Distinct Hox protein sequences determine specificity in different tissues

Sophie Chauvet*†, Samir Merabet*†, David Bilder‡§, Matthew P. Scott‡, Jacques Pradel*, and Yacine Graba*¶

*Laboratoire de Génétique et Biologie du Développement, Institut de Biologie du Développement de Marseille, Centre National de la Recherche Scientifique/Université de la Méditerranée, Parc Scientifique de Luminy, Case 907, 13288 Marseille Cedex 9, France; and ‡Departments of Developmental Biology and Genetics, Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305-5427

Contributed by Matthew P. Scott, February 2, 2000

***Hox* genes encode evolutionarily conserved transcription factors that control the morphological diversification along the anteroposterior (A/P) body axis. Expressed in precise locations in the ectoderm, mesoderm, and endoderm, Hox proteins have distinct regulatory activities in different tissues. How Hox proteins achieve tissue-specific functions and why cells lying at equivalent A/P positions but in different germ layers have distinctive responses to the same Hox protein remains to be determined. Here, we examine this question by identifying parts of Hox proteins necessary for Hox function in different tissues. Available genetic markers allow the regulatory effects of two Hox proteins, Abdominal-A (AbdA) and Ultrabithorax (Ubx), to be distinguished in the *Drosophila* embryonic epidermis and visceral mesoderm (VM). Chimeric Ubx/AbdA proteins were tested in both tissues and used to identify protein sequences that endow AbdA with a different target gene specificity from Ubx. We found that distinct protein sequences define AbdA, as opposed to Ubx, function in the epidermis vs. the VM. These sequences lie mostly outside the homeodomain (HD), emphasizing the importance of non-HD residues for specific Hox activities. Hox tissue specificity is therefore achieved by sensing distinct Hox protein structures in different tissues.**

homeotic genes | tissue specificity | Abdominal-A | Ultrabithorax | homeodomain

Hox genes encode evolutionarily conserved transcription factors controlling morphogenetic traits along the anteroposterior (A/P) axis (1). The encoded proteins share a 60-aa DNA-binding motif, the homeodomain (HD), which folds into three α-helices. The N-terminal region of the HD just upstream of helix 1 contacts the minor groove of DNA, whereas helix 3 and particularly the HD amino acid 50 within this helical structure interact with the major groove. In addition, several residues throughout the HD contact the DNA phosphate backbone. The mode of HD/DNA interaction appears to be highly conserved within Hox and more largely within HD-containing proteins (2). Consequently, the DNA-binding properties of Hox proteins are very similar, raising the question of how proteins with equivalent biochemical properties reach *in vivo* distinct regulatory effects to ultimately initiate distinct developmental programs (3).

The study of the *Drosophila extradenticle* (*exd*), and of the related vertebrate *Pbx* genes, has provided important support for the idea that interaction with protein cofactors critically contributes to distinguishing the functions of Hox proteins. The major established outcome of the Hox-Exd/Pbx interactions is to increase the DNA-binding specificity, allowing different Hox-Exd/Pbx complexes to select distinct target genes (4). These interactions control not only target specificity, but also control whether a Hox protein function will act as a transcriptional activator or a repressor (5, 6). Other potential cofactors have been identified, suggesting that protein–protein interaction provides an important mechanism for conferring specificity on Hox proteins (7, 8).

Understanding *Hox* gene specificity also requires learning how a single Hox protein can have different influences on cells of different types. The large domains of *Hox* gene expression, in multiple tissues, contrast with the distinctive effects the genes have at each location. For example, the *abdA* expression domain in the visceral mesoderm (VM) encompasses the third and fourth *Drosophila* midgut chambers (9). In each chamber, the Abdominal-A (AbdA) protein has distinct transcriptional specificity. In the third chamber, AbdA activates the zinc finger protein Odd-paired, whereas in the fourth chamber AbdA activates the *pointed* gene, which encodes an Ets transcription factor. We previously examined the genetic basis for these different effects of AbdA on posterior midgut development. We found that Wnt (wingless, Wg) and TGF-β (decapentaplegic, Dpp) signals subdivide the AbdA Hox domain and regionalize AbdA transcriptional activity along the A/P axis (10).

The diverse effects of the Hox proteins are especially evident for cells lying at equivalent A/P coordinates but in distinct embryonic germ layers. Accordingly, the regulation of most identified *Drosophila* Hox downstream target genes is tissue specific; a target regulated in one tissue is usually not in another (11). A Hox protein might be active only in one tissue because of competing or inhibiting factors that differ in the two tissues, or because of a need for collaborating protein that is present only in one tissue. In either case, specific parts of the Hox protein might be required for interacting with the relevant cofactor or antagonist. If this reasoning is correct, it should be possible to identify parts of a Hox protein sequence that are required in one tissue but not in another.

In the present study, we searched for such tissue-dedicated determinants of Hox protein specificity. Our approach has been to construct Ultrabithorax (Ubx)/AbdA chimeric proteins and analyze their activity in the epidermis and VM. We identify the relevant protein sequences that distinguish AbdA from Ubx. The results emphasize the functional importance of residues lying outside of the HD and demonstrate that distinct protein sequences are required for AbdA specific actions in the epidermis and VM.

Materials and Methods

Construction of Chimeric Genes and Establishment of Transgenic Lines. Chimeras A–F were generated by introducing point mutations in the *Ubx Ia* cDNA. The 2.2-kb *Bam*HI/*Eco*RI fragment containing the entire ORF of *Ubx Ia* was subcloned in the *Bam*HI/*Eco*RI site of the vector pBluescript KS(+) (Stratagene). Single-stranded DNA from this phagemid vector was prepared according to standard procedures and used as a template to introduce the desired mutations by using the sculptor mutagenesis kit (Amersham Pharmacia). Mutagenic oligonucleotides used to generate chimeras A–F

Abbreviations: A/P, anteroposterior; HD, homeodomain; exd, extradenticle; AbdA, Abdominal-A; VM, visceral mesoderm; Wg, Wingless; Dpp, Decapentaplegic; Ubx, Ultrabithorax; SOE, splicing by overlap extension; CterC, C-terminal conserved; NterC, N-terminal conserved; Antp, Antennapedia; PS, parasegment; Dfd, Deformed.

†S.C. and S.M. contributed equally to this work.

§Present address: Department of Genetics, Howard Hughes Medical Institute, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115.

¶To whom reprint requests should be addressed. E-mail: graba@lgpd.univ-mrs.fr.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.070046997. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.070046997

were as follows: Mutagenic aa 1: 5′-P GT ACA AAT GGT CCG CGA AGA CGC; Mutagenic aa 12: 5′-P CA TAC ACC CGC TTC CAG ACG CTC; Mutagenic aa 1 + 12: 5′-P GT ACA AAT GGT CCG CGA AGA CGC GGC CGA CAG ACA TAC ACC CGC TTC CAG ACG CTC; Mutagenic aa CterC (61 + 32 + 64 + 67): 5′-P G AAG AAG GAG CTC CGG GCG GTC AAG GAG ATC AAC GAA CAG; Mutagenic aa 23 + 35: 5′-P G AAG GAG TTC CAC TTT AAT CAT TAT CTG ACC CGC AGA CGG AGA ATC GAG ATC GCG CAC CCG. Chimeras A–F were sequenced to select mutant clones. Each construct was cloned in the P-Casper Hsp70 vector by replacing the wild-type 2.2-kb *Bam*HI/*Eco*RI fragments by the mutated versions of it. The switch of the 2.2-kb *Bam*HI/*Eco*RI was verified by sequence analysis.

Chimeras G and H were generated according to the splicing by overlap extension (SOE) procedure (12). For both constructs, two primers lying at the N terminus of Ubx (UbxN-ter: 5′-CAATGAACTCGTACTTTGAACAGGC) and at the C terminus of AbdA (AbdA C-ter: 5′-CTTAGGAGTTGACTTTGCTGACCGCC) were used. For chimera G, overlapping junction primers were Ubx junction, 5′-CTTCGTGGACCATTTGTACCTAGCCAGTC; AbdA junction, 5′-GGTACAAATGGTCCACGAAGGCGCGGTC. For chimera H, overlapping junction primers were Ubx junction, 5′-GCCCCTGGCATCGAGATGGTGCAATTGGCAT; AbdA junction, 5′-CATCATCTCGATGCCAGGGGCAGGGGGAGC.

Products of the SOE reaction were cloned in the PgemT-easy vector (Promega) and sequenced to verify the gene fusions. Chimeras G and H were introduced into the P-Casper Hsp70 vector by cloning the *Spe*I Klenow-filled/*Eco*RI fragments containing the gene fusions into the *Eco*RI/*Stu*I sites of P-Casper Hsp70.

Plasmid DNA for each construct was prepared and used for P element-mediated germ line transformation, as described by Rubin and Spradling (13). The P insertions were genetically mapped, and at least two balanced lines were established for each chimera, except for chimeras E and H, where only single lines were recovered, presumably because of the interfering potency of these proteins.

Flies, Egg Collections, Heat Shock Conditions, and Preparation of Cuticles. Oregon R was used as a standard. The Ubx^{C1} stock was provided by M. Akam (Cambridge Univ., Cambridge, U.K.) and J. Casanova (Consejo Superior de Investigaciones Científicas, Barcelona). The *HS.dpp* stock was obtained from the Bloomington Stock Center. To analyze the activity of chimeras in the epidermis, embryos from transgenic chimeric lines were collected for 3 h, aged for 3 h, dechorionated in 50% bleach for 90 s, and heat-shocked for 30 min at 37°C in a water bath. Embryos were allowed to develop for 30–36 h and subsequently mounted in Hoyer's mountant (14). All cuticle preparations were examined and photographed with dark-field optics. To analyze the activity of chimeras in the VM, embryos from transgenic chimeric lines or *UbxC1; HS.dpp* were collected and prepared as described by Mathies *et al.* (15).

***In Situ* Hybridization to, and Immunostaining of, Whole-Mount Embryos.** *In situ* hybridization to whole embryos using digoxigenin DNA and RNA-labeled probes was performed according to Tautz and Pfeiffle (16) and Vincent *et al.* (17). After alkaline phosphatase detection, embryos were mounted in 90% glycerol/100 mM Tris (pH 7.5) and observed under an Axiophot Zeiss microscope using Nomarski optics. Digoxigenin RNA- and DNA-labeled probes were generated according to the manufacturer's protocol (Roche Molecular Biochemicals). Template DNAs to prepare labeled RNA probes were as follows: a Bluescript KS containing the *wg* cDNA digested with *Xba*I and reverse transcribed with T3 RNA polymerase; a Bluescript KS containing the *dpp* cDNA digested with *Bam*HI and reverse transcribed with T3 RNA polymerase. A DNA-labeled probe recognizing only the N-terminal part of *abdA* was generated by using a 637-bp *Eco*RI/*Pst*I fragment as template. Immunostainings were performed according to Alexandre *et al.* (18). The monoclonal antibodies Dm.Abd-A.1 and Ubx FP3.38 were generous gifts from D. Mattson-Duncan (Washington Univ., St. Louis) and R. White (Cambridge Univ., Cambridge, U.K.). Both antibodies recognize an epitope lying in the N-terminal part of each protein and were used at 1/1,000 dilution and 20 μg/ml, respectively.

Results

Design of Ubx/AbdA Chimeras. As a guide to what protein sequences may be most important for AbdA specificity, we compared AbdA sequences from evolutionarily close species within the insect phylum. We observed that sequence conservation extends beyond the HD, including twelve amino acids C-terminal and adjacent to the HD. This region is also well conserved among Ubx proteins (19). Within this region, the ten first amino acids are mostly conserved between AbdA and Ubx (Fig. 1*A*). This sequence will be referred to as CterC for C-terminal conserved. Only a few residues distinguish AbdA from Ubx in the conserved HD and in the CterC, although the proteins look mostly different in other regions. The distinguishing residues are the HD amino acids 1, 12, 23, and 35 and amino acids 61, 62, 64, and 67 in the CterC (Fig. 1*A*).

Examination of AbdA protein sequences from the evolutionarily distant insects *Drosophila melanogaster* and *Tribolium castaneum* shows that the 40 amino acids preceding the HD have been significantly conserved in AbdA. This region, referred to as NterC for N terminal conserved (Fig. 1*A*), contains the hexapeptide, a motif required for appropriate Hox-Exd/Pbx interaction (20). *D. melanogaster* and *T. castaneum* AbdA proteins have the same NterC amino acids at 63% of the positions; that number increases to 78% with conservative changes taken into consideration. This score is significantly higher than the identity/similarity score (34%/47%) found in the remaining N-terminal sequences (N-ter). Apart from the hexapeptide, the NterC region of Ubx and AbdA have largely diverged.

We reasoned that the few residues that differ between Ubx and AbdA in otherwise evolutionarily conserved sequences would be good candidates for specificity control. To identify sequence elements critical for AbdA character, conserved parts of the AbdA sequence were introduced into the Ubx protein and analyzed for their ability to confer AbdA-like activity. Chimeras A–F (Fig. 1*B*) contain combinations of point mutations introduced into Ubx to test the functional importance of the HD and CterC residues that are found specifically in AbdA. Chimeras G and H are protein sequence switches addressing the function of the NterC and of the region following the AbdA CterC sequence (C-ter). All of the protein-coding constructs were fused to a heat-inducible promoter and introduced into the fly by using *P*-mediated germ-line transformation. Eggs from these transgenic lines were immunostained with an antibody directed against the N-terminal sequences of Ubx, allowing detection of chimeras A–H. After heat induction, all chimeric proteins accumulate at similar levels in the nucleus (data not shown).

The HD and CterC Region Confer AbdA Specificity in the Epidermis. Eight amino acids distinguish Ubx and AbdA within the HD and CterC region. Chimera A consists of a Ubx protein where all eight amino acids have been changed into the corresponding AbdA residues. The effects on cuticle development of ubiquitous expression of chimera A were compared with those of ubiquitous expression of *Ubx* and *abdA* (21, 22). The number, identity, and spatial organization of denticles readily distinguish these two segments (Fig. 2*A*). Uniform *Ubx* expression transforms anterior segments into extra A1 segments (Fig. 2*B*), whereas *abdA* turns them into A2 (Fig. 2*C*). Chimera A has the same effect as ubiquitous *abdA* expression (Fig. 2*D*). The three thoracic segments (T1–T3) and the first abdominal segment A1 have been transformed toward an A2 identity. In the epidermis, therefore, chimera A has acquired AbdA-like character.

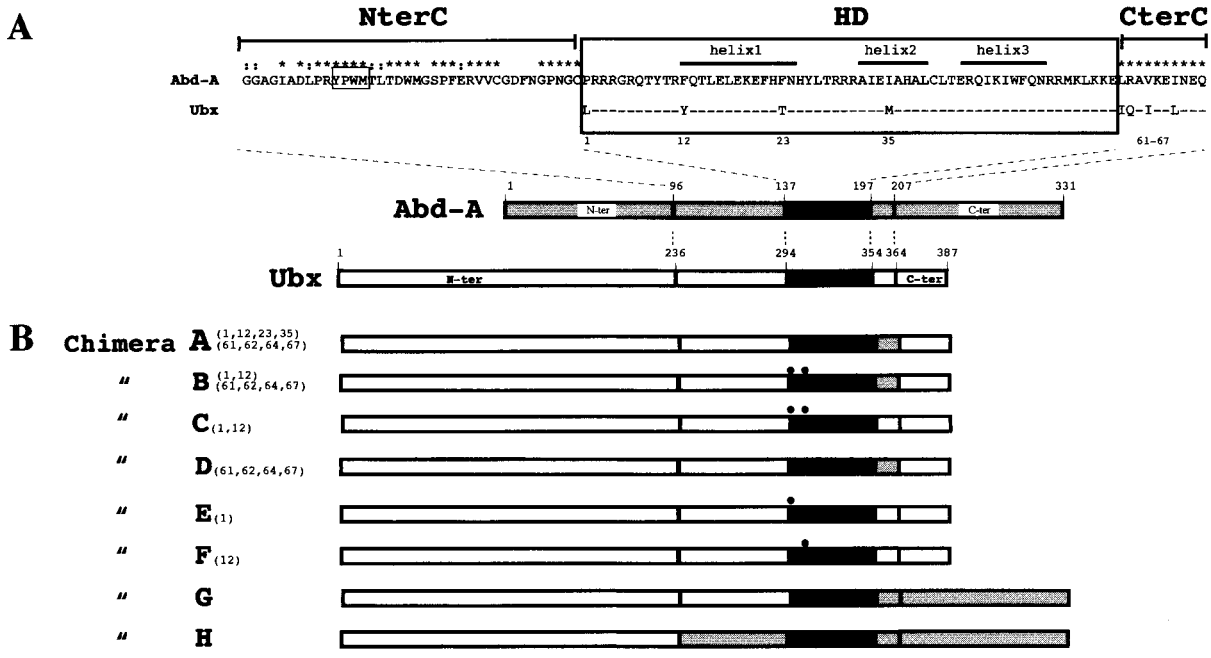

Fig. 1. Design of the Ubx/AbdA chimeras. (*A*) The AbdA and Ubx proteins are divided in five portions designated N-ter, NterC, HD, CterC, and C-ter. Dashes indicate conserved residues between *Drosophila melanogaster* Ubx and AbdA proteins. *, Identities, and :, similarities between *D. melanogaster* and *Tribolium castaneum* AbdA proteins. (*B*) Diagrams of Ubx/AbdA chimeric proteins. Dotted boxes represents AbdA sequences. The numbers in parentheses from chimeras A–F refer to the amino acids that have been changed to those found in equivalent position in AbdA.

Segments posterior to A2 are not affected by ubiquitous expression of either AbdA or chimera A. This is similar to other cases of "posterior prevalence" (23), where more posterior Hox functions suppress the activity (not just expression) of anterior *Hox* genes. We noticed a slight difference between the consequences of ectopic *abdA* and chimera A expression. *abdA* transforms all thoracic segments and A1 equally well (Fig. 2*C*), but chimera A preferentially transforms T1 and A1 (Fig. 2*D*). The segments that are less completely transformed, T2 and T3, correspond to the functional domain of Antennapedia (Antp) protein, suggesting that the activity of chimera A is partially suppressed by Antp.

The HD and CterC Region Do Not Confer AbdA Specificity in the VM. *Ubx* and *abdA* differentially activate the target genes *wg* and *dpp* during midgut morphogenesis (24–27). In the midgut mesoderm, *dpp* is expressed in parasegment (PS) 3–4 and PS7, with Ubx activating it in PS7. When *Ubx* is expressed ubiquitously, *dpp* is ectopically activated in PS anterior to PS7 (Fig. 3*A*). Ubiquitous *abdA* expression represses *dpp* in all of the VM (Fig. 3*C*). Ectopic *Ubx* does not affect *wg* transcription that continues to occur only in its normal place, PS8 (Fig. 3*B*). Ubiquitous *abdA* expression does induce ectopic *wg* transcription in anterior regions (Fig. 3*D*). Chimera A behaves as Ubx, because it still activates *dpp* (Fig. 3*E*) while leaving *wg* expression unaffected (Fig. 3*F*). We noted, however, that the activation of *dpp* by chimera A does not occur in all VM cells anterior to PS7. This suggests that the chimera is transcriptionally less potent and/or that it is more sensitive to phenotypic suppression by the more anteriorily expressed *Scr* and *Antp* genes. In any case, these results unambiguously show that in the

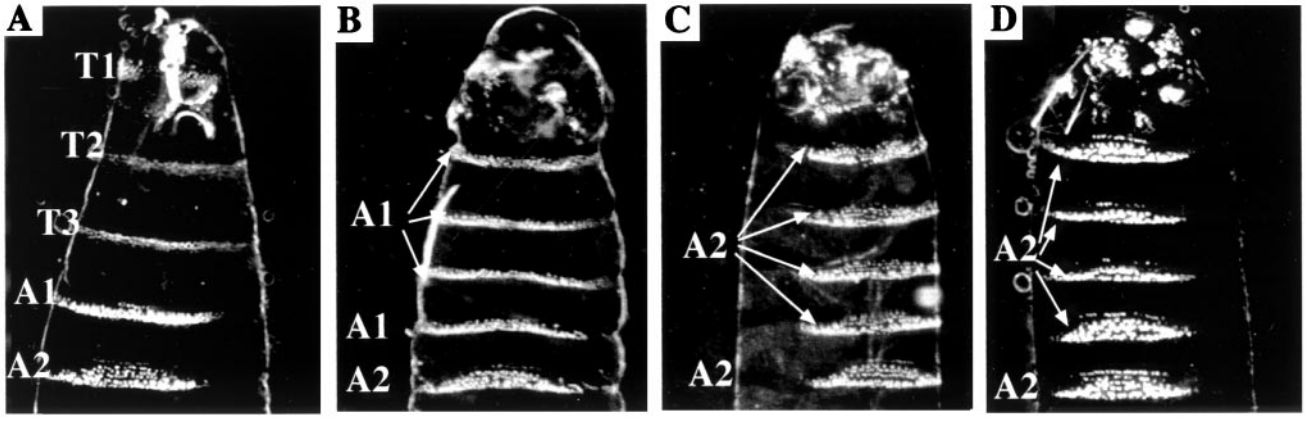

Fig. 2. The HD and CterC region confer AbdA activity in the epidermis. (*A*) Wild-type cuticle indicating the morphology of the thoracic (T1, T2, T3) and first two abdominal segments (A1, A2). Cuticular transformations associated with ubiquitous expression of *Ubx* (*B*), *abdA* (*C*), and chimera A (*D*). *Ubx* and *abdA*, respectively, transform segments anterior to their normal expression domains in A1 and A2 metameres. Chimera A phenocopies AbdA: it transforms the identity of thoracic and the first abdominal segments in an A2 identity.

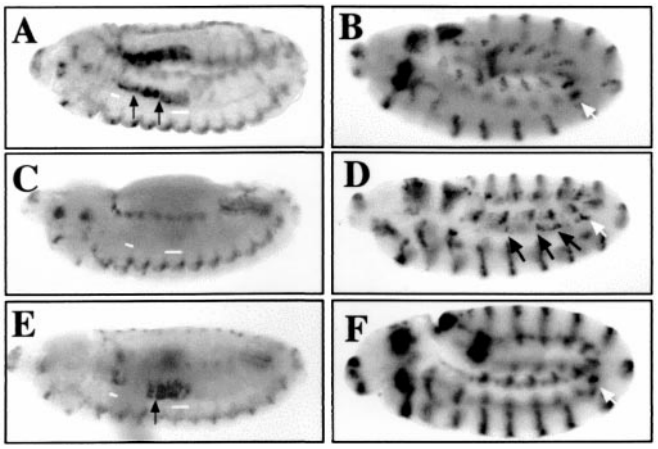

Fig. 3. The HD and CterC region do not confer AbdA character in the VM. Effects of uniform expression of *Ubx*, *abdA*, and chimera A on *dpp* and *wg* transcription in the VM. *Ubx* induces ectopic *dpp* expression (*A*) and leaves *wg* expression unaffected (*B*). *abdA* represses *dpp* (*C*) and induces anterior ectopic expression of *wg* (*D*). White bars and arrows indicate the wild-type expression domain of *dpp* (PS3–4 and PS7) and *wg* (PS8), respectively. Black arrows show the sites of ectopic expression. In the VM, chimera A phenocopies the effects of uniform *Ubx* expression: it ectopically activates *dpp* (*E*), although somewhat less efficiently, whereas leaving *wg* expression unaffected (*F*).

VM, chimera A retains Ubx specificity and has not acquired, as in the epidermis, an AbdA-like character.

The HD N-Terminal Arm and the CterC Region Cooperatively Define AbdA Function in the Epidermis. The activities of chimeras B–F in the epidermis were studied to assess the individual contributions of the eight amino acids that distinguish AbdA from Ubx in the HD and CterC region. Changing only one residue in the N-terminal arm of the HD, either amino acid 1 (chimera E) or 12 (chimera F), is not sufficient to give the protein AbdA character in epidermal patterning. Both chimeras induce, like Ubx, ectopic A1 metameres (Fig. 4 *A* and *B*). However, simultaneously changing both amino acids 1 and 12 (chimera C) directs the formation of A2-like segments in place of thoracic and A1 segments (Fig. 4*C*). The transformation toward an A2 identity is less complete than with chimera A (Fig. 2*D*). Such weak A2-like transformations are also observed when the changes in the protein concern the four residues in the CterC region (chimera D; Fig. 4*D*). Among chimeras B–F, the only protein that has an efficient AbdA effect is chimera B (Fig. 4*E*). This indicates that amino acids 1 and 12 of the HD, as well as amino acids 61, 62, 64, and 67 of the CterC region, are collectively required for full-potency AbdA effects. HD residues 23 and 35 appear to be unnecessary for AbdA-like activity.

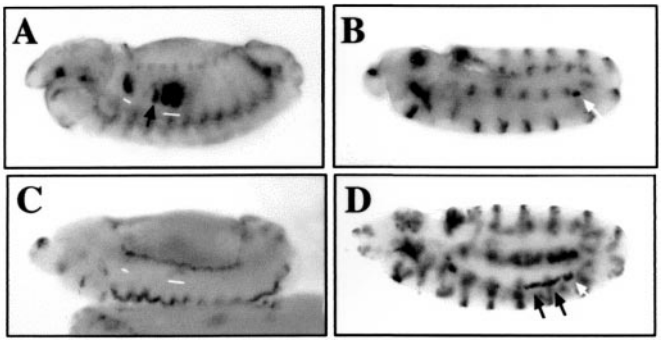

Fig. 5. The NterC region defines AbdA activity in the VM. Effects of ubiquitous expression of chimeras G and H were analyzed in the VM. (*A* and *B*) Chimera G still retains Ubx activity: although less efficiently than AbdA, it activates *dpp* (*A*) and leaves *wg* expression unaffected (*B*). (*C* and *D*) Chimera H has gained AbdA activity: it represses *dpp* (*C*) and ectopically induces *wg* expression (*D*).

AbdA Function in the VM Requires the Hexapeptide-Containing NterC Region. Because chimera A does not have AbdA-like effects in the VM, we investigated the possible contributions of two additional protein sequences. The importance of the AbdA C-ter region, which is significantly longer than the Ubx C-ter region, was tested by using chimera G. In chimera G, the Ubx HD and downstream sequences are replaced by AbdA sequences. Chimera G behaves like a Ubx protein in the VM: it retains the ability to activate *dpp* while leaving *wg* expression unaffected (Fig. 5 *A* and *B*). Sequences downstream of the AbdA HD are therefore not sufficient to convey AbdA function in the VM.

The function of the NterC sequence was analyzed by using chimera H, which has Ubx sequences through amino acid 234 and AbdA sequences thereafter, including the AbdA NterC, HD, and C-tail. Chimera H therefore tests the contribution of the NterC region to AbdA activity in the VM, particularly in comparison to chimera G. The results show that chimera H has AbdA effects but not Ubx effects: it represses *dpp* expression and induces anterior ectopic expression of *wg* (Fig. 5 *C* and *D*). In the VM, therefore, the hexapeptide-containing NterC sequence confers AbdA activity, in

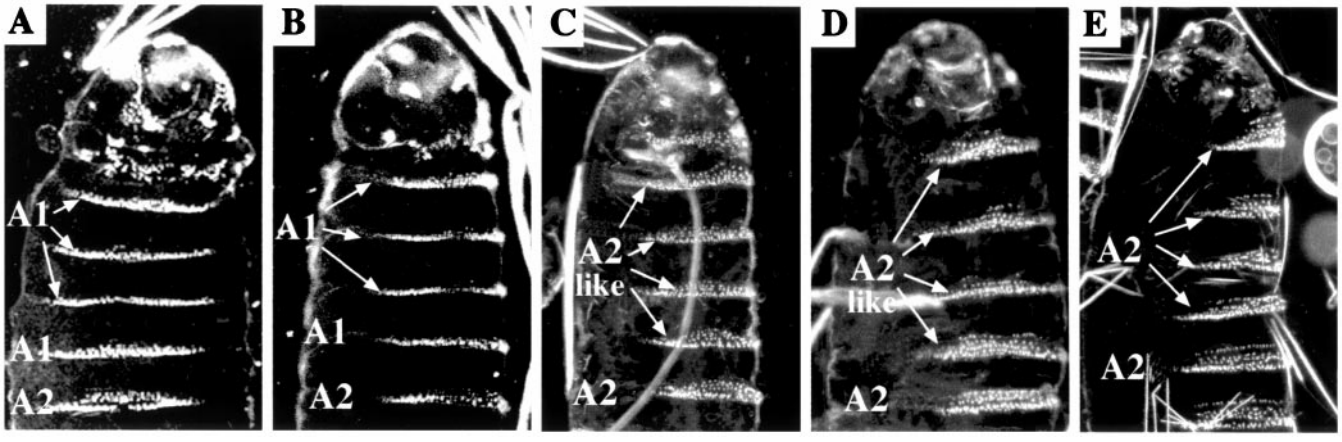

Fig. 4. HD residues 1 and 12 cooperate with the CterC region to define AbdA activity in the epidermis. Effects of ubiquitous expression of chimeras B–F were analyzed in the epidermis. (*A* and *B*) Chimeras E and F transform thoracic metameres into A1 segments. (*C* and *D*) Chimeras C and D impose weak A2-like transformations to segments anterior to A2. (*E*) Chimera B behaves as AbdA: segments anterior to A2 acquire an A2 identity.

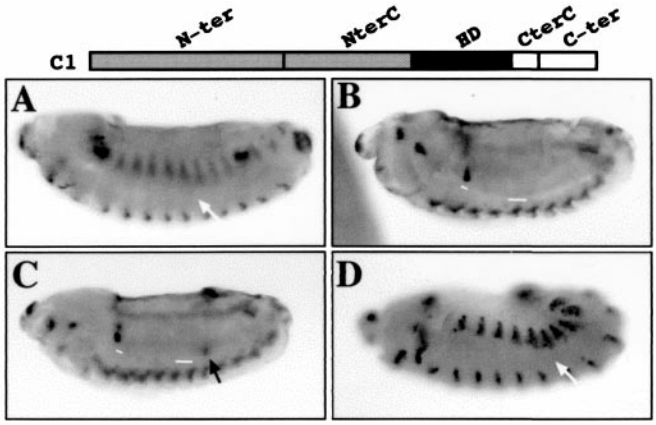

Fig. 6. The NterC region requires additional HD and/or C-terminal sequences to confer AbdA mesodermal activity. The effect of endogenous expression of the C1 fusion product (the structure of the protein is shown, *Top*) on *wg* and *dpp* expression in the VM was analyzed. (*A*) *wg* is not expressed in homozygous *UbxC1*. (*B* and *C*) Most Homozygous Ubx^{C1} embryos show a severe reduction or absence of *dpp* expression in the central midgut VM, whereas expression in the anterior midgut is not altered (*B*). A minority exhibits some posterior ectopic expression in the expression domain of the C1 fusion protein. (*C* and *D*) Providing high level of Dpp signaling with a heat-inducible transgene (*HS.dpp*) to homozygous Ubx^{C1} embryos does not restore *wg* expression.

contrast to the HD and CterC sequences important for AbdA character in the epidermis.

Contribution of the HD and/or C-Terminal Sequences to the Function of the NterC Sequence. Is the NterC region on its own sufficient to confer to an otherwise Ubx protein an AbdA activity in the VM? The Ubx^{C1} mutation is a chromosome rearrangement that leads to an AbdA/Ubx fusion product (28). The resulting protein consists of AbdA N-ter and NterC sequences joined to the Ubx HD, CterC, and C-ter. The hybrid gene is expressed in the posterior VM in the *abdA* expression domain (29). The mutation can therefore be used to determine whether the AbdA NterC region is sufficient to provide AbdA activity in the VM. In homozygous Ubx^{C1} embryos, *wg* is not activated in VM PS8 by the C1 fusion product (Fig. 6*A*). In the same mutant context, expression of *dpp* in the anterior VM (PS3–4) is never affected, whereas expression in the central part of the midgut is severely reduced or abolished (Fig. 6*B*). A minority of mutant embryos display posterior ectopic expression within the expression domain of the fusion protein C1 (Fig. 6*C*). These observations indicate that the C1 fusion protein has retained some transcriptional activity in the VM and that the level of Dpp signaling is very low in most mutant embryos.

wg activation by *abdA* critically depends on Dpp signaling in the central midgut (24). The absence of *wg* expression in Ubx^{C1} homozygous embryos might thus be the result of reduced Dpp signaling rather than the inability of the C1 fusion protein to activate *wg*. To discriminate between these possibilities, we analyzed *wg* expression in homozygous Ubx^{C1} embryos, where a high level of Dpp activity was provided by a heat-inducible transgene (*HS.dpp*). Even in such a context, no *wg* expression is observed (Fig. 6*D*). Taken together, these experiments indicate that the C1 fusion protein is not capable of providing AbdA-like function in the VM. The AbdA NterC region, although necessary to impart AbdA-like function, is not sufficient on its own. Additional AbdA sequences lying in the HD and/or C-terminal sequences are required as well.

The Endogenous *abdA* Gene Does Not Mediate Mesodermal and Epidermal Activity Switches. The effects of ectopically expressed Hox or Hox chimeric proteins depend in some instances on ectopic activation of endogenous *Hox* gene activity (30, 31). We tested whether the switches to AbdA-like function obtained with chimera A in the epidermis and chimera H in the VM require the endogenous *abdA* gene. Embryos uniformly expressing chimera A or chimera H were stained with a probe corresponding to the N-ter sequences of AbdA, allowing the detection of the endogenous *abdA* gene exclusively. In both cases, no ectopic expression of *abdA* was observed, either in the epidermis or in the VM (data not shown). The AbdA-like activities of chimeras A and H therefore do not rely on activation of the endogenous *abdA* gene. Instead, both chimeras have acquired an AbdA activity.

Discussion

Importance of Sequences Lying Outside of the HD for Hox Functional Specificity. Most of the Hox protein specificity information has been found to lie within the HD. Indeed, in all but two of the chimeras already analyzed, the HD must be switched to change a Hox protein's function into that of another Hox protein (19, 32–34). Residues in the N-terminal arm of the HD have been shown necessary and sometimes sufficient to define functional specificity (19, 32–34). In agreement with these findings, our results indicate that amino acids 1 and 12, in the N-terminal arm, significantly contribute to defining AbdA character.

For several reasons, the idea that the HD contains most of the specificity information appears overemphasized. First, the switch of specificity is often only partial. For example, when the Deformed (Dfd) HD is replaced by that of Ubx, the chimeric protein activates the *Antp* gene, a target of Ubx, whereas failing to carry out a Dfd function, autoactivation of *Dfd* transcription. The chimeric protein has therefore acquired the target specificity of Ubx. However, the chimera did not have all of the regulatory specificity of Ubx because Ubx normally represses *Antp* (32).

Second, the switch of the HD alone sometimes does not confer an identity switch: replacing the HD of Ubx by that of Antp results in a chimera that behaves like Ubx, promoting A1 identity in the epidermis (19). Similarly, the effect of a Hox-A4/Hox-C8 chimera on vertebral patterning does not follow the identity of the HD (35). Our Ubx/AbdA chimera data show that, *in vivo*, switching the HD is not necessary for changing Hox protein specificity. Chimera A activates *dpp* in the VM and has no effects on *wg* expression, indicating that it has Ubx character, despite having the AbdA HD.

Third, sequences outside the HD have been shown to provide important function for *Hox* gene activity. This includes the C-tail for Ubx, Antp, Scr, and Dfd (19, 32, 33) along with an acidic N-region preceding the HD in the Dfd protein (36). Our results demonstrate that two sets of non-HD sequences contribute to Hox specificity. On the C-terminal side, four residues adjacent to the HD are necessary to impart AbdA-like character in the epidermis. This region is evolutionarily conserved, has been proposed to be part of a coiled–coil structure, and might constitute an interface for interacting proteins. On the N-terminal side, the NterC region is required to define mesodermal specificity.

Intrinsic Determinants of Hox Tissue Specificity: Role of the Hexapeptide-Containing NterC Region. Hox proteins clearly have distinct functions in different tissues, but the biological mechanisms underlying the tissue specificity remain unknown. In fact, this holds more generally for the restriction of widely expressed regulatory activities (37). Two mechanistic hypotheses are commonly proposed. The first one assumes the existence of tissue-specific cofactors, whereas the second one postulates that tissue-specific alteration of chromatin structure plays an important role (38). No tissue-specific Hox cofactor has yet been identified, and experimental support for the second hypothesis is limited (39). If Hox tissue-specific cofactors exist, it should be possible to identify parts of Hox proteins involved in the interaction with such proteins. Functional dissections of Hox proteins, using chimeric approaches, have in principle the potential to identify protein determinants of tissue-specific Hox actions.

In contrast to previous chimeras studied, Ubx/AbdA chimeras

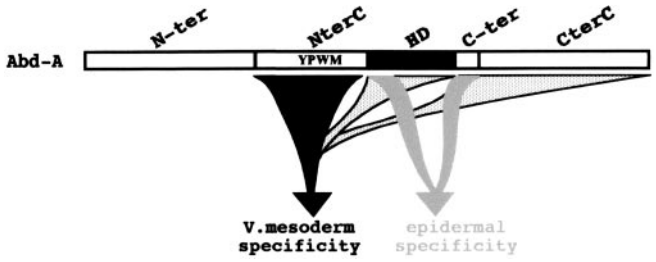

Fig. 7. Distinct regions cooperate to define AbdA specificity in the epidermis and VM. Amino acids 1 and 12 of the HD and the CterC sequence together define the AbdA epidermal specificity. The NterC hexapeptide-containing region defines the mesodermal specificity of AbdA. It, however, requires additional HD and/or C-terminal sequences for accurate mesodermal activity.

have special usefulness because available genetic markers allow an assessment of distinctive regulatory functions of the parental Hox proteins in distinct tissues. Our conclusion is that different protein sequences contribute to AbdA character in the epidermis and VM (Fig. 7). The HD and CterC region are sufficient to impart AbdA activity in the epidermis, and the hexapeptide-containing NterC region is necessary to convey AbdA-like character in the VM. The NterC sequence is not sufficient to confer AbdA character on its own. The dramatic functional differences exhibited by chimeras G, H, and the fusion protein C1 show that the HD and/or sequences C-terminal to the HD (CterC and C-ter) are also required to create AbdA specificity in the VM.

The importance of Exd for AbdA activity in the VM (10, 40) and the role of residues lying within the HD for the Hox/Exd interaction (20, 41) suggest that the HD most likely contributes to AbdA mesodermal specificity. The existence of intrinsic structural determinants in AbdA suggests that tissue-specificity is achieved by the differential readout of its protein structure in different tissues. In a simple scenario, a prediction would be that this differential readout is achieved by tissue-specific cofactors that remain to be discovered. The NterC domain constitutes an attractive bait for the identification of such proteins.

The role of the hexapeptide YPWM in the formation of Hox-Exd/Pbx heterodimers *in vitro* has been extensively studied (4, 41), yet its *in vivo* function has not been established. An interesting feature of the NterC region is that it contains the hexapeptide motif. The requirement of Exd for target gene regulation in the VM suggests that the motif is involved in AbdA mesodermal function. Several lines of evidence, however, do not favor the hypothesis that the YPWM motif itself provides the key for AbdA VM specificity. First, Ubx and AbdA share the YPWM motif, so sequences other than the motif itself must contribute to distinguishing AbdA from Ubx. Second, Exd that contacts the motif is required for Hox protein activity in several tissues. Third, the conservation within NterC regions of AbdA proteins evolutionarily as distant as *D. melanogaster* and *T. castaneum* is not restricted to the hexapeptide motif. One would thus expect these evolutionarily conserved sequences in AbdA, that have diverged in Ubx, to provide the VM-specificity information. The NterC domain therefore likely coordinates the contribution of Exd and putative tissue-specific cofactors for accurate AbdA mesodermal activity.

We thank M. Akam and R. White for comments on the manuscript, M. Akam for sending the Ubx^{C1} stock, I. Duncan and D. Mattson-Duncan for the monoclonal antibody Dm.AbdA.1, and K. Matthews and the Bloomington Stock Center for the *HS.dpp* stock. We also thank S. Sohn for his contribution to the early part of this study. This work was supported by the Centre National de la Recherche Scientifique and grants from l'Association pour la Recherche contre le Cancer (ARC) and La Ligue Nationale Contre le Cancer (LNCC). D.B. was supported by a National Science Foundation predoctoral grant, and M.P.S. is an Investigator of the Howard Hughes Medical Institute.

1. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68,** 283–302.
2. Gehring, W. J., Qian, Y. Q., Billeter, M., Furokubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. & Wütrich, K. (1994) *Cell* **78,** 211–223.
3. Hayashi, S. & Scott, M. P. (1990) *Cell* **63,** 883–894.
4. Mann, R. S. & Chan, S.-K. (1996) *Trends Genet.* **12,** 258–262.
5. Pinsonneault, J., Florence, B., Vaessin, H. & McGinnis, W. (1997) *EMBO J.* **16,** 2032–2042.
6. Li, X., Murre, C. & McGinnis, W. (1999) *EMBO J.* **18,** 198–211.
7. Mann, R. S. & Affolter, M. (1998) *Curr. Opin. Genet. Dev.* **8,** 423–429.
8. Ryoo, H. D., Marty, T., Casares, F., Affolter, M. & Mann, R. S. (1999) *Development (Cambridge, U.K.)* **126,** 5137–5148.
9. Tremml, G. & Bienz, M. (1989) *EMBO J.* **8,** 2677–2685.
10. Bilder, D., Graba, Y. & Scott, M. P. (1998) *Development (Cambridge, U.K.)* **125,** 1781–1790.
11. Graba, Y., Aragnol, D. & Pradel, J. (1997) *BioEssays* **19,** 379–388.
12. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989) *Gene* **77,** 61–68.
13. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218,** 348–353.
14. van der Meer, J. M. (1977) *Drosophila Inf. Serv.* **52,** 160.
15. Mathies, L. C., Kerridge, S. & Scott, M. P. (1994) *Development (Cambridge, U.K.)* **120,** 2799–2809.
16. Tautz, D. & Pfeiffle, C. (1989) *Chromosoma* **98,** 81–85.
17. Vincent, F., Solloway, M., O'Neil, J. W., Emery, J. & Bier, E. (1994) *Genes Dev.* **8,** 2602–2616.
18. Alexandre, E., Graba, Y., Fasano, L., Gallet, A., Perrin, L., De Zulueta, P., Pradel, J., Kerridge, S. & Jacq, B. (1996) *Mech. Dev.* **59,** 191–204.
19. Chan, S. K. & Mann, R. S. (1993) *Genes Dev.* **7,** 796–811.
20. Chan, S. K., Jaffe, L., Capovilla, M., Botas, J. & Mann, R. S. (1994) *Cell* **78,** 603–615.
21. Mann, R. S. & Hogness, D. S. (1990) *Cell* **60,** 597–610.
22. Sanchez-Herrero, E., Guerrero, I., Sampedro, J. & Gonzalez-Reyes, A. (1994) *Mech. Dev.* **46,** 153–167.
23. Gonzalez-Reyes, A. & Morata, G. (1990) *Cell* **61,** 515–522.
24. Immerglück, K., Lawrence, P. A. & Bienz, M. (1990) *Cell* **62,** 261–268.
25. Reuter, R., Panganiban, G. E. F., Hoffmann, F. M. & Scott, M. P. (1990) *Development (Cambridge, U.K.)* **110,** 1031–1040.
26. Capovilla, M., Brandt, M. & Botas, J. (1994) *Cell* **76,** 461–475.
27. Capovilla, M. & Botas, J. (1998) *Development (Cambridge, U.K.)* **125,** 4949–4957.
28. Casanova, J., Sanchez-Herrero, E. & Morata, G. (1988) *EMBO J.* **7,** 1097–1105.
29. Rowe, A. & Akam, M. (1988) *EMBO J.* **7,** 1107–1114.
30. Kuziora, M. A. & McGinnis, W. (1988) *Cell* **55,** 477–485.
31. Kuziora, M. A. (1993) *Mech. Dev.* **42,** 125–137.
32. Lin, L. & McGinnis, W. (1992) *Genes Dev.* **6,** 1071–1081.
33. Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. & Scott, M. P. (1993) *Development (Cambridge, U.K.)* **118,** 339–352.
34. Furukubo-Tokunaga, K., Flister, S. & Gehring, W. J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6360–6364.
35. Sreenath, T. L., Pollock, R. A. & Bieberich, C. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 9636–9640.
36. Zhu, A. H. & Kuziora, M. A. (1996) *Development (Cambridge, U.K.)* **122,** 1577–1587.
37. de Celis, J. F. (1999) *BioEssays* **21,** 542–545.
38. Getzenberg, R. H. (1994) *J. Cell. Biochem.* **55,** 22–31.
39. Ch'ng, Q. & Kenyon, C. (1999) *Development (Cambridge, U.K.)* **126,** 3303–3312.
40. Rauskolb, C. & Wieschaus, E. (1994) *EMBO J.* **13,** 3561–3569.
41. Johnson, F. B., Parker, E. & Krasnow, M. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 739–743.