillary fate: a clear example of classical homeosis (Bateson, 1894). In the labral and dorsal pharyngeal region, *cnc* is required for the maintenance of the pattern of segment polarity gene expression similar to the requirement for the homeotic gene *fkh* in the adjacent esophagus. We conclude that *cnc* functions as a homeotic gene in *Drosophila*, namely a segment-specific selector gene which is continuously expressed in those segments (labral and mandibular) in which it is required.

The tissues lost in *cnc* mutant embryos form a set of structures essentially distinct from those affected by other homeotic genes. Moreover, double mutant combinations of *cnc* mutations and those in other homeotic genes reveal an independent additivity of mutant defects. This additivity indicates that *cnc* acts in parallel with the other homeotic genes and does not serve to mediate their actions. This independence of *cnc* action reflects its independent regulation, in which *cnc* transcription is not dependent on any of the other homeotic genes for either activation or maintenance (Mohler, 1993).

**Control of segment identity in the mandibular region**

*cnc* mutant embryos exhibit a partial transformation of the mandibular segment toward maxillary development, as
mutant embryos, mandibular development is also abnormal: the mandibular lateral gräte are malformed, possibly due to transformation into hypopharyngeal ventral arms (see above; Merrill et al., 1987; Regulski et al., 1987). This combinatorial which is normally restricted to the maxillary segment. In mutant embryos, cnc expression is present in the posterior mandibular, maxillary and labial segments, and tsh is expressed in thoracic and abdominal segments. Expression of lab drives intercalary development (IC), whereas Dfd expression promotes development of maxillary structures (MX). Independently regulated expression of cnc in the anterior compartment of parasegment 0 is hypothesized to have two roles: promotion of hypopharyngeal development (HY) in portions of the anterior mandibular segment where Dfd is not expressed (the hypopharyngeal lobe) and promotion of mandibular development (MD) in combination with Dfd in the mandibular lobe, where Dfd is expressed. Segment identity in the Scr expression domain is dependent on the presence of tsh: Scr in the absence of tsh drives labial development (LA); in conjunction with tsh, Scr drives prothoracic development (T1).

evidenced by duplicated maxillary cuticular structures (mouth-hooks and cirri) and the mandibular expression of the maxillary-specific gene, 1.28. In the mandibular segment, expression of cnc overlaps the expression domain of the homeotic gene Dfd. Dfd is expressed in the maxillary segment and the portion of the mandibular segment that includes the mandibular lobe. cnc is expressed in the anterior compartment of the mandibular segment, including the hypopharyngeal lobe and the anterior portion of the mandibular lobe. This overlapping expression of cnc and Dfd appears to direct a combinatorial control of developmental fate in this region. Where Dfd alone is expressed, Dfd drives maxillary development, as occurs normally in the maxillary segment or when Dfd is ectopically expressed in the trunk (Kuziora and McGinnis, 1988). Tissues expressing cnc but not Dfd adopt a hypopharyngeal fate (e.g. T-ribs and ventral arms). Where both cnc and Dfd are expressed, this combination of genes drive the development of mandibular lobe-derived structures, such as the lateral gräte. In cnc mutants, the mandibular lobe fails to follow a mandibular developmental program, adopting instead a maxillary fate. This can be seen in the Dfd-dependent transcription of 1.28 in mandibular lobe of cnc mutant embryos, which is normally restricted to the maxillary segment. In Dfd mutant embryos, mandibular development is also abnormal: the mandibular lateral gräte are malformed, possibly due to transformation into hypopharyngeal ventral arms (see above; Merrill et al., 1987; Regulski et al., 1987). This combinatorial regulation serves to divide the Dfd expression domain into separate mandibular and maxillary primordia, by whether or not cnc is also expressed.

A simple model for control of segment identity in the head may involve combinatorial action by homeobox-containing ANTC homeotic genes in concert with non-homeobox homeotic genes expressed in overlapping domains (Fig. 9). Three ANTC genes, lab, Dfd and Scr, are expressed in domains that overlap three non-homeobox homeotic genes, cnc, sal and tsh. The simplest segmental control is in the intercalary segment, where lab expression, unique to that segment, alone promotes intercalary development. The Dfd expression domain extends over two segments and is divided into mandibular and maxillary specificities, as discussed above, by the presence or absence of cnc: Dfd with cnc driving mandibular development, Dfd without cnc driving maxillary development, while cnc alone in the anteriormost portion of the mandibular segment specifies hypopharyngeal fates. The homeotic gene, sal, while it is also expressed in the posterior mandibular and maxillary segments, does not appear to be required for mandibular or maxillary development (see above; Jürgens, 1988). In the two-segment-wide Scr expression domain, tsh appears to play a similar role in subdividing this region as cnc does in the Dfd domain. tsh is expressed in the trunk of the embryo,
further elaboration of that pattern by subdivision of the foregut hh expression domain at the interface of the cnc and fkh functional domains. Strictly speaking, the cnc and fkh genes do not function as ‘homeotic’ genes in the foregut region; rather they act as region-specific selector genes to maintain their respective primordia through development.

The role of the cnc and fkh homeotic genes to refine further blastoderm segment polarity gene patterns in the foregut might serve, in general, as a prototype for cephalic segmentation. The three preoral segments (intercalary, antennal and preantennal) are formed from the trifurcation during stage 9 of broad wg and hh expression domains established at blastoderm stages from the action of the otd, ems and btd gap genes (Baker, 1987; 1988; Cohen and Jürgens, 1990; Mohler and Vani, 1992; Tashiro et al., 1993; J. Mohler, unpublished results). It is possible that the homeotic genes for this region might function in a process similar to the action of fkh and cnc in the foregut, both to carve out individual segments and to specify their identity. Unfortunately, with the exception of lab, homeotic genes active in this region have yet to be identified.

**cnc is a NF-E2 related protein**

Molecularly, cnc encodes a member of the bZIP class of transcription factors. cnc shares approximately 60% similarity over 160 amino acids, about half the protein, with the p45 subunit of NF-E2, an erythroid-specific transcription factor regulating hemoglobin synthesis (Andrews et al., 1993a). This homology includes 95% identity in the 20 amino acid DNA-binding domain. The long 160 a.a. extent of similarity between NF-E2 p45 and cnc is especially remarkable when compared to the degree of similarity between conserved mammalian and Drosophila bZIP proteins, such as jun and fos (Hurst, 1994; Perkins et al., 1990), where the similarity is limited to just a few amino acids either side of the DNA-binding domain. NF-E2 p45 belongs to a small family of closely related mammalian proteins, which share equal similarity to cnc and each other, including nrf1 and h26A, both of unknown function (Chan et al., 1993a,b). It is not known whether NF-E2 or any of the other members of the NF-E2 family may serve an analogous role to cnc in vertebrate embryonic development.

The high degree of similarity suggests that cnc may function similarly to NF-E2, biochemically. Functional NF-E2 transcription factor, which binds to an extended AP-1 site (Andrews et al., 1993a), is a dimer of NF-E2 p45 and one of any of a number of small maf-related bZIP proteins (mafK, mafF, and mafG; Andrews et al., 1993b; Fujinata et al., 1993; Igarashi et al., 1994). It is tempting to speculate that the involvement of cnc and related proteins in differential heterodimers, as appears to occur among mammalian bZIP proteins, may provide differential developmental specificities (e.g. labral vs. mandibular) within the regions cnc is expressed.

The authors thank Vangga Ly and Elizabeth Bonwith for help in making cnc deletions and Allan Spradling for providing the cnc^- strain. This work was supported by grants from the National Institutes of Health and the National Science Foundation.

**REFERENCES**


(Accepted 19 October 1994)