Development 125, 4553-4564 (1998)
Printed in Great Britain © The Company of Biologists Limited 1998
DEV5229

A cap ‘n’ collar protein isoform contains a selective Hox repressor function

Nadine McGinnis, Erlend Ragnhildstveit*, Alexey Veraksa and William McGinnis‡

Department of Biology, University of California, San Diego, La Jolla, CA, 92093, USA
*Present address: Biotechnology Centre of Oslo, PO Box 1125, N-0316 Oslo, Norway
‡Author for correspondence (e-mail: wmcginnis@ucsd.edu)

Accepted 11 September; published on WWW 20 October 1998

SUMMARY

We have characterized a protein isoform (CncB) from the *Drosophila cap ‘n’ collar* locus that selectively represses *cis*-regulatory elements that are activated by the Hox protein Deformed. Of the three Cnc protein isoforms, CncB is expressed in a localized pattern in mandibular and labral cells of the head during mid-stages of embryogenesis. When CncB protein is absent or reduced, mandibular cells are homeotically transformed toward maxillary identities. This transformation is associated with persistent *Deformed* expression in anterior mandibular cells, since the *Deformed* autoactivation circuit is normally antagonized by CncB function in these cells. Heat-shock-induced ectopic expression of CncB in mid-stages of embryogenesis is sufficient to attenuate the activity of Dfd response elements in maxillary epidermal cells, but appears to have no effect in trunk epidermal cells on either the function or the response elements of other Hox proteins. CncB provides a mechanism to modulate the specificity of Hox morphogenetic outcomes, which results in an increase in the segmental diversity in the *Drosophila* head.

Key words: *cap ‘n’ collar, Deformed*, Homeotic, Hox repressor, *Drosophila*

INTRODUCTION

The homeotic selector (Hox) genes are important components in a developmental pathway that diversifies morphology on the anterior-posterior body axis of many animals (Kenyon, 1994; Lawrence and Morata, 1994; Manak and Scott, 1994). Loss-of-function mutations in Hox genes, or their ectopic expression, can cause homeotic transformations in which body structures are duplicated in abnormal positions on the anterior-posterior axis (McGinnis and Krumlauf, 1992; Krumlauf, 1994). The protein products of Hox genes are DNA-binding transcriptional regulators that apparently accomplish their developmental functions by activating and repressing many target genes (Laughon, 1991; Botas, 1993; Graba et al., 1997).

The similar in vitro DNA-binding properties of Hox proteins (Ekker et al., 1994) and the relative paucity of Hox genes compared to the diverse morphologies under their control, suggest that other genetic functions that act in parallel to Hox proteins play a crucial role in diversifying segmental morphology. Recently, a few *Drosophila* genes (*extradenticle, teashirt, homothorax, cap ‘n’ collar*) have been discovered that can mutate to give homeotic transformations that are largely independent of their effects on *Hox* patterns of transcription (Peifer and Weischaus, 1990; Roder and Kerridge, 1992; Mohler et al., 1995; Rieckhof et al., 1997). These genes appear to modulate the functional activity and specificity of the Hox proteins, thus we refer to them as the Hox modulator class. The understanding of Hox modulator functions is beginning to provide insight into the mechanisms of how Hox proteins achieve functional specificity, as well as broadening our knowledge of the developmental genetic circuitry that controls anterior-posterior body axis patterning. The modulator proteins so far characterized play a variety of biochemical roles, including enhancing the DNA-binding affinity of Hox proteins (Mann and Chan, 1996), regulating their transcriptional activities (Pinsonneault et al., 1997) and regulating the nuclear entry of Hox cofactors (Rieckhof et al., 1997).

An interesting example of the Hox modulator class is encoded in the *cap ‘n’ collar* (*cnc*) locus. *cnc* was identified and named (Mohler et al., 1991) based on its striking expression pattern in the anteriormost labral segment (cap) and the mandibular segment (collar) of embryos. Deletion mutants of *cnc* coding sequences indicate that *cnc* functions are required for the normal development of both labral and mandibular structures (Mohler et al., 1995). In place of the missing mandibular structures, some maxillary structures – mouth hooks and cirri – are ectopically produced (Harding et al., 1995; Mohler et al., 1995). The genetic function of the homeotic gene *Deformed (Dfd)* is required in the *cnc* mutant background to produce ectopic mouth hooks, and Mohler et al. (1995) have proposed that *Dfd* and *cnc* function in combination to specify mandibular identity. Previous research on the molecular genetics of *cnc* (Mohler et al., 1991) showed that the locus encodes a protein that is one of the founding members of the CNC/bZIP class of transcription factors, which includes NF-E2 (Andrews et al., 1993a), LCR-F1 (Farmer et al., 1997), Skn-1 (Bowerman et al., 1992) and Nrf-1 (Chan et al., 1993).

We recently isolated EMS-induced mutations in the *cnc* gene

in a genetic screen for modifiers of *Dfd* (Harding et al., 1995) and are interested in the mechanism by which *cnc* alters the morphogenetic function of *Dfd*. To that end, we have further explored the molecular stucture of the *cnc* locus and found that three different protein isoforms are encoded by *cnc*. One of these isoforms, the CncB protein, exhibits spatially localized expression that is limited to the mandibular and labral segments. In mandibular cells, CncB antagonizes the ability of Dfd protein to transcriptionally activate response elements of downstream genes. Ectopic expression of CncB in embryos results in an ablation of normal maxillary structures, while having mild or no effects on the functions of homeotic proteins of the trunk. We propose that the CncB protein has properties that allow it to selectively repress many regulatory elements that are normally activated by Dfd.

MATERIALS AND METHODS

Embryonic cuticle preparations

Collections were made from cages containing flies heterozygous for mutant chromosomes over wild-type chromosomes to avoid balancer chromosome effects. Embryos were collected on apple juice/agar plates for 4 hours, aged at 25°C for 24 hours, then harvested and dechorionated in 100% bleach. They were devitellinized, cleared, mounted and analyzed as described in Harding et al. (1995).

Northern analysis

Polyadenylated RNA was isolated from 0-2, 2-8, 8-12 and 12-24 hour embryos and 4 μg of each staged aliquot was loaded onto a pre-run 0.8% agarose formaldehyde gel. The RNA was blotted onto nitrocellulose filters and probed as in Sambrook et al. (1989). Radioactively labeled probes were made by the method of nick translation from the following DNA fragments. (i) *cncA* probe, a 900 bp *Xho*I-*Eco*RI fragment from a genomic clone that includes exon A1. (ii) *cncB* probe, a 1.1 kb *Bst*XI-*Eco*RV fragment from the cDNA clone pNBcnc27, which was isolated from a 8-12 hour embryonic cDNA library (Brown and Kafatos, 1988). The *cncB* probe includes part of exon B1, all of exons B2 and B3, and most of exon B4. (iii) *cncB*-specific probe, a 2.2 kb *Eco*RI genomic fragment in clone pBstB1spec, which includes exon B1 sequences and no other B exon sequences. (iv) *cncC* probe, a 2.3 kb *Bam*HI-*Eco*RI fragment from the cDNA clone pNBcnc23, which was isolated from a 8-12 hour embryonic cDNA library (Brown and Kafatos, 1988). This probe contains part of exon C1, all of exons C2 and C3 and most of exon C4. (v) The *cnc* common probe is a 310 bp *Bam*HI fragment entirely contained within exon A2, which was isolated from the cDNA clone cnc1A10 (Mohler et al., 1991). See Fig. 2 for the location of the probes on the map of the *cnc* locus.

Library screens

Several *D. melanogaster* libraries were screened for both cDNA and genomic clones. The 4-8 hour, 8-12 hour and 12-24 hour embryonic cDNA libraries of Brown and Kafatos (1988) were screened using the 310 bp *Bam*HI fragment from exon A2. The 8-12 hour library was also screened with an 83 bp PCR product made solely of exon A1 and the 4-8 hour and 12-24 hour libraries were additionally screened with a PCR fragment covering the 5′ end of exon B2 to the middle of exon B4. An *iso-1* genomic library consisting of *Sau*3A partial digests ligated into λEMBL3 (Tamkun et al., 1992) was screened with a 2.3 kb *Bam*HI-*Eco*RI fragment from pNBcnc23, a *cncC* cDNA clone. A randomly primed λgt10 cDNA library (Clontech, Cat # IL1010a) was screened with a 4 kb *Eco*RV-*Hin*dIII fragment from pNBcnc27, which includes sequences from within exon B2 to beyond the translation stop. All potential positives obtained from the primary screens were picked, replated at lower density and rescreened according to standard protocols to ensure plaque/colony purity. PCR screening was employed as a rapid method to detect and characterize the structure of cDNA clones containing different exons and/or to assay the extent of 5′ exon sequence.

Genomic DNA sequencing

For two mutant alleles, cnc^{2E16} and cnc^{C7}, and the parental chromosome *Ki* Dfd^{rV8} p^{p}, DNA sequence was obtained for chromosomal regions corresponding to known exons, as well as all exon/intron boundaries including 40-50 bp beyond each boundary into the introns. All mutant and parental chromosomes were heterozygous and balanced over *TM3 Sb 35UZ* (Irvine et al., 1991). DNA was extracted from 1 g of flies according to standard protocols and isolated on a CsCl gradient. Regions of the genome were PCR amplified using nested primers chosen from cDNA and genomic sequences. Primers were made approximately 100 bp from intron/exon boundaries so that all boundaries could be sequenced from both directions. The PCR products were then either sequenced using ^{35}S and the Sequenase PCR Product Sequencing Kit from Amersham (Catalogue no. US70170) or they were agarose gel purified and sequenced according to kit protocols without the use of Exonuclease I and Shrimp Alkaline Phosphatase. At least one strand was sequenced for 100% of the regions, both strands were sequenced for the majority of the open reading frames and 7-deaza dGTP reactions were used to uncover any ambiguities due to compression. No nucleotide substitutions were present in any *cnc* open reading frames or splice junctions of the cnc^{2E16} and cnc^{C7} mutant chromosomes when compared to the parental chromosome sequence.

In situ detection of transcript and protein expression

Polyclonal antiserum was raised against the common Cnc protein domain produced as a glutathione S-transferase (GST)-fusion protein. A cDNA clone (pNBcnc10) containing a full-length CncA open reading frame was digested with *Bst*EII and *Sac*I and blunt ended. The resulting 1721 bp fragment was ligated into the *Sma*I site of pGEX-4T-3 (Pharmacia Biotech) to produce an in-frame fusion with GST. The GST-Cnc common region polypeptide was purified on glutathione-Sepharose columns according to the manufacturer's protocol, and the antiserum raised in rabbits at the Pocono Rabbit Farm and Laboratory, Inc.

For affinity purification of antibodies, pNBcnc10 was digested with *Hin*cII and *Sac*I and the resulting 1827 bp fragment that included the entire CncA open reading frame was ligated into the pQE-32 vector (Qiagen) which had been digested with *Bam*HI, blunt-ended and then cut with *Sac*I. The fusion protein containing a 6xHis-tag at the N terminus was purified on the Qiagen Ni-NTA Agarose column under denaturing conditions and refolded. The protein was coupled to Actigel ALD beads and antibodies directed at the common Cnc protein domain were purified using the Quickpure system (Sterogene Bioseparations, Inc.). Antibody staining for Dfd and Cnc proteins were done as described in Zeng et al. (1994). For fluorescence microscopy, FITC-conjugated anti-rabbit secondary antibodies were used to detect Cnc protein and Cy-5-conjugated anti-guinea pig antibodies were used to detect Dfd protein (Jackson ImmunoResearch Laboratories, Inc.). Optical sections of fixed and stained embryos were taken every 0.2 μm using a DeltaVision microscope system (Applied Precision, Inc.) with a computer-controlled stage. An Olympus 60×/1.40 objective was used. Following image acquisition, out-of-focus blur was removed using constrained iterative deconvolution (Agard et al., 1989).

For RNA in situ staining, sense and antisense digoxigenin-labeled RNA probes were produced from subclones of cDNA or genomic fragments based on Tautz and Pfeifle (1989). The *lacZ* reporter and other gene expression patterns were detected by staining of whole-mount embryos as described in Bergson and McGinnis (1990).

hsp70-*cnc* cDNA fusion genes

hs-cncA

A *cncA* cDNA clone, 1A10 (Mohler et al., 1991) was digested with

*Eco*RI and cloned into the *Eco*RI site of pHSBJ-CaSpeR (Jones and McGinnis, 1993). The resulting pHSBJ-CaSpeR-*cncA* contained 100 bp of 5′ UTR and 800 bp of 3′ UTR from exon A3.

hs-cncB

pNBcnc27 cDNA was digested with *Eco*RV and ligated with *Eco*RV-digested pHSBJ to produce pHST17. The heat-shock cassette was cut out with *Not*I and inserted into the *Not*I-digested pCaSpeR4 vector to produce pHSBJ-CaSpeR4-*cncB*. This construct contained 390 bp of 5′ UTR and 1230 bp of 3′ UTR from exon A3.

hs-cncC

pNBcnc23 cDNA was digested with *Hin*dIII and *Nco*I, as was the pHST17 plasmid. Corresponding DNA fragments were ligated to produce pHSBJ-*cncC*. The resulting cassette was cut out with *Not*I and ligated with *Not*I-digested pCaSpeR4 to create pHSBJ-CaSpeR4-*cncC*. The first ATG codon in this *cncC* expression construct is 37 bp from the *Hin*dIII site and it encodes the methionine residue found at position 21 in the conceptual translation of the *cncC* open reading frame shown in Fig. 3. This construct also contains 1230 bp of the normal *cnc* 3′ UTR.

RESULTS

The *cap 'n' collar* locus encodes three protein isoforms

We recovered three EMS-induced mutant alleles of *cnc* (*cnc* ^{2E16}, *cnc*^{C7} and *cnc*^{C14}) in a screen for mutations that interact with the Hox gene *Dfd* (Harding et al., 1995). Embryos homozygous for these EMS-induced alleles have ectopic duplications of maxillary mouth hooks and cirri, but retain normal labral structures and some normal mandibular structures, e.g. the lateralgräten and median tooth (Fig. 1). This contrasts with the phenotype of deletion mutants of *cnc*, which lack all mandibular and labral derivatives (Mohler et al., 1995). The difference between the phenotypes of the EMS-induced alleles when compared to the deletion alleles prompted us to consider the possibility that multiple functions are encoded in the *cnc* locus.

Previous studies detected one transcript isoform at *cnc* (Mohler et al., 1991), but our molecular analyses of the locus indicate that three transcript and protein isoforms are produced from the *cnc* gene. As shown in Fig. 2C, a probe homologous to the region that encodes the b-ZIP region of *cnc* detects three different sizes of polyadenylated RNAs on embryonic northern blots. We will refer to these as the *cncA*, *cncB* and *cncC* transcripts. No other embryonic transcripts were detected with genomic probes that spanned the region from −35 to +5 kb shown in Fig. 2A. The 3.3 kb *cncA* transcript is present in 0-2 hour embryos, presumably from maternal stores and is also abundantly expressed in 12-24 hour embryos. The 5.4 kb *cncB* transcript is absent from 0-2 hour embryos, but present at all other embryonic stages. The 6.6 kb *cncC* transcript is present in 0-2 hour embryos, is barely detected in 2-12 hour embryos and is detected at relatively higher levels in 12-24 hour embryos (Fig. 2C).

To identify cDNAs corresponding to the *cncA*, *cncB* and *cncC* transcripts, 212 cDNA clones from libraries covering all stages of *Drosophila* embryonic development were isolated and characterized. The first class of cDNAs corresponded to the *cncA* transcript. This is the same class characterized by Mohler et al. (1991), and is distinguished by the incorporation of exon A1. Exons A2 and A3, which encode the CNC and b-ZIP domains, are present in *cncA* and the other two isoforms of *cnc*. A probe containing exon A1 sequences specifically hybridizes the 3.3 kb *cncA* transcript on northern blots (Fig. 2C). The *cncA* open reading frame begins with an ATG codon near the 5′ end of exon A2 and is predicted to encode a 533 amino acid protein (Mohler et al., 1991).

A second class of cDNAs from the locus corresponded to *cncB* transcripts. Such cDNAs lacked sequences from exon A1, but did contain five additional exons (B1-B5) spliced onto the 5′ end of exon A2. A probe containing the B1-B4 exons detects the 5.4 kb *cncB* transcript and the 6.6 kb *cncC* transcript on northern blots (Fig. 2C). The total extent of the *cncB* transcription unit is approximately 17 kb (Fig. 2B). Since exon A2 sequences contain no stop codons upstream of the initiating ATG for the CncA codons, the open reading frame in *cncB* transcripts includes the entirety of the CncA protein, as well as an additional 272 codons from exons B3, B4, B5 and A2 (Fig.

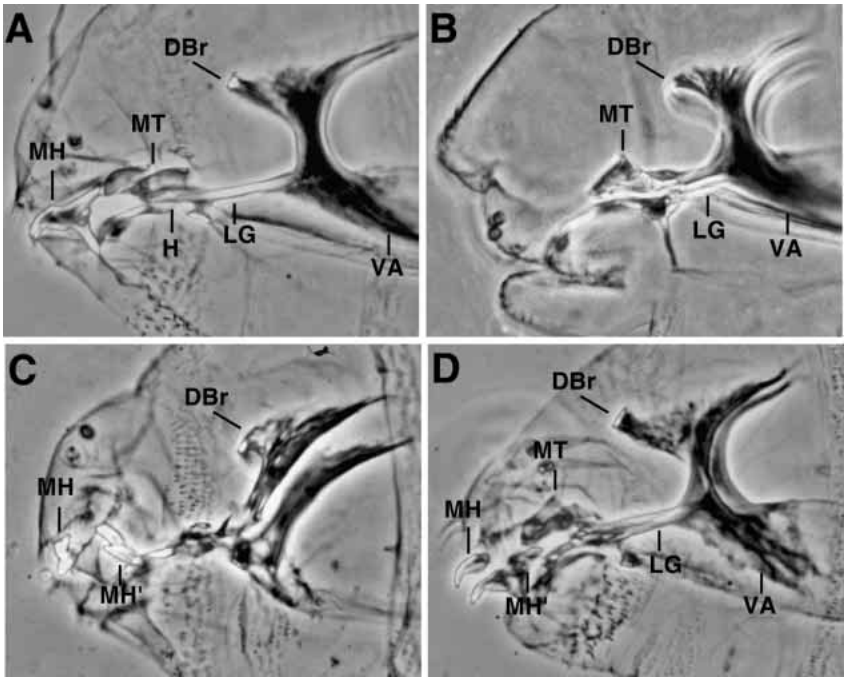

Fig. 1. *cap 'n' collar* mutant phenotypes. (A) Wild-type embryonic head cuticle. MH, mouth hooks; MT, median tooth; LG, lateralgräten; DBr, dorsal bridge; H, H-piece; VA, ventral arms. (B) *Dfd*^{w21}/*Dfd*^{w21} mutant cuticle. Mouth hooks and cirri, which are of maxillary origin, are missing and the lateralgräten, of mandibular origin, are shortened. (C) *cnc*^{VL110}/*cnc*^{VL110} mutant. The median tooth (labral structure) and lateralgräten are both absent and ectopic mouth hooks (MH′) are formed in the pharynx. (D) *cnc*^{2E16}/*cnc*^{2E16} mutant. Ectopic mouth hooks and cirri are formed, but the median tooth develops normally. The lateralgräten are shortened slightly, resembling those seen in *Dfd* mutants.

2B). The predicted 805 amino acid CncB protein thus is distinguished from CncA by a 272 amino acid region that includes His-Pro repeats, Ala-repeats, a Pro-repeat and Val-Gly repeats, but exhibits no extended sequence similarity to other proteins in database searches besides the CNC/b-ZIP domain that it shares with CncA (Fig. 3).

The third class of cDNAs from the locus corresponded to *cncC* transcripts. These cDNAs have identical sequence to the *cncB* cDNAs, except that exon B1 is absent, and five additional exons (C1-C5) are spliced onto the 5′ end of exon B2. A probe containing the C1-C4 exons detects the 6.6 kb *cncC* transcript on northern blots (Fig. 2C). Since exon B2 and the 5′ end of exon B3 contain no stop codons upstream of the initiating ATG for the CncB codons, the ATG-initiated open reading frame in *cncC* transcripts includes the entirety of the CncB protein, as well as an additional 491 codons that derive from the C3, C4, C5, B2 and B3 exons. The extent of the entire *cncC* transcription unit is approximately 39 kb. The 491 amino acid CncC-specific domain at the N terminus of the predicted 1296 residue CncC protein includes regions that are rich in Ser and Thr residues, other regions with abundant concentrations of Glu and Asp residues, but exhibits no extended sequence similarity to other proteins in database searches (Fig. 3). Interestingly, the *fuzzy onions* gene, which encodes a testis protein required for mitochondrial fusion in *Drosophila* spermatids (Hales and Fuller, 1997), is encoded in the sequence interval between the C5 and B1 exons (Fig. 2).

Expression patterns of the *cnc* isoforms

We next were interested in defining which of the *cnc* isoforms were involved in modulating *Dfd* function. As a first step, the expression patterns of the three transcript isoforms were analyzed using *cncA*, *cncB*, or *cncC* exon-specific probes both on wild-type and our EMS-induced *cnc* mutants. In wild-type embryos, a *cncB* probe detects cytoplasmic transcripts limited to the mandibular and labral segments from cellular blastoderm to the end of embryogenesis (Fig. 4A,B). The *cncB* transcripts are expressed throughout both anterior and posterior regions of the mandibular lobes. In contrast, a *cncA*-specific probe detects a ubiquitous distribution of presumably maternal RNA at syncytial and early cellular blastoderm stages (Fig. 4F,G). After cellular blastoderm, *cncA* transcripts are not detectable until stage 14, when the level of ubiquitous cytoplasmic transcript increases in abundance and remains high for the remainder of embryogenesis. *cncC*-specific probes also detect a ubiquitous distribution of RNA in syncytial stage embryos and a low level ubiquitous expression pattern in embryos after stage 14 (data not shown).

Based on the above results, the labral and mandibular stripes of transcription that were detected by Mohler et al. (1991) using a probe including the *cnc* common exons (A2 and A3) correspond primarily to *cncB* transcripts. Since *cncB* is the transcript isoform that is expressed throughout the entire mandibular segment during mid-embyronic stages, our working hypothesis is that *cncB* encodes the principal function that modulates *Dfd* function in the mandibular segment. To further test this hypothesis, we assayed whether *cncB* transcript or protein abundance was altered in embryos homozygous for the cnc^{2E16} and cnc^{C7} mutations. We found that the pattern of zygotic RNA expression detected with a *cncB* probe is unaltered in the EMS-induced *cnc* mutant embryos (Fig. 4C). The signal due to cncA and cncC transcripts was also unchanged in these mutants (data not shown).

However, the use of polyclonal antiserum raised against the common domain of the *cnc* isoforms (anti-Cnc) indicates that CncB protein expression is strikingly reduced in both cnc^{2E16} and cnc^{C7} mutant embryos. In wild-type embryos, the anti-Cnc antiserum exhibits a low-level ubiquitous staining in syncytial embryos, presumably due to maternally deposited CncA and

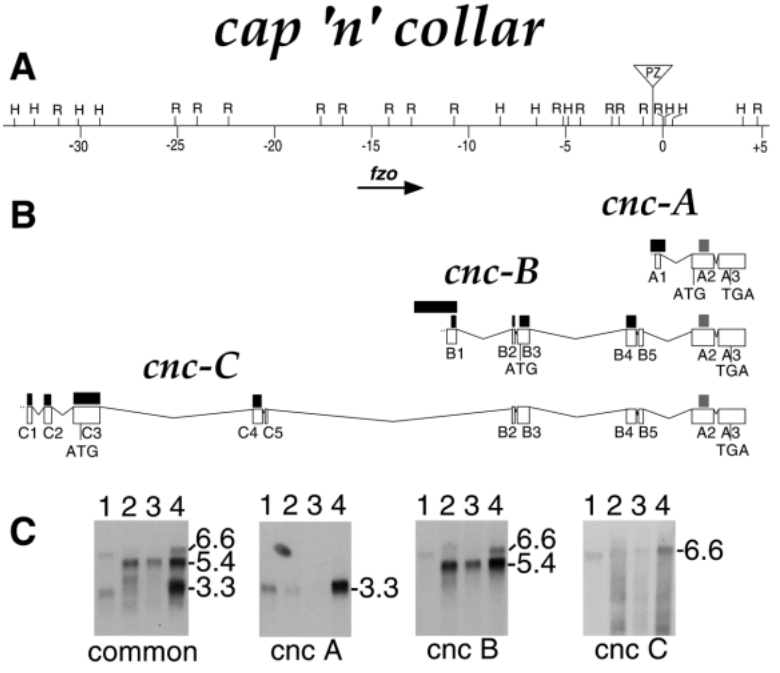

Fig. 2. *cap 'n' collar* molecular genetics. (A) A structural map of the 94E genomic region that contains the *cnc* locus. The location of the *fuzzy onions* transcription unit (Hales and Fuller, 1997) is indicated. H, *Hin*dIII; R, *Eco*RI. (B) Exon maps of three *cnc* transcript isoforms. Open boxes indicate the exons of each transcript and are numbered accordingly. The black boxes above the exons indicate the origins of the *cnc* isoform probes used for northerns and in situ hybridizations; the grey shaded box above exon A2 indicates the origin of the *cnc* common probe. The positions of start (ATG) and stop (TGA) codons for the open reading frames in each isoform are indicated. (C) Embryonic developmental northerns probed with the common *cnc* probe and the isoform probes for *cncA*, *cncB* and *cncC*. The northern contains polyadenylated RNA from: lane 1, 0-2 hour; lane 2, 2-8 hour; lane 3, 8-12 hour; lane 4, 12-24 hour embryos. The common probe hybridizes RNAs of sizes 6.6 kb, 5.4 kb and 3.3 kb, representing *cncC*, *cncB* and *cncA*, respectively. The *cncA*-specific probe hybridizes the 3.3 kb transcript, which is most abundant at 12-24 hours and at lower abundance in 0-2 hour embryos. The *cncB* probe hybridizes the 5.4 kb transcript, detected in embryos from 2-24 hours. The *cncB* probe used in this panel also detects the *cncC* transcript (6.6 kb) since it shares exons B2, B3, B4 and B5 with *cncC*. When a *cncB*-specific probe is used, consisting of a 2.2 kb *Eco*RI genomic fragment in clone pBstB1spec, which includes exon B1 sequences and no other B exon sequences, only the 5.4 kb transcript is detected (data not shown). The *cncC*-specific probe hybridizes the 6.6 kb transcript, which is detected in 0-2 hour embryos and in 12-24 hour embryos.

```
*Cnc-C                                                                                                  100
MANGIGGCKLPPRFNGSTFVMNLHN TTGNSSVQTAALQDVQSTSAAATGA TMVVGTGGAPTSSGQTSGSALGEIH IDTASLDPGNANHSPLHPTSELDTF
                                                                                                        200
LTPHALQDQRSIWEQNLADLYDYND LSLQTSPYANLPLKDGQPQPSNSSH LDLSLAALLHGFTGGSGAPLSTAAL NDSTPHPRNLGSVTNNSAGRSDDGE
                                                                                                        300
ESLYLGRLFGEDEEEDYEGELVGGV ANACEVEGLTTDEPFGSNCFANEVE IGDDEEESEIAEVLYKQDVDLGFSL DQEAIINGSYASGNSAATNVKSKPE
                                                                                                        400
DETKSSDPSISESSGFKDTDVNAEN EASAASVDDIEKLKALEELQQDKDK NNENQLEDITNEWNGIPFTIDNETG EYIRLPLDELLNDVLKLSEFPLQDD
                                                                                                  *Cnc-B
LSNDPVASTSQAAAAFNENQAQRIV SETGEDLLSGEGISSKQNRNEAKNK DNDPEKADGDSFSVSDFEELQNSVG SPLFDLDEDAKKELDEMLQSTVPSY
                                                                                                        600
HHPHPHHGHPHAHPHSHHHASMHHA HAHHAAAAAAAHQRAVQQANYGGGV GVGVGVGVGVGSGTGSAFQRQPAAG GFHHGHHQGRMPRLNRSVSMERLQD
                                                                                                        700
FATYFSPIPSMVGGVSDMSPYPHHY PGYSYQASPSNGAPGTPGQHGQYGS GANATLQPPPPPPPPHHAAMLHHPN AALGDICPTGQPHYGHNLGSAVTSS
                                                                                  *Cnc-A                800
MHLTNSSHEADGAAAAAAAYKVEHD LMYYGNTSSDINQTDGFINSIFTDE DLHLMDMNESFCRMVDNSTSNNSSV LGLPSSGHVSNGSGSSAQLGAGNPH
                                                                                                        900
GNQANGASGGVGSMSGSAVGAGATG MTADLLASGGAGAQGGADRLDASSD SAVSSMGSERVPSLSDGEWGEGSDS AQDYHQGKYGGPYDFSYNNNSRLST
                                                                                                       1000
ATRQPPVAQKKHQLYGKRDPHKQTP SALPPTAPPAAATAVQSQSIKYEYD AGYASSGMASGGISEPGAMGPALSK DYHHHQPYGMGASRSAFSGDYTVRP
                                                                                                       1100
SPRTSQDLVQLNHTYSLPQGSGSLP RPQARHKKPLVATKTASKGASAGNS SSVGGNSSNLEEEHLTRDEKRARSL NIPISVPDIINLPMDEFNERLSKYD
                                                                                                       1200
LSENQLSLIRDIRRRGKNKVAAQNC RKRKLDQILTLEDEVNAVVKRKTQL NQDRDHLESERKRISNKFAMLHRHV FQYLRDPEGNPCSPADYSLQQAADG
                                                                                                       1296
SVYLLPREKSEGNNTATAASNAVSS ASGGSLNGHVPTQAPMHSHQSHGMQ AQHVVGGMSQQQQQQSRLPPHLQQQ HHLQSQQQQPGGQQQQQHRKE*
```

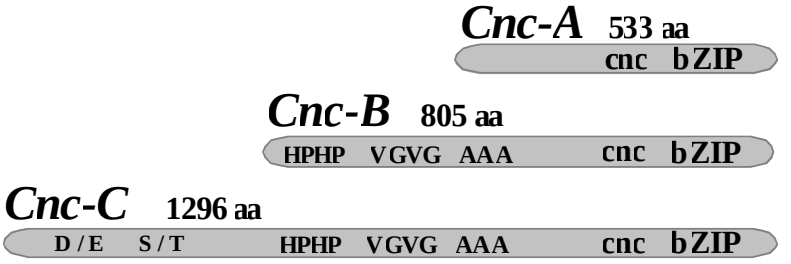

Fig. 3. Cnc protein sequences. The predicted protein sequences of the three cnc isoforms are shown. The presumed initiating methionine for each is indicated by a bold **M** below an asterisk that denotes the N termini of the CncA, CncB and CncB proteins. The basic-leucine zipper (bZIP) consensus domain is underlined. CncB and CncC both contain His-Pro repeats (HP), Val-Gly repeats (VG) and Ala-repeats (AAA) in the sequence they share. The N-terminal amino acid sequence that is unique to CncC is enriched in scattered Ser and Thr (S/T) residues, and also contains a high proportion of acidic residues, including a Asp-Glu repeat (D/E). GenBank accession numbers for the cDNA sequences that provided the predicted protein sequences shown in this figure are AF070062 for *cncA*, AF070063 for *cncB*, AF070064 for *cncC*.

CncC isoforms. From cellular blastoderm (stage 5) until stage 14, the staining detected by the anti-Cnc antiserum is localized in the nuclei of mandibular and labral cells (Fig. 4D). Although the anti-Cnc antiserum used in these experiments cross-reacts with all three Cnc proteins, only *cncB* RNA expression is localized in mandibular and hypopharyngeal regions from stages 6 through 14. From stage 14 to the end of embryogenesis, the antiserum detects a low level global staining, upon which is superimposed much stronger levels of staining in mandibular and labral cells. As can be seen in Fig. 4E, cnc^{2E16} mutants (and cnc^{C7} mutants, not shown) accumulate much lower levels of Cnc antigen in both mandibular and labral cells of stage 11 embryos. These results provide further evidence that the cnc^{2E16} and cnc^{C7} mutations result in a loss of *cncB* function, and is consistent with the idea that CncB protein is required to prevent the maxillary-promoting function of Dfd from being active in mandibular cells.

We determined the sequence of all of the coding exons and exon/intron boundaries for all isoforms on the cnc^{2E16} and cnc^{C7} mutant chromosomes (see Materials and Methods) in an attempt to find the molecular lesion responsible for the decreased amount of CncB protein in the mutant embryos. However, no nucleotide substitutions were detected when the coding and splice site sequences were compared with parental chromosome sequence. Though we do not yet know the location of the mutations that alter CncB protein expression, they could plausibly reside in translational regulatory sequences for *cncB*.

Heat-shock phenotypes of *cnc* isoforms

In another test of the functions of the Cnc protein isoforms, we placed each of the *cncA*, *cncB* and *cncC* open reading frames under the control of the heat-shock promoter in P-element vectors and generated transgenic fly strains carrying these constructs. Using the Cnc common-region antiserum to stain heat-shocked embryos, it appears that all three isoforms are produced at similar levels, localized in nuclei and possess similar stabilities after ectopic expression (Fig. 5C,E,G). However, their morphogenetic and regulatory effects are very dissimilar. Heat-shock-induced ectopic expression of CncA during embryogenesis has no effect on embryonic morphology. Nearly all of the *hs-cncA* embryos hatch and proceed through larval development, and many eclose as viable adults.

In contrast, ectopic expression of CncB at mid-stages (4-10 hours) of embryonic development is lethal. When ectopic expression is induced at 6 to 8 hours after egg lay, a defective embryonic head phenotype, which resembles the mutant phenotype of strong *Dfd* hypomorphs is produced (Fig. 5I, compare with Fig. 1B). These *hs-cncB* embryos develop with rudimentary mouth hooks, H-piece and cirri. In addition, the anterior portion of the lateralgräten are truncated. All of these structures are components of the head skeleton that are absent or abnormal in *Dfd* mutant embryos (Merrill et al., 1987; McGinnis et al., 1990). The head defects seen in the *hs-cncB* embryos also include an absent or abnormal dorsal bridge, a structure that is usually unaffected in *Dfd* mutant embryos. Many other head structures that develop in a *Dfd*-independent

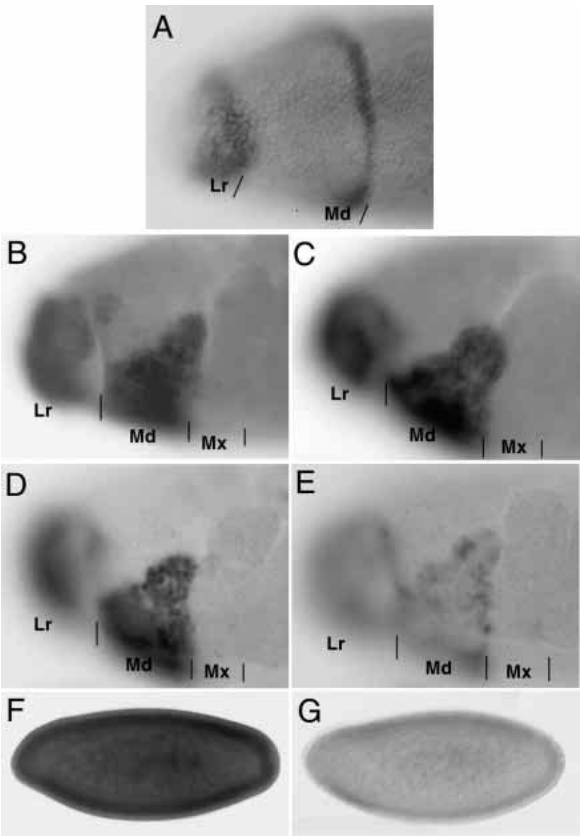

Fig. 4. *cnc* expression patterns in wild type and mutants. (A) Dorsal view of a stage 5 wild-type embryo showing the earliest RNA expression detected with a *cncB*-specific RNA probe. The cells at the anterior tip (left) of the embryo are labral progenitors, the cells in the more posterior stripe are primordia of the mandibular segment. The *cncB*-specific probe consisted of a 2.2 kb *Eco*RI genomic fragment in clone pBstB1spec, which includes exon B1 sequences and no other B exon sequences. (B) A stage 11 wild-type embryo hybridized with the *cncB* probe. Cytoplasmic transcript signals are abundant in mandibular (Md) and labral cells (Lr). (C) A stage 11 cnc^{2E16}/cnc^{2E16} mutant embryo hybridized with the *cncB*-specific probe. The pattern and abundance of transcript signal is indistinguishable from wild-type embryos as in B. (D) A stage 11 wild-type embryo incubated with the common Cnc antibody. Cnc antigen is abundant in mandibular and labral nuclei. (E) A stage 11 cnc^{2E16}/cnc^{2E16} mutant embryo incubated with the common Cnc antibody. Protein staining is severely reduced in all mandibular and labral nuclei compared to wild-type embryos. Since *cncA* and *cncC* transcripts are not detected at this stage, we believe the reduced protein levels are due to a reduction of CncB levels in these homozygous mutants. (F) A stage 5 wild-type embryo hybridized with the *cncA*-specific RNA probe. *cncA* transcripts are ubiquitously distributed in the cytoplasm of embryos at this stage. (G) A stage 5 wild-type embryo hybridized with a *cncA* sense control probe.

manner, such as the antennal sense organ, vertical plates and T-ribs (Jurgens et al., 1986) develop normally in the *hs-cncB* embryos. The *hs-cncB* head defects are produced at high penetrance (>95%) by heat shocks in mid-embryogenesis (4-10 hours). In 10-70% of these embryos, depending on the stage of heat shock, abdominal denticles near the ventral midline are replaced with naked cuticle (Fig. 5L). We have not observed the ectopic formation of mandibular or hypopharyngeal structures such as lateralgräten, dorsolateral papillae or T-ribs (Jurgens et al., 1986) in other head or trunk segments in *hs-cncB* embryos.

Ectopic induction of *hs-cncC* at 6-8 hours of development also results in highly penetrant defects in head development that include the loss of maxillary mouth hooks and cirri (Fig. 5J) as well as head involution defects that are more profound than those induced by *hs-cncB*. In addition to the morphological defects described for CncB, ectopic CncC induces the formation of an abnormal head sclerite that develops as an extension of the normal lateralgräten. The position and appearance of this extra fragment of head skeleton suggests that it might correspond to ectopic production of lateralgräten or longitudinal arms of the H-piece.

Effects of overexpression of CncB on downstream targets

Since CncB encodes a function that is required and sufficient to antagonize the maxillary-promoting effects of the Hox gene *Dfd*, we next were interested in whether CncB protein acts upstream to repress *Dfd* transcription, or in parallel to inhibit Dfd protein function. It is possible for CncB to do both, since Dfd protein function is required to establish an autoactivation circuit that provides persistent *Dfd* transcription in maxillary and mandibular cells (Zeng et al., 1994). In wild-type embryos at stage 9, both Dfd and CncB proteins are expressed throughout the entire mandibular segment (Fig. 6A). By stage 11, Dfd protein is present at lower levels in the anterior when compared to posterior mandibular nuclei, while CncB protein persists at relatively high levels throughout the segment (Fig. 6B). Finally, at stage 13, Dfd protein expression is no longer detected in anterior mandibular nuclei, although it is still abundant in posterior nuclei (Fig. 6C). *cnc* is required for this progressive repression of *Dfd* expression in the anterior mandibular segment, since *cnc* null mutants as well as the EMS-induced mutants show inappropriate persistence of *Dfd* transcripts and protein after stage 11 in anterior mandibular cells (Fig. 6E). All of this data suggests that CncB is not capable of repressing Dfd expression before stage 11. But after that stage CncB represses the maintenance phase of *Dfd* transcription in mandibular cells, perhaps by repressing the autoactivation circuit that is normally established during stages 9 and 10 (Zeng et al., 1994).

We also found that CncB is sufficient to repress *Dfd* transcription outside the mandibular segment. When CncB is ectopically expressed in embryos, *Dfd* transcript levels in the maxillary segment are reduced, especially in the anterior region of the segment (Fig. 6G). Note that these transcript expression assays were done at a time after heat shock when Dfd protein is still present at wild-type levels (Fig. 7A,B). Only the CncB isoform is capable of repressing *Dfd* transcription. Neither the ectopic expression of CncA nor CncC have an effect on the abundance or pattern of *Dfd* transcripts in the maxillary epidermis (Fig. 6F,H). Since the phenotypic effect of *hs-cncC* in epidermal cells strongly resembles that of *hs-cncB*, this indicates that the effect of Cnc gene products on maxillary epidermal development may not require repression of *Dfd* transcription per se.

The CncB repressive effect on *Dfd* expression might be mediated, at least in part, through autoactivation elements. To test this, we assayed the activity of three subregions from the

Dfd-epidermal autoactivation element (*Dfd*-EAE) in *hs-cncB* embryos, using *hs-cncA* and wild-type embryos as controls. At 30 minutes after heat-shock induction of CncB expression, when no change in Dfd protein abundance is detectable (Fig. 7A,B), there is a decrease in the activity of various modules of the *Dfd*-EAE. One of these, module F, consists of a 471 bp fragment at the 3′ end of the *Dfd* EAE (Zeng et al., 1994). This element activates *Dfd*-dependent reporter expression in posterior maxillary cells of wild-type embryos and is equally active in *hs-cncA* controls (Fig. 7C). Ectopic expression of CncB nearly abolishes the activity of this element (Fig. 7D). Another *Dfd*-EAE fragment, 570 bp module C, activates *Dfd*-dependent reporter transcription in most maxillary epidermal cells (Zeng et al., 1994; Fig. 7E). When compared with *hs-cncA* embryos, module C exhibits lower activity in *hs-cncB* embryos (Fig. 7F). The smallest known module of the *Dfd*-EAE with a significant amount of autonomous activity is the 120 bp module E, which is directly targeted by both Dfd and Exd proteins (Zeng et al., 1994; Pinsonneault et al., 1997). In *hs-cncB* embryos, expression levels from module E are lower than those observed in wild-type or *hs-cncA* embryos (Fig. 7G,H). Conversely, in cnc^{2E16}, cnc^{C7} or cnc^{VL110} embryos (shown), module E is ectopically expressed in mandibular cells (Fig. 7K,L). From these experiments, we conclude that *cnc* function is required to repress a variety of Dfd response elements in mandibular cells and that ectopic CncB is sufficient to reduce the activity of all of those elements in maxillary cells. The *hs-cncB*-induced repression of these elements occurs at a time after heat shock when Dfd protein levels in the maxillary segment are unchanged, evidence that the CncB effect on these elements is not indirectly caused by CncB repression of Dfd protein levels produced from the endogenous *Dfd* locus.

Although the cuticular phenotype conferred by ectopic expression of *hs-cncB* suggested it does not generally antagonize the function of Hox proteins that specify trunk regional identities, we wished to test whether *cis*-regulatory elements that are activated by other Hox proteins exhibited any response to ectopic CncB. To address this question, we tested the activity of the Ubx-activated *dpp*674 element (Capovilla et al., 1994; Sun et al., 1995), the Abd-B-activated *ems*-1.2 kb filzkörper element (Jones and McGinnis, 1993) and an element that is activated by the Hox protein Labial (Popperl et al., 1995; Chan et al., 1996). All of these elements exhibited patterns and amounts of reporter expression in *hs-cncB* embryos that were indistinguishable from controls (Fig. 7I,J and data not shown).

Another experiment to address whether CncB represses the maxillary-promoting function of *Dfd* by antagonizing its function on downstream target elements involved testing the effects of CncB on *Dfd*-dependent structures when *Dfd* expression is driven by an exogenous promoter. Heat-shocked embryos that are heterozygous for both *hs-Dfd* (Kuziora and McGinnis, 1988) and *hs-cncA* develop ectopic maxillary cirri and mouth hooks in 69% (*n*=94) of the first thoracic segments. However, when a *hs-cncB* is substituted for *hs-cncA* in the same genetic background, only 25% of the first thoracic segments bear thoracic cirri. Thus, even when *Dfd* expression

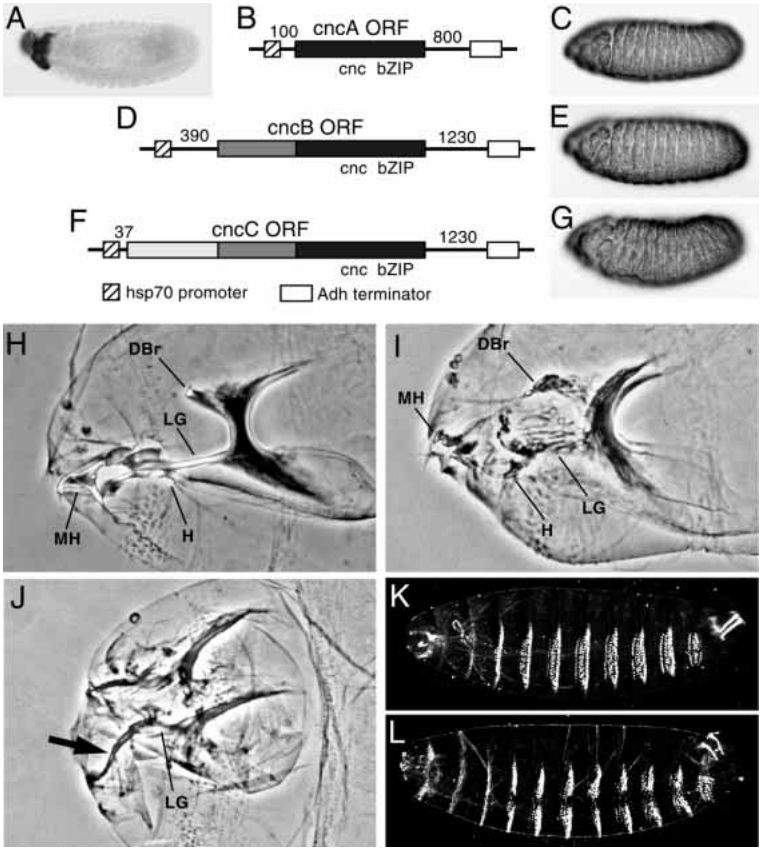

Fig. 5. Phenotypes produced by ectopic expression of Cnc isoforms. (A) A heat-shocked wild-type control embryo reacted with the common Cnc antibody. (B) A map of the heat-shock promoter-*cncA* construct. (C) A heat-shocked *hs-cncA* embryo reacted with the common Cnc antibody. (D) A map of the heat-shock promoter-*cncB* construct. (E) A heat-shocked *hs-cncB* embryo reacted with the common Cnc antibody. (F) A map of the heat-shock promoter-*cncC* construct. The first methionine in frame with the long ORF that includes the CNC and b-ZIP codons corresponds to the codon for methionine #21 in Fig. 3. (G) A heat-shocked *hs-cncC* transformed embryo reacted with the common Cnc antibody. All embryos shown were fixed 15 minutes post heat shock. Head cuticles at terminal stages of embryogenesis are shown for (H) wild-type, (I) *hs-cncB* and (J) *hs-cncC*. One hour heat shocks were performed at 6-8 hours of embryogenesis (see Materials and Methods). *hs-cncB* cuticles have only fragments of mouth hook material (MH) at the anterior tip of the head and fewer cirri. The H-piece (H) is also either fragmented or absent. The dorsal bridge (DBr) and truncated lateralgräten (LG) are recognizable, although more diffuse in appearance. The median tooth is also recognizable although abnormally shaped. *hs-cncC* cuticles do not develop mouth hooks. The cirri and dorsal bridge are greatly reduced or absent. The median tooth is recognizable although abnormally shaped. There are narrow lateralgräten (LG) in their proper position in the head but there is also a sclerotic extension of the lateralgräten (arrow) that may represent a duplication of lateralgräten or H-piece arms. (K) Dark-field view of denticle belts in a wild-type first instar larva. (L) First instar denticle belts of a *hs-cncB* cuticle after a 1 hour heat shock at 6-8 hours of development. In the heat-shocked embryos, the abdominal denticle pattern is disrupted along the ventral midline in approximately 50% of the embryos.

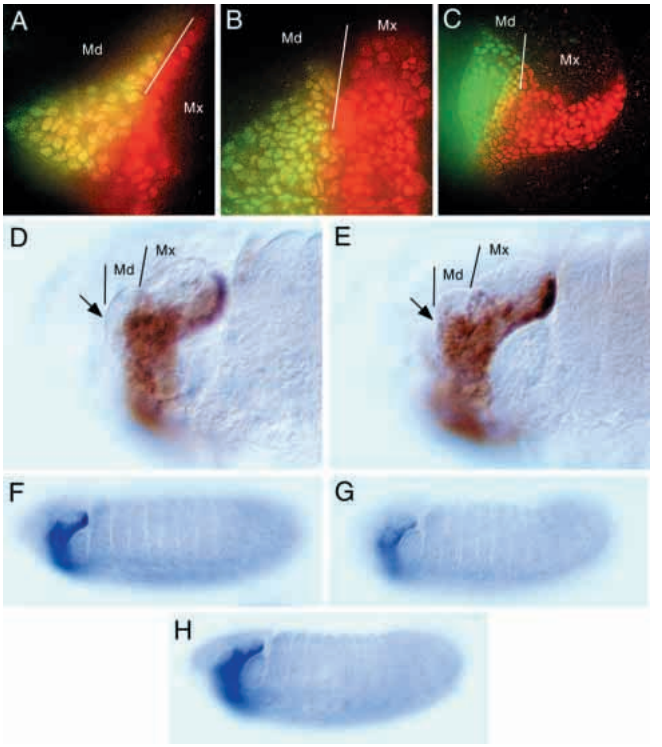

Fig. 6. *Dfd* and *cnc* expression in wild-type and mutant backgrounds. (A-C) Double labeling for Cnc protein (green) and Dfd protein (red) in wild-type embryos at successive stages of embryonic development. Overlapping regions of protein expression are yellow. Embryos are oriented with ventral down and anterior to the left. The white line indicates the boundary between the maxillary and mandibular lobes. (A) At stage 9, CncB protein is limited to mandibular nuclei while Dfd protein is present in all mandibular and maxillary segment nuclei. (B) At stage 11, Dfd protein is still abundant in posterior mandibular cells overlapping with CncB protein, but Dfd levels are lower in anterior mandibular cells. (C) At stage 13, Dfd protein is limited to the maxillary segment and an approximately 2-cell-wide stripe in the posterior mandibular segment where it overlaps with CncB expression. In the anterior mandibular segment, Dfd protein expression has been repressed. CncB protein is still abundant throughout the mandibular lobe, except on the ventral aspect of the embryo, where CncB is excluded from the posterior mandibular compartment (Mohler et al., 1995). (D) Stage 12 wild-type, and (E) stage 12 cnc^{2E16}/cnc^{2E16} mutant embryos reacted with anti-Dfd antibodies. In wild-type, *Dfd* transcript and protein expression (shown) is limited to the maxillary segment and a 2-cell-wide stripe in the posterior mandibular segment. In cnc^{2E16}/cnc^{2E16} mutants, *Dfd* transcript and protein expression (shown) persists in anterior mandibular cells (arrow) in a pattern that resembles the pattern in maxillary cells. (F) *Dfd* transcript expression pattern in stage 12 *hs-cncA* embryos. (G) *Dfd* transcript expression pattern in stage 12 *hs-cncB* embryos. Both embryos were heat shocked for 1 hour at 37°C and fixed 30 minutes later (see Materials and Methods). Levels of *Dfd* transcripts in *hs-cncA* embryos and wild-type embryos after heat shock are indistinguishable, but are reduced in *hs-cncB* embryos. (H) *Dfd* transcripts in *hs-cncC* embryos fixed 30 minutes after heat shock. The pattern of *Dfd* transcripts is identical to wild-type or *hs-cncA* controls.

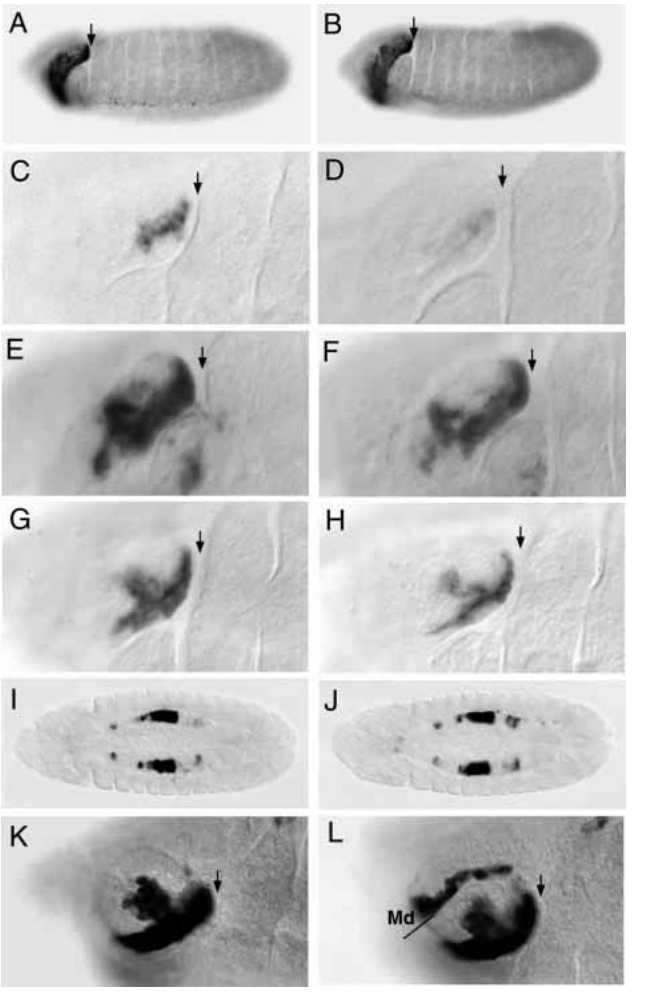

Fig. 7. Hox response elements in *cnc* mutants and *hs-cnc* genetic backgrounds. Wild-type and *hs-cnc* embryos were heat shocked for 1 hour at 37°C, allowed to recover for 30 minutes, then fixed and either hybridized with a digoxigenin-labeled RNA probe for *lacZ* reporter transcripts (C-J) or reacted with anti-Dfd antiserum (A,B). Arrows indicate the posterior boundary of the maxillary segment. The derivation of the Dfd response modules shown here are described in Zeng et al. (1994). (A) Dfd protein expression in *hs-cncA*. (B) Dfd protein expression in *hs-cncB* embryos. Dfd protein abundance is unchanged in *hs-cncB* embryos at 30 minutes after heat shock, although *Dfd* RNA transcripts are reduced at this time point (see Fig. 6G). (C) *Dfd*-EAE module F activity in *hs-cncA* embryos. (D) *Dfd*-EAE module F activity in *hs-cncB* embryos. (E) *Dfd*-EAE module C activity in *hs-cncA* embryos. (F) *Dfd*-EAE module C activity in *hs-cncB* embryos. (G) *Dfd*-EAE 4X module E activity in *hs-cncA* embryos at stage 12. (H) *Dfd*-EAE 4X module E activity in *hs-cncB* embryos at stage 12. (I) Embryonic expression pattern of *dpp*674, a Ubx response element, in *hs-cncA* embryos. (J) *dpp*674 activity in *hs-cncB* embryos, which is indistinguishable from *hs-cncA* or wild type. (K) Expression pattern of *Dfd*–EAE 4X module E detected with anti-β-gal antibodies in wild-type stage 13 embryos. (L) In cnc^{VL110} mutant embryos at stage 13, *Dfd*–EAE 4X module E is activated ectopically in posterior mandibular (Md) cells.

is driven in ectopic positions by a heat-shock promoter, the maxillary-promoting function of Dfd protein is reduced in the presence of CncB. This could either be due to CncB-mediated repression of the *Dfd* autoactivation circuit in ectopic positions or to CncB repression of downstream target elements of Dfd protein, or to both of these effects.

cnc is required to repress mandibular *Dll* expression

One of the downstream genes that is activated by *Dfd* in maxillary cells is *Distal-less* (*Dll*). *Dll* is required for the formation of the larval appendage primordia and the distal regions of adult appendages (Cohen et al., 1989). In the maxillary segment, *Dll* is expressed in two patches of cells: a dorsal patch that gives rise to the maxillary sense organ and a ventral patch that consists of the primordia for the maxillary cirri (Fig. 8A). The dorsal maxillary domain of *Dll* expression is largely independent of *Dfd* function, while the ventral maxillary patch of *Dll* is activated by *Dfd* through a 3′ enhancer (the ETD6 element) (O'Hara et al., 1993). In cnc^{2E16} mutant embryos, *Dll* is ectopically expressed in ventral mandibular cells, suggesting that *cncB* represses *Dll* transcription in mandibular cells (Fig. 8B). In *hs-cncB* embryos 30 minutes after heat shock, when Dfd protein abundance is normal (Fig. 7A,B), *Dll* expression is repressed in the ventral maxillary segment (Fig. 8D) but other domains of *Dll* expression in the head and thorax are relatively unaffected, indicating that CncB selectively represses the *Dfd*-dependent portion of the *Dll* expression pattern. In *hs-cncC* and *hs-cncA* embryos, the ventral maxillary expression of *Dll* is not selectively repressed (Fig. 8C). Reporter gene expression from the *Dll* ETD6 enhancer follows the expression of *Dll* as the enhancer is ectopically activated in the ventral mandibular region in *cnc* mutants and repressed in *hs-cncB* embryos (data not shown).

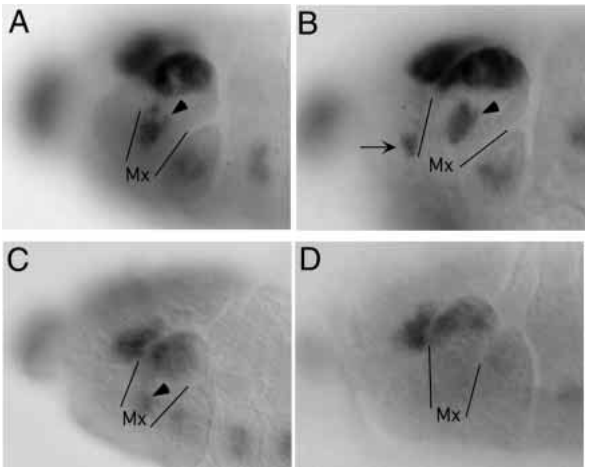

Fig. 8. *Distal-less* expression in *cnc* mutants and *hs-cnc* genetic backgrounds. The panels show expression patterns obtained using a digoxigenin-labeled RNA probe that detects transcripts from the *Dfd* downstream gene *Dll* (O'Hara et al., 1993). (A) *Dll* expression in a stage 12 wild-type embryo. Note the dorsal-lateral patch in the maxillary segment as well as the *Dfd*-dependent patch in the center of the maxillary segment (arrowhead). (B) *Dll* expression in a stage 12 cnc^{2E16} homozygote. The wild-type *Dll* expression pattern is obtained as well as ectopic expression in the mandibular segment (arrow). (C) *Dll* expression in stage 12 *hs-cncA* embryos at 30 minutes post heat shock. In the *hs-cncA* background, or in wild-type embryos, the pattern of *Dll* is normal, although *Dll* transcript levels are globally lower in all *hs-cnc* backgrounds after heat shock. (D) *Dll* expression in *hs-cncB* embryos. The dorsal lateral maxillary expression pattern of *Dll* is observed, but the ventral *Dfd*-dependent patch of *Dll* expression is repressed.

DISCUSSION

The *cnc* gene has been proposed to encode a spatial determinant required for proper segmental identity in the posterior head (Mohler et al., 1995). Here we find that a crucial part of this segmental identity function is provided by the localized expression of the CncB protein isoform. In loss-of-function mutants in which CncB function is reduced or absent, mandibular epidermal cells assume the identity of maxillary epidermis. This is in part because Dfd function is no longer antagonized in mandibular cells, which our current evidence suggests is mechanistically accomplished by CncB exerting a repressive effect on many Dfd response elements. These elements include 'downstream' autoactivation elements at the *Dfd* locus and Dfd response elements at downstream genes such as *Dll* and *1.28* (Mohler et al., 1995). When CncB function is reduced or lost, these downstream target genes and Dfd response elements are inappropriately activated in mandibular cells. CncB is one of the growing list of genetic functions, other examples being Exd/Pbx class proteins, the products of the *teashirt* and *homothorax* genes and other Hox genes, that modulate the functional activity and specificity of the Hox system, and thereby diversify A/P body axis morphology (Peifer and Weischaus, 1990; Gonzalez-Reyes and Morata, 1990; Roder and Kerridge, 1992; Jurgens and Hartenstein, 1993; Mohler et al., 1995; Rieckhof et al., 1997; Kurant et al., 1998; Pai et al., 1998). Unlike Exd/Pbx, there is as yet no evidence that the modulatory effect is mediated by direct binding of CncB to Dfd protein.

CncB is not only required to repress Dfd target genes and Dfd response elements in most mandibular cells, but also is sufficient to partially repress such Dfd targets in the maxillary segment. When CncB is ectopically expressed during mid-embryogenesis, many of the normal *Dfd*-dependent cuticular structures produced by maxillary cells are reduced or absent. This is associated with the ability of CncB to reduce the levels of *Dfd* transcription in maxillary cells. This is due, at least in part, to CncB action on *Dfd* epidermal autoactivation elements, since the activity of some of these elements is rapidly repressed by ectopic CncB. No ectopic mandibular structures (e.g. lateralgräten, dorsolateral papillae) are detected in the *hs-cncB* cuticular preparations, therefore it appears that the heat-shock-induced expression of CncB, coupled with Dfd function, is not sufficient to specify mandibular segmental identity in place of maxillary identity.

Among the many known and suspected targets of Hox proteins, the CncB repression function appears to be highly selective for Dfd targets. In normal embryos, this is partly due to the highly restricted expression pattern of CncB, which overlaps the expression pattern of *Dfd*, but not the expression of other Hox genes except for *proboscipedia* (Pultz et al., 1988). However, even when *hs-cncB* is ectopically expressed, the epidermal phenotype indicates that the function of Hox genes that act in the trunk (e.g. *Scr*, *Antp*, *Ubx*, *abd-A* and *Abd-B*) are largely unaffected. This selective activity might be due to the presence of CncB-binding sites in many *Dfd* response elements and the absence of such sites from downstream elements that are activated by other Hox proteins. All of the *Drosophila* Cnc protein isoforms are closely related in their CNC and b-ZIP domains to the well-studied p45 subunit of the mammalian NF-E2 transcription factor. Based on this, it is

expected that one mode of CncB action will involve heterodimerization with small Maf-class proteins on sequences with the consensus GCTGANTCAT (Andrews et al., 1993a,b). The underlined nucleotides designate the palindrome found in many sites bound by homodimers or heterodimers of b-ZIP proteins. None of the elements (modules C, E or F of the *Dfd* epidermal autoregulatory enhancer) that show ectopic activity in *cnc* mutants and are repressed by ectopic expression of CncB, have sequences that match the TGANTCA core of the b-ZIP-binding site. Module E has been previously subjected to systematic mutageneses that involved clustered substitutions throughout its length (Zeng et al., 1994; Gross and McGinnis, 1995; Pinsonneault et al., 1997). None of the mutant versions of module E exhibited ectopic activity in the mandibular segment, although discrete subregions were required for the activity of this Dfd response element in the maxillary segment. The sum of this evidence suggests that the sequences in module E that transduce the *cnc* repressive effect overlap with the sequences required for module E activation. At present, the mechanism by which CncB function is directed to Dfd target elements is unknown. Since the ectopic expression of CncA protein has no detectable effect on morphology or viability, even though it possesses the same DNA-binding domain as CncB, it seems unlikely that a Cnc DNA-binding function alone is sufficient to repress the activity of Dfd response elements.

One region of *Dfd* expression that appears to be immune to the repressive effects of CncB is in posterior mandibular epidermis. Throughout most of embryogenesis, these cells maintain abundant levels of CncB protein and levels of Dfd protein that are only somewhat lower than the levels detected in maxillary cells (Fig. 6). One possible reason for this is that some of the autoactivation enhancers that supply *Dfd* expression in this region are not completely repressed by CncB. It is only a subset of such elements that are immune to CncB, since other autoregulatory enhancers that contribute to Dfd expression in this region are strongly activated when CncB function is absent (Fig. 7K,L). We believe that the most-likely possibility is that some of the persistent *Dfd* transcription in the posterior regions of the maxillary and mandibular segments is supplied by elements that require Dfd protein only to achieve the normal levels of expression and have independent sources of regulatory input that determine their spatial limits of expression. There is evidence for the existence of such elements, which act in the posterior region of the maxillary segment (Zeng et al, 1994), and this also applies to the posterior mandibular segment at early stages of embryogenesis (G. Gellon and W. McG., unpublished results).

It is possible that protein-protein interactions play a role in the selectivity of CncB action. The difference between the CncA protein and the CncB protein resides in the 272 amino acids found at the N terminus of CncB. This 272 amino acid domain, required for the repressive effect on Dfd response elements, has no significant similarity to non-repetitive amino acid sequence of proteins that are currently listed in public databases. It does, however, share a number of repetitive amino acid motifs that are found in a variety of other transcription factors. These include His-Pro repeats, Ala repeats, Val-Gly repeats and Pro repeats. Alanine-rich regions have been associated in a few instances with transcriptional repression functions (Licht et al., 1990; Han and Manley, 1993 Hanna-Rose and Hansen, 1996). Thus, it is possible that the 272 amino acid domain has a relatively generic repression function that accounts for its ability to interfere with Dfd protein on target elements. This 272 amino acid domain might also contain sequences that allow it to specifically repress Dfd response elements, while having little effect on other Hox response elements.

The predicted CncC protein also possesses this 272 amino acid domain, but in addition has an additional 491 amino acids at its N terminus that are unique to this isoform. At present, there are no known mutant alleles that remove only the *cncC* function and leave the function of the other isoforms intact, but the *hs-cncC*-induced defects allow us to speculate about the role of *cncC* in normal development. Heat-shock-induced ectopic expression of CncC protein has a profound influence on head development, removing all *Dfd*-dependent maxillary epidermal structures and affecting the normal morphogenesis of many other cuticular structures from a variety of head segments. Interestingly, it ablates the maxillary-promoting function of *Dfd* without any detectable affect on *Dfd* transcription levels. The *cncC* transcript isoform is apparently maternally deposited, and expressed in all or virtually all cells during embryogenesis at levels that vary at different stages. Using staining intensity as a guide with probes of similar sizes, the *cncC* transcript levels appear to be much lower than those of *cncB* in the mandibular and hypopharyngeal regions. One model consistent with all these results is that the CncC function sets a threshold level of repressor that must be overcome in order for Dfd and other head-patterning functions to activate transcription from some or many of their downstream target genes. Conceivably, the threshold level of this repressor might even change the segment identity function of Dfd protein by regulating its ability to activate different numbers of target genes. Consistent with this model, when the level is dramatically raised by heat-shock-induced expression of CncC, the ability of *Dfd* and other head-patterning genes to promote the development of head structures is lost. Further experiments with *cncC*-specific null mutations, in combination with markers for target genes of Dfd and other head-patterning factors, will be required to explore this and other potential explanations for the *cncC* gain-of-function and loss-of-function phenotypes.

CncB and CncC may act in the mandibular segment in a manner that resembles how posterior Hox proteins act to influence the function of anterior Hox genes during the process of phenotypic suppression (Gonzalez-Reyes and Morata, 1990; Macias and Morata, 1996). For example, the Hox protein Ultrabithorax (Ubx) can suppress the thoracic-promoting function of Antennapedia (Antp) in a manner that is independent of Ubx regulatory effects on *Antp* transcription. One mechanism that has been proposed to explain phenotypic suppression is competition for common binding sites by the different members of the Hox family, although other mechanisms are equally plausible (Duboule and Morata, 1994). Binding-site competition seems highly unlikely to be sufficient for the CncB suppressive effect on Dfd response elements, since CncA also possesses the same DNA binding and dimerization motif as CncB, and has no detectable influence on Dfd expression or function. Though binding site competition in itself seems to be an unlikely mechanism, it is intriguing that the half site that is recognized by CNC class b-

Zip proteins, $^{G}/_{A}$TCAT, resembles the preferred half-site (ATCA) for proteins of the PBC class (e.g *Drosophila* Exd, mammalian Pbx). The Exd/Pbx proteins bind cooperatively to DNA with many Hox proteins (Mann and Chan, 1996) and the *exd* function in *Drosophila* appears to be required for many of the Hox proteins to activate downstream target genes, but not for their repression effects on targets (Pinsonneault et al., 1997). Perhaps one way in which CncB acts is by antagonizing the function of this known Hox coactivator on certain Dfd response elements.

In many interesting ways, the interactions of CncB with Dfd resemble those of the *teashirt* (*tsh*) gene with Hox genes of the trunk, particularly *Sex combs reduced (Scr)*. *Scr* is normally expressed both in the labial and 1st thoracic segments, while *tsh* expression overlaps only the 1st thoracic (T1) portion of the *Scr* domain (Fasano et al., 1991). Both *tsh* and *Scr* are required for T1 identity, and in order to achieve the normal morphology of this segment, *tsh* represses *Scr* transcription and function in T1 (Fasano et al., 1991; Roder and Kerridge, 1992). *Teashirt* is also capable of repressing some of the morphogenetic functions of *Scr* in other segments when *Scr* expression is driven by heterologous promoters (Andrew et al., 1994). The mechanism that integrates the functions of *tsh* and *Scr* is not yet known, but presumably occurs on common downstream target elements, since the *tsh* gene encodes a zinc-finger protein with a sequence-specific DNA-binding function (Alexandre et al., 1996).

CncB is one of three isoforms produced from the *cnc* locus. The other two, CncA and CncC, are expressed maternally, as their transcripts are present in early syncytial embryos. Thus either or both of these isoforms may play a role in oocyte development or in early stages of embryogenesis. Analysis of a P-element insertional allele of *cnc* indicates that a function at the locus is required for germ cell viability or early oogenesis (Perrimon et al., 1996). Both the CncA and cncC isoforms are also expressed at later stages in most or all embryonic cells. Studies of their roles in development, and how these roles are integrated with the role of CncB, await mutations that selectively eliminate their functions.

We thank Jym Mohler, Xuelin Li and Juan Botas for sharing clones, flies and regulatory elements, Brian Florence and Elizabeth Wiellette for critical readings of the manuscript and Raffi Aroian for instruction on the confocal microscope. The Bloomington *Drosophila* stock center and Flybase supplied invaluable support. This research was supported by an NIH grant (W. M.), a Fulbright fellowship (E. R.) and by an HHMI pre-doctoral fellowship (A. V.).

REFERENCES

Agard, D., Hiraoka, Y., Shaw, P. and Sedat, J. (1989). Fluorescence microscopy in three dimensions. *Meth. Cell. Biol.* **30**, 353-377.

Alexandre, E., Graba, Y., Fasano, L., Gallet, A., Perrin, L., De Zulueta, P., Pradel, J., Kerridge, S. and Jacq, B. (1996). The *Drosophila* teashirt homeotic protein is a DNA-binding protein and modulo, a HOM-C regulated modifier of variegation, is a likely candidate for being a direct target gene. *Mech. Dev.* **59**, 191-204.

Andrew, D. J., Horner, M. A., Petitt, M. G., Smolik, S. M. and Scott, M. P. (1994). Setting limits on homeotic gene function: restraint of Sex combs reduced activity by *teashirt* and other homeotic genes. *EMBO J.* **13**, 1132-1144.

Andrews, N., Erdjument-Bromage, H., Davidson, M., Tempst, P. and Orkin, S. (1993a). Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* **362**, 722-728.

Andrews, N. C., Kotkow, K. J., Ney, P. A., Erdjument-Bromage, H., Tempst, P. and Orkin, S. (1993b). The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. *Proc. Natn. Acad. Sci., USA* **90**, 11488-11492.

Bergson, C. and McGinnis, W. (1990). The autoregulatory enhancer element of the *Drosophila* homeotic gene *Deformed*. *EMBO J.* **9**, 4287-4297.

Botas, J. (1993). Control of morphogenesis and differentiation by HOM/Hox genes. *Current Opin. Cell Biol.* **5**, 1015-1022.

Bowerman, B., Eaton, B. and Priess, J. R. (1992). *skn-1*, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. *Cell* **68**, 1061-1075.

Brown, N. H. and Kafatos, F. C. (1988). Functional cDNA libraries from *Drosophila* embryos. *Mol. Biol.* **203**, 425-437.

Capovilla, M., Brandt, M. and Botas, J. (1994). Direct regulation of *decapentaplegic* by Ultrabithorax and its role in *Drosophila* midgut morphogenesis. *Cell* **76**, 461-475.

Chan, J. Y., Han, X.-L. and Kan, Y. W. (1993). Cloning of Nrfl, and NF-E2-related transcription factor, by genetic selection of yeast. *Proc. Natn. Acad. Sci., USA* **90**, 11371-11375.

Chan, S. K., Popperl, H., Krumlauf, R. and Mann, R. S. (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.* **15**, 2476-2487.

Cohen, S. M., Brönner, G., Küttner, F., Jürgens, G. and Jäckle, H. (1989). Distal-less encodes a homoeodomain protein required for limb development in *Drosophila*. *Nature* **338**, 432-434.

Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**, 358-364.

Ekker, S., Jackson, D., von Kessler, D., Sun, B., Young, K. and Beachy, P. (1994). The degree of variation in DNA sequence recognition among four *Drosophila* homeotic proteins. *EMBO J.* **13**, 3551-3560.

Farmer, S., Sun, C., Winnier, G., Hogan, B. and Townes, T. (1997). The bZIP transcription factor LCR-F1 is essential for mesoderm formation in mouse development. *Genes Dev.* **11**, 786-798.

Fasano, L., Roder, R., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S. (1991). The gene *teashirt* is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. *Cell* **64**, 63-79.

Gonzalez-Reyes, A. and Morata, G. (1990). The developmental effect of overexpressing a Ubx product in *Drosophila* embryos is dependent on its interactions with other homeotic products. *Cell* **61**, 515-522.

Graba, Y., Aragnol, D. and Pradel, J. (1997). *Drosophila* Hox complex downstream targets and the function of homeotic genes. *BioEssays* **19**, 379-88.

Gross, C. and McGinnis, W. (1995). DEAF-1, a novel protein that binds an essential region in a Deformed response element. *EMBO J.* **15**, 1961-1970.

Hales, K. and Fuller, M. (1997). Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. *Cell* **90**, 121-129.

Han, K. and Manley, J. L. (1993). Functional domains of the *Drosophila* Engrailed protein. *EMBO J.* **12**, 2723-2733.

Hanna-Rose, W., and U. Hansen. (1996). Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet*. **12**, 229-234

Harding, K. W., Gellon, G., McGinnis, N. and McGinnis, W. (1995). A screen for *Dfd* modifier mutations in *Drosophila*. *Genetics* **140**, 1339-1352.

Irvine, K. D., Helfand, S. L. and Hogness, D. S. (1991). The Large Upstream Control Region of the *Drosophila* Homeotic Gene *Ultrabithorax*. *Development* **111**, 407-424.

Jones, B. and McGinnis, W. (1993). The regulation of *empty spiracles* by Abdominal-B mediates an abdominal segment identity function. *Genes Dev.* **7**, 229-240.

Jurgens, G. and Hartenstein, V. (1993). The terminal regions of the body pattern. In *The Development of Drosophila melanogaster*, vol. 1 (ed. M. Bate and A. Martinez-Arias), pp. 687-746. Cold Spring Harbor: Cold Spring Harbor Press.

Jurgens, G., Lehman, R., Schardin, M. and Nusslein-Volhard, C. (1986). Segmental organization of the head in the embryo of *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.* **195**, 359-377.

Kenyon, C. (1994). If birds fly, why can't we? Homeotic genes and evolution. *Cell* **78**, 175-180.

Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* **78**, 191-201.

Kurant, E., Pai, C. Y., Sharf, R., Halachmi, N., Sun, Y. H. and Salzberg, A. (1998). *Dorsotonals/homothorax*, the *Drosophila* homologue of meis1,

interacts with extradenticle in patterning of the embryonic PNS. *Development* **125**, 1037-48.

Kuziora, M. A. and McGinnis, W. (1988). Autoregulation of a *Drosophila* homeotic selector gene. *Cell* **55**, 477-485.

Laughon, A. (1991). DNA binding specificity of homeodomains. *Biochemistry* **30**, 11357-11367.

Lawrence, P. A. and Morata, G. (1994). Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell* **78**, 181-189.

Licht, J. D., Grossel, M. J., Figge, J. and Hansen, U. M. (1990). *Drosophila* Kruppel protein is a transcriptional repressor. *Nature* **346**, 76-78.

Macias, A. and Morata, G. (1996). Functional hierarchy and phenotypic suppression among *Drosophila* homeotic genes: the *labial* and *empty spiracles* genes. *EMBO J.* **15**, 334-343.

Manak, J. R. and Scott, M. P. (1994). A class act: conservation of homeodomain protein functions. *Development* **1994 Supplement**, 61-71.

Mann, R. S. and Chan, S. K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends in Genetics* **12**, 259-262.

McGinnis, W., Jack, T., Chadwick, R., Regulski, M., Bergson, C., McGinnis, N. and Kuziora, M. A. (1990). Establishment and maintenance of position-specific expression of the *Drosophila* homeotic selector gene *Deformed*. *Advances In Genetics* **27**, 363-402.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

Merrill, V. K. L., Turner, F. R. and Kaufman, T. C. (1987). A genetic and developmental analysis of mutations in the *Deformed* locus in *Drosophila melanogaster*. *Dev. Biol.* **122**, 379-395.

Mohler, J., Mahaffey, J. W., Deutsch, E. and Vani, K. (1995). Control of *Drosophila* head segment identity by the bZIP homeotic gene *cnc*. *Development* **121**, 237-247.

Mohler, J., Vani, K., Leung, S. and Epstein, A. (1991). Segmentally restricted, cephalic expression of a leucine zipper gene during *Drosophila* embryogenesis. *Mech. Dev.* **34**, 3-10.

O'Hara, E., Cohen, B., Cohen, S. M. and McGinnis, W. (1993). Distal-less is a downstream gene of Deformed required for ventral maxillary identity. *Development* **117**, 847-856.

Pai, C. Y., Kuo, T. S., Jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A., Salzberg, A. and Sun, Y. H. (1998). The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in *Drosophila*. *Genes Dev.* **12**, 435-46.

Peifer, M. and Weischaus, E. (1990). Mutations in the *Drosophila* gene *extradenticle* affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev.* **4**, 1209-1223.

Perrimon, N., Lanjuin, A., Arnold, C. and Noll, E. (1996). Zygotic lethal mutations with maternal effect phenotypes in *Drosophila* melanogaster. II. Loci on the second and third chromosomes identified by P-element-induced mutations. *Genetics* **144**, 1681-92.

Pinsonneault, J., Florence, B., Vaessin, H. and McGinnis, W. (1997). A model for extradenticle function as a switch that changes Hox proteins from repressors to activators. *EMBO J.* **16**, 2032-2042.

Popperl, H., Bienz, M., Studer, M., Chan, S., Aparicio, S., Brenner, S., Mann, R. S. and Krumlauf, R. (1995). Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. *Cell* **81**, 1031-1042.

Pultz, M., Diederich, R. J., Cribbs, D. L. and Kaufman, T. C. (1988). The *proboscipedia* locus of the Antennapedia Complex: a molecular and genetic analysis. *Genes Dev.* **2**, 901-920.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. and Mann, R. S. (1997). Nuclear translocation of extradenticle requires *homothorax*, which encodes an extradenticle-related homeodomain protein. *Cell* **91**, 171-183.

Roder, V. and Kerridge (1992). The role of the *teashirt* gene in trunk segmented identity in *Drosophila*. *Development* **115**, 1017-1033.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*, 2nd Ed. Cold Spring Harbor, New York: Cold Spring Harbor Press.

Sun, B., Hursh, D. A., Jackson, D. and Beachy, P. A. (1995). Ultrabithorax protein is necessary but not sufficient for full activation of decapentaplegic expression in the visceral mesoderm. *EMBO J.* **14**, 520-535.

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C. and Kennison, J. A. (1992). brahma: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68**, 561-572.

Tautz, D. and Pfeifle, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* **98**, 81-85.

Zeng, C., Pinsonneault, J., Gellon, G., McGinnis, N. and McGinnis, W. (1994). Deformed protein binding sites and cofactor binding sites are required for the function of a small segment-specific regulatory element in *Drosophila* embryos. *EMBO J.* **13**, 2362-2377.