Direct regulation of the muscle-identity gene *apterous* by a Hox protein in the somatic mesoderm

Maria Capovilla¹,*, Zakaria Kambris¹ and Juan Botas²

¹Institut de Biologie Moléculaire et Cellulaire, UPR 9022 du CNRS, 15 rue René Descartes, 67084 Strasbourg, France
²Departments of Molecular and Human Genetics and Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

*Author for correspondence (email: m.capovilla@ibmc.u-strasbg.fr)

Accepted 21 January; published on WWW 22 March 2001

SUMMARY

Hox genes control segment identity in the mesoderm as well as in other tissues. Most evidence indicates that Hox genes act cell-autonomously in muscle development, although this remains a controversial issue. We show that *apterous* expression in the somatic mesoderm is under direct Hox control. We have identified a small enhancer element of *apterous* (apME680) that regulates reporter gene expression in the LT1-4 muscle progenitors. We show that the product of the Hox gene *Antennapedia* is present in the somatic mesoderm of the second and third thoracic segments. Through complementary alterations in the *Antennapedia* protein and in its binding sites on apME680, we show that *Antennapedia* positively regulates *apterous* in a direct manner, demonstrating unambiguously its cell-autonomous role in muscle development. Finally, we determine that LT1-4 muscles contain more nuclei in the thorax than in the abdomen and we propose that one of the segmental differences under Hox control is the number of myoblasts allocated to the formation of specific muscles in different segments.

Key words: *antennapedia, apertous, Drosophila, Hox, Muscles*

INTRODUCTION

Hox genes encode transcription factors that have been highly conserved through evolution. The phenotypic analysis of their loss- and gain-of-function mutations has shown that Hox genes are required in all animals to establish cellular identities specific for each segment along the rostrocaudal axis (reviewed by Botas, 1993; Kenyon, 1994; McGinnis and Krumlauf, 1992) and to pattern the vertebrate limb (reviewed by Duboule, 1992). In *Drosophila*, Hox genes were first identified through their role in patterning the ectoderm, the tissue most accessible to phenotypic analysis, where they function in a cell-autonomous fashion (reviewed by Lawrence and Morata, 1994). It was later found that Hox genes have a primary role in determining the differences between segments also in internal tissues. In the peripheral nervous system (PNS) and in the central nervous system (CNS), Hox mutations affect the identity of specific neurons (Heuer and Kaufman, 1992; Prokop et al., 1998; Ghysen et al., 1985). It is not yet known whether these effects are completely cell autonomous. In the visceral mesoderm, *Antennapedia* (Antp), *Ultrabithorax* (Ubx) and *abdominal-A* (abd-A) are required for the formation of specific constrictions at precise positions along the midgut, corresponding to the domains in which each gene is expressed. UBX and ABD-A have been shown to directly regulate the *decapentaplegic* (*dpp*) gene, which is required for formation of the second midgut constriction, demonstrating their cell-autonomous function in midgut morphogenesis (Capovilla et al., 1994; Capovilla and Botas, 1998). As *dpp* expression in the visceral mesoderm is required for *labial* expression in the underlying endoderm (Immerglück et al., 1990), *Ubx* and *abd-A* have also a non cell-autonomous function in this tissue. We will focus on Hox expression and function in somatic mesoderm development.

*Drosophila* muscles are multinucleated cells derived from the fusion of a special class of myoblasts called ‘founder cells’ with neighboring fusion-competent myoblasts (reviewed by Bate et al., 1999). At the end of embryogenesis, each myofiber presents unique characteristics with respect to its position, size, orientation, innervation and attachment to the epidermis. The pattern of muscles has been clearly described (Bate, 1993). The number and identity of muscles differ along the rostrocaudal axis of the trunk. A common set of approximately 30 muscles is present in each trunk hemisegment, with the exception of the first thoracic segment (T1) and the eighth abdominal segment (A8), which present fewer and more diversified muscles. Superimposed on the muscle scaffold common to the T2-A7 segments is a muscle pattern specific for T2, T3, A1 and A2-A7. So far this has been determined through the presence or absence of specific muscles. For example, A1 bears a unique muscle VII whereas T2 lacks muscles VA1 and VO3 (nomenclature based on Bate, 1993). For reasons of simplicity, most of the studies on muscle development have been carried out on the A2-A7 segments, which present the same muscle pattern. Very little work has been done to determine how a segment-specific muscle pattern is established. It is likely that Hox genes are, directly or indirectly, involved in this process.
Hooper first showed that in Ubx mutant larvae the abdominal muscles are transformed into T3 muscles, with A1 and A2 being the most affected segments (Hooper, 1986). No alterations are observed in the third thoracic segment. This indicates that the realm of action of Ubx in the somatic mesoderm is shifted posteriorly with respect to the epidermis where Ubx functions primarily in the T3 and A1 segments (Lewis, 1978; Sánchez-Herrero et al., 1985). This result has been taken as indirect evidence for the cell-autonomous role of Ubx in muscle development, as migration of embryonic muscle precursors has never been observed. However, a shortcoming of this experiment is that in the Ubx mutant larvae analyzed all the tissues are transformed. In addition, the phenotypes were observed at a late stage, when many cellular interactions have occurred in order to complete muscle differentiation.

Additional evidence for the cell-autonomous functions of Hox genes in muscle development comes from the effects of their overexpression specifically in the embryonic mesoderm. The ubiquitous mesodermal overexpression of abd-A, normally expressed in the posterior abdomen, leads to the appearance of abdominal-type muscle progenitors (Greig and Akam, 1993; Michelson, 1994) and of abdominal-type muscles (Michelson, 1994) in the thorax. Similarly, the ectopic presence of UBX in the thoracic mesoderm results in a thoracic to abdominal transformation of muscles and muscle precursors (Michelson, 1994). Overexpression did not cause transformation of the overlying ectoderm. These mesoderm-specific ectopic expression studies indicate that Hox genes have the ability to act cell-autonomously in the mesoderm, but do not demonstrate whether they do so in the wild-type animal.

In spite of this evidence for Hox genes functioning cell-autonomously, some observations suggest that Hox expression in the nervous system and in the epidermis influence the formation of specific muscles. In adults, the development of a male-specific muscle depends on Hox function in its innervating motoneuron and on the segmental identity of the epidermis (Lawrence and Johnston, 1986). As shown by the study of the mutant four-winged fly, the number of myoblasts and their migration pattern in the haltere disc are controlled by Ubx expression in the ectoderm (Fernandes et al., 1994). Finally, it has been reported that muscle development in T2 depends on Antp expression in the embryonic ectoderm (Roy et al., 1997). Thus the autonomous versus non-autonomous role of Hox genes in muscle development is still a matter of controversy.

Muscle progenitors and founders present a stereotyped

---

**Fig. 1.** apME680 is a muscle-specific enhancer of ap. The arrows point to the first abdominal segment. Detection of ap transcripts by in situ hybridization (A,B) and of β-galactosidase by immunohistochemistry (C-E). (A) Stage 10 embryo. In addition to segmental clusters composed of PNS and muscle progenitors, ap is also expressed in the mandibular lobe and in the terminalia. (B) Stage 13 embryo. Expression is observed in the CNS, PNS and mesoderm. (C) Stage 11 apME680lacZ embryo. β-galactosidase is detected in the T2-A8 clusters. Note that thoracic clusters are bigger than abdominal clusters. The difference in size among the abdominal clusters is due to differences in the focal plane. (D) Stage 13 apME680lacZ embryo. The thoracic clusters are continuous, while each abdominal cluster has split in two smaller clusters, a dorsal and a ventral one. β-galactosidase is also detected in few cells in T1, but only in the lines with strongest expression. (E) Stage 16 apME680lacZ embryo. β-galactosidase accumulates in muscles LT1-4 (1-4). Thoracic muscles are stained more heavily than abdominal muscles. The arrowheads point to muscles LT4, which extend more dorsally and ventrally in the thorax than in the abdomen.

---

**Fig. 2.** Antp positively regulates apME680 in T2 and T3. Detection of β-galactosidase in wild-type (A,B) and AntpW10 (C-E) apME680lacZ embryos at stage 11 (A,C), at stage 13 (B,D) and at stage 16 (E). Fewer cells express lacZ in T2 and T3 in Antp mutants than in wild-type embryos. In Antp mutants, muscles LT1-4 are undetectable in segments T2 and T3 (E). The arrows point to the first abdominal segment.
pattern of gene expression which contributes to their identity. Several genes known as ‘muscle-identity genes’ are selectively expressed in specific progenitors/founders and have a role in the formation of specific muscles. These genes encode primarily transcription factors and include apterous (ap), S59, nautilus (nau), Krüppel, muscle segment homeobox (msh), even skipped, vestigial, ladybird and collier (knot – FlyBase; reviewed by Paululat et al., 1999; Bate et al., 1999; Frasch, 1999). It is likely that the segmental differences in the normal muscle pattern and the muscle transformations caused by the lack or abnormal expression of Hox genes are the consequence of an alteration in the expression of muscle-identity genes (Michelson, 1994).

ap encodes a LIM-homeodomain-containing protein best known for its role in dorsoventral patterning in the wing (Diaz-Benjumea and Cohen, 1993; Hobert and Westphal, 2000). In the embryo, ap is expressed in the CNS, in the PNS and in the somatic mesoderm, where its expression is limited to the muscle progenitors that contribute to the formation of muscles LT1-4, VA2 and VA3 (Bourginou et al., 1992). In ap mutants, a variable loss of these muscles is observed. Conversely, ap overexpression leads to the formation of ectopic LT1-LT3 muscles, suggesting that ap contributes to muscle identity (Bourginou et al., 1992).

Hox genes are thought to regulate directly or indirectly a large number of genes (reviewed by Pradel and White, 1998; Graba et al., 1997; Weatherbee et al., 1999). However, only a few Hox response elements that mediate direct regulation by Hox proteins have been identified and investigated in detail at a molecular level. In Drosophila these are decapentaplegic (dpp) (Capovilla et al., 1994; Capovilla and Botas, 1998), Distal-less (Dll) (Vachon et al., 1992), fork head (Ryoo and Mann, 1999), teashirt (McCormick et al., 1995) and the Hox genes themselves (Thuringer et al., 1993; Li et al., 1999; Grier et al., 1997; Dessain et al., 1992). These Hox response elements are tools required to answer many questions about Hox function that are still poorly understood – How does each Hox protein select specific target genes in order to build specific body structures? What is the molecular basis of the activating/repressing functions of Hox proteins? How do Hox proteins control pattern formation in different tissues?

We present the identification of a muscle-specific enhancer of ap and demonstrate that this element is directly regulated by ANTP in vivo. We also show that ANTP is normally present in the thoracic cells in which the enhancer is active. These results provide strong evidence for the cell-autonomous function of Hox proteins in muscle development. In addition, we propose that one of the differences between segments under Hox control is the number of myoblasts that give rise to specific muscles.

MATERIALS AND METHODS

Transformation constructs

apME680 is a 680 bp XhoI-Sau3A fragment that we obtained from phage ϕAP2D of the ap walk (Cohen et al., 1992) and is located in the second largest ap intron, based on available sequences (http://flybase.bio.indiana.edu:82). To generate wild-type P[apME680-lacZ] (apME680[5BCD-lacZ]), the apME680 fragment was cloned in the XhoI/BamHI sites of pBluescript KS+, excised with XhoI and XbaI, and subcloned in the XhoI/SpeI sites of P[CaSpeR-hs43-AUG-betaGal] vector (Chab; V. Pirrotta, unpublished observations). To generate the P[apME680.1-5BCD-lacZ] (apME680[5S-lacZ]) transgene, the mutated apME680 fragment cloned in pBluescript (see below) was excised with XhoI and NotI, and cloned in the XhoI/NotI sites of Chab. To construct the P[UAS-Antp.Q50K.C] (UAS:AntpK50) plasmid, the ANTP cDNA cloned in pBluescript KS+ (a gift of T. Kaufman) was mutagenized to replace the 50th codon of the homeodomain (encoding Gln) with the codon AAG (encoding Lys) and the insert was excised with EcoRI and cloned in the EcoRI site of the pUAST vector.

Site-directed mutagenesis

Site-directed mutagenesis was performed using the Muta-Gene Phagemid in vitro mutagenesis kit (BioRad, Richmond, CA). As a template we used the apME680 fragment cloned in the XhoI/BamHI sites of pBluescript KS+. The oligonucleotides used to mutagenize each site are

SITE 1: TTTGGGATATATCATAATAATCCCCGATTTACAC;
SITE 2: GCCCAAGAGGGAATATAATCCCCAACATGATAG;
SITE 3: GCTACCAAGGGATATAAAAATGACCC;
SITE 4: GTTAAGGGATTACAAGCCACCATAC;
SITE 5: GAAAGAAGGGATTATTTAATCCCATGGATAAG.

Protein and antisera production

To produce ANTP protein in bacteria, the ANTP cDNA cloned in pBluescript (see above) was excised with SmaI and HindIII and cloned in the PvuII/HindIII sites of pRSV-ANT (Invitrogen). The ANTP protein produced lacks the first 30 amino acids. The pRSET:ANTP plasmid was transformed in BL21 (DE3) LysS Escherichia coli cells. After 3 hours of induction with 1 mM IPTG, the protein was purified on a Ni-NTA agarose column (Qiagen) according to manufacturer’s instructions.

To produce anti-ANTP antibodies, 60 µg of purified ANTP were

![Fig. 3. ANTP protein accumulates in ap-expressing cells in the T2 and T3 mesoderm. Stage 11 apME680lacZ embryo stained with anti-ANTP (red) and anti-β-galactosidase (green) antibodies shown as single channels (A,B) or merged channels (C). The arrowheads point to cells that are positive for ANTP and anti-β-galactosidase in T2 and in T3. Note that in T1 there are cells positive for β-galactosidase and negative for ANTP (arrow).](image-url)
injected in rats every 2 weeks and blood was collected 10-14 days after the third to the fifth injection. ELISA tests were performed on each bleed and the sera with highest titer were used.

**DNase footprinting**
To obtain the footprinting probes, the apME680 fragment cloned in pBluescript (see above) was end-labeled by filling with Klenow the XhoI (sense strand) or XhoI (antisense strand) sites. DNase footprinting and chemical cleavage sequencing reactions were done as described in (Vachon et al., 1992). The products of the reactions were separated by electrophoresis on a 6% acrylamide gel containing 8 M urea.

**Fly crosses and embryos staining**
For each experiment, two to four independent transgenic lines were analyzed (identification symbols are available upon request). To analyze apME680lacZ expression in Antp mutants, stocks carrying the apME680lacZ transgene on the second chromosome and the mutant chromosome generated by standard crosses. Males of these stocks were crossed to P{en1}wg[en11] and P{GAL4-twi.G}*; P{GawB}how[24B] balancers were balanced with the TM6B, P{35UZ}DB1, Tb[1] balancer were generated by standard crosses. Mutual embryos were identified from the lack of the UbxlacZ expression pattern. To generate the embryos shown in Fig. 6, stocks containing both the apME680<sup>1-5BCD</sup>lacZ and the UAS:Antp<sup>K50</sup> or UAS:Antp<sup>K50</sup> transgenes balanced with CyO-P[en1]wg[en11] and TM6B, P{35UZ}DB1, Tb[1] balancers were generated by standard crosses. Males of these stocks were crossed to P[GAL4-twi.G]<sup>+</sup>; P[GawB]how[24B] (twiGal4; 24BGal4) females and the progeny embryos were analyzed.

In situ hybridization was performed as described (Tautz and Pfeifle, 1989) with minor modifications. Histochemical detection of β-galactosidase was carried out using biotinylated horse anti-mouse IgG (1:500 dilution) and avidin-horseradish peroxidase (Vectorstain Elite Kit, Vector labs). Mouse anti-β-galactosidase antibodies were used at a 1:1000 dilution for immunohistochemistry and at 1:500 for immunofluorescence. Rat anti-ANTP and rabbit anti-MEF2 (a gift of H. Nguyen) antiseras were used at a 1:1000 dilution. For immunofluorescence, the following secondary antibodies from Molecular Probes were used at a 1:200 dilution: Alexa Fluor 488 goat anti-mouse IgG, Alexa Fluor 546 goat anti-rat IgG and Alexa Fluor 546 goat anti-rabbit IgG.

**RESULTS**

**apME680 is a muscle-specific enhancer of the *ap* gene**
In the embryo, *ap* is first expressed at stage 10 in a patch of cells in each thoracic and abdominal hemisegment (Bourgouin et al., 1992; Cohen et al., 1992; Fig. 1A). These cells correspond to muscle progenitors and to cells associated with the PNS (Bourgouin et al., 1992; Cohen et al., 1992; see below). Very few cells express *ap* in the first thoracic segment. The size of the *ap*-positive patches in the second and third thoracic segments is bigger than in the first to eighth abdominal segments (Fig. 1A). At germ band retraction, *ap* is expressed in the brain, in the ventral nerve cord, in the CNS and in the somatic mesoderm (Fig. 1B).

To identify the mesoderm-specific enhancer of *ap* and to study the mechanisms that lead to differential expression of *ap* along the rostrocaudal axis, a series of restriction fragments covering the *ap* gene were cloned in the P[CaSpeR-hs43-AUG-betakal] reporter vector and transgenic animals were generated. A XhoI/Sau3A fragment of 680 bp, located in the second largest *ap* intron (see Fig. 4A), was capable of directing lacZ expression starting from stage 10 in clusters of cells very similar to those expressing *ap* at this stage (compare Fig. 1C with Fig. 1A). We call this fragment apME680 (for *ap*-muscle-enhancer-680) because it directs muscle-specific reporter gene expression (see below). At stage 13, β-galactosidase is detected in one continuous cluster in T2 and T3, while two smaller clusters, located at the dorsal and ventral limits of the thoracic clusters, are detected in segments A1-A7 (Fig. 1D). In segment A8, a unique smaller cluster is detected. As shown in Fig. 1E, these β-galactosidase-positive cells contribute to the formation of muscles LT1-4 in segments T2-A7 and to muscle LT1 in A8. These are a subset of the muscles originating from *ap*-expressing cells, as *ap* is expressed also in the progenitors of muscles VA2 and VA3 (Bourgouin et al., 1992). We note that thoracic muscles LT1-4 are slightly different than the same abdominal muscles (see below). In particular, muscle LT4 extends more dorsally and ventrally in the thorax than in the abdomen (see arrowheads in Fig. 1E).

**Antp positively regulates apME680**
The differential expression of *ap* in thoracic versus abdominal segments, starting from early germ band extension, suggests that *ap* is under Hox control. Antp is the only Hox gene expressed in the mesoderm of the posterior thorax (Kaufman et al., 1990). We investigated *ap* regulation in the cells giving rise to muscles LT1-4 by analyzing apME680lacZ expression in Antp loss of function mutants. At stages 10-11, lacZ expression in T2 and T3 is reduced (compare Fig. 2C with Fig. 2A), indicating that Antp positively regulates apME680 in T2 and T3. At stage 13, apME680lacZ expression is unaltered in the abdomen, while only a few β-galactosidase-positive cells are detected in T2 and T3 (compare Fig. 2D with Fig. 2B). These cells are located at the same position as the ventralmost abdominal clusters. In Antp mutants, thoracic muscles LT1-4 in T2 and T3 do not form (Roy et al., 1997 and Fig. 2E).

It has been reported that Antp is not expressed in the T2 mesoderm, suggesting that the role of Antp on the formation of muscles LT1-4 is non cell-autonomous (Roy et al., 1997). We produced polyclonal anti-ANTP antibodies in rats and used them to perform double immunofluorescence labeling on apME680lacZ embryos together with anti-β-galactosidase antibodies, to specifically label the LT1-4 muscle progenitors expressing *ap*. These anti-ANTP antibodies are specific for ANTP as they do not label Antp mutant embryos (data not shown). As shown in Fig. 3, co-localization of ANTP and β-galactosidase is observed in T3 as well as in T2. In the mesoderm as well as in the ectoderm, the subcellular localization of the protein appears to be nuclear but also cytoplasmic. Overall, ANTP protein is detected in more cells in the T3 mesoderm than in the T2 mesoderm and its level is higher in T3 than in T2. Nevertheless ANTP is clearly present in the T2 mesoderm, particularly at the germ band extension stage and, most importantly, in the apME680lacZ-expressing cells.

These results indicate that Antp is a positive regulator of *ap* expression in the somatic mesoderm and suggest that Antp has a cell-autonomous role in muscle development.

**ANTP regulates apME680 directly in vivo**
The cell-autonomous role of Antp on *ap* regulation in the somatic
Direct regulation of ap by ANTP

mesoderm was further investigated at the molecular level. To determine whether ANTP binds to apME680, the enhancer was sequenced and used as a probe in DNaseI footprinting experiments with ANTP protein produced in bacteria. In vitro, Hox proteins bind to sequences composed of a TAA T core followed by G, T or A nucleotides (Ekker et al., 1994). In the apME680 sequence, several TAA T cores and putative ANTP binding sites are present (Fig. 4B). As shown in Fig. 5, ANTP protects five regions of apME680 in DNaseI footprinting experiments. These regions contain a total of at least eight sequences closely matching or resembling ANTP consensus binding sites and we collectively name them sites 1-5.

In order to determine whether the binding of ANTP to apME680 is important for activating reporter gene expression in vivo, we used a strategy of altering the binding sites and making compensatory mutations in the protein to restore high affinity binding (Schier and Gehring, 1992). First, we altered the sequences of the ANTP-binding sites by replacing the eight consensus sites protected by ANTP in vitro with TAA TCCC sequences (Fig. 4B). These sequences have extremely low affinity for wild-type Hox proteins, but have high affinity for homeodomains containing Lys in position 50, such as that of BICOID (BCD) (Hanes and Brent, 1989; Treisman et al., 1989). This mutated enhancer (apME6801-5BCD) is not bound detectably by ANTP using in vitro DNaseI footprinting as an assay (data not shown) and loses the ability to activate lacZ expression in cells in which the wild-type enhancer is active (Fig. 6). The mutant enhancer is still functional, as novel expression is seen in the head and in the central part of the endoderm starting from stage 11 (Fig. 6A). Interestingly, endodermal expression was observed also with another direct Hox target enhancer containing mutant BCD sites (Capovilla et al., 1994; Sun et al., 1995). It is likely that, in these cells, transcription factors containing a homeodomain (possibly of BCD type) are responsible for lacZ expression directed by these mutated enhancers. At stage 12, expression is observed also in some ventral cells (Fig. 6C) and at stage 16 β-galactosidase has accumulated in ventral muscles, but is...
undetectable in muscles LT1-4 (Fig. 6E). These results indicate that sites 1-5, which are bound in vitro by ANTP, are required for enhancer function in the progenitors of muscles LT1-4, the cells in which wild-type apME680 is positively regulated by Antp.

ANTP may regulate apME680 directly by binding to sites 1-5. Alternatively, it may regulate a gene encoding a protein with a homeodomain of HOX-type, which in turn regulates apME680. To distinguish between these possibilities, we investigated whether mutant ANTPK50 protein (containing Lys in position 50 of the homeodomain), that can bind to the apME680-5BCD enhancer, is able to activate apME680-5BCD lacZ expression. We used the UAS/Gal4 system to produce wild-type ANTP and mutant ANTPK50 proteins in the mesoderm using the twiGal4; 24BGal4 driver line and tested their effect on apME680-5BCD lacZ expression. Overexpression of wild-type Antp does not have any effect on apME680-5BCD lacZ expression while ANTPK50 restores lacZ expression in clusters of lateral cells in each thoracic and abdominal hemisegment (Fig. 6). At stage 16, β-galactosidase is detected in muscles LT1-4 (Fig. 6F), indicating that ANTPK50 restores apME680-5BCD lacZ expression in the appropriate cells. Nevertheless, some of these muscles present abnormal shapes and orientation, suggesting that overexpression of ANTPK50 has a deleterious effect on muscle development. This is further confirmed by the observation that ANTPK50 overexpression in the mesoderm leads to larval lethality. We note that ANTPK50 is able to activate apME680-5BCD lacZ expression in T1 (where ANTP is normally absent), although in fewer cells than in T2 and T3, and that the size of the clusters in T2 and T3 is larger than in A1-A8. This last effect is probably due to the fact that T2 and T3 contain a higher number of LT muscle progenitors than the abdominal segments (see below).

Altogether, these results demonstrate that ANTP regulates apME680 directly and not through an intermediate homeodomain-containing transcription factor.

Role of ANTP in muscle segment identity
As shown above, ap and apME680lacZ are expressed in more cells in the thorax than in the abdomen, and ANTP directly regulates ap in the thorax. Next we addressed the question of the significance of the homeotic regulation of ap by ANTP.

The perdurance of β-galactosidase allows us to label thoracic and abdominal LT1-4 mature muscles originating from the cells expressing ap starting from the early germ band.
Direct regulation of ap by ANTP

Extended stage (see Fig. 1E). LT1-4 muscles present different characteristics in the thorax and in the abdomen. In the thorax, they contain more β-galactosidase, they are more tightly packed and, at least in the case of muscle LT4, extend more dorsally and ventrally (Figs 1E, 7). These differences may be a consequence of more myoblasts contributing to the thoracic muscles than to the corresponding abdominal muscles. To investigate this hypothesis, we performed double labeling experiments using anti-β-galactosidase to label muscles LT1-4 and anti-MEF2 antibodies, which label all muscle nuclei (Bour et al., 1995). Fig. 7 shows that in wild-type embryos, LT1-4 thoracic muscles do contain more MEF2-positive nuclei than the same abdominal muscles. The number of nuclei was compared in the T2, T3 and A1 hemisegments of ten independent embryos. This quantitative analysis shows that, on average, T3 muscles contain a total of 28 nuclei, while A1 muscles contain 19 nuclei (Fig. 7D). This difference is statistically significant \( P<0.001 \). No significant differences were observed between the number of nuclei in T2 and T3 (data not shown). Consistently, highly packed nuclei are present in the medial portion of T2 and T3 muscles, but are absent in the same region of abdominal muscles (compare Figs 7B,C).

**DISCUSSION**

*Antp* has a cell-autonomous function in muscle development

We have identified a muscle-specific enhancer of *ap* (apME680) that is positively regulated by *Antp* in the second and third thoracic segments. This enhancer drives expression in the progenitor cells of muscles LT1-4. In the T2 and T3 segments, *Antp* is required for the formation of these muscles, and the LT1-4 muscle progenitors accumulate ANTP protein. By altering the binding specificity of the ANTP protein and

**Fig. 6.** ANTP directly regulates apME680. All embryos carry the *apME680-5B+8B lacZ* and the *twlGal4:24BGal4* transgenes and have been stained in parallel with anti-β-galactosidase antibodies. (A,C,E) Embryos overexpressing wild-type *Antp lacZ* is weakly expressed in the head, in the endoderm (arrowhead) and in ventral muscles, but is not detected in muscles LT1-4. The same pattern is observed in *apME680-5B+8B lacZ* embryos not carrying the *UAS:Antp* transgene. (B,D,F) Embryos overexpressing mutant *AntpK50 lacZ* is activated in segments T1-A8 in mesodermal clusters that resemble the *ap*-expressing clusters. At stage 16, β-galactosidase is detected in muscles LT1-4 (F). The arrows point to the first abdominal segment.

**Fig. 7.** LT1-4 muscles in the thorax contain more nuclei than the same muscles in the abdomen. (A) Confocal image of a wild-type *apME680 lacZ* embryo labeled by immunofluorescence with anti-β-galactosidase (green) and anti-MEF2 (red) antibodies. A higher number of nuclei (yellow/orange) is present in the thoracic LT1-4 muscles than in the abdominal ones. (B,C) High magnification scans of segments T2 (B) and A1 (C) of embryo shown in A. Each LT muscle is indicated by its corresponding number. The arrowhead in B points to an individual nucleus. Nuclei are present in the medial areas of the T3 muscles, but absent from the corresponding areas of the A1 muscles. (D) Grey bars indicate the average number of nuclei in the LT1-4 muscles of T3 and A1. Error bars represent standard deviations \( n=10 \).
making compensatory mutations in the ANTP-binding sites in apME680, we show that ANTP directly regulates ap expression in the mesoderm.

Our results contrast with a previous report claiming that Antp acts non cell-autonomously in T2 muscle development (Roy et al., 1997). These authors based their conclusion on two main observations. First, they were able to detect ANTP protein in T3 but not in T2 mesoderm. However, we have clearly shown that ANTP is detectable in the T2 mesoderm using polyclonal anti-ANTP antibodies. One possibility to explain these contradictory observations is that the monoclonal antibody used by Roy et al. (1997) might not recognize an ANTP isoform present in the T2 mesoderm. Alternatively, their detection method might not be sensitive enough to detect the low levels of ANTP in the T2 mesoderm. The second observation by Roy and colleagues is that expression of Antp in the ectoderm from the e22cGal4 driver restores the LT1-4 muscles that are lacking in Antp mutants. This led to the conclusion that Antp expression in the ectoderm patterns the underlying mesoderm non autonomously. However, we must consider the possibility that the e22cGal4 line expresses Gal4 in the mesoderm (Lawrence et al., 1995).

We would like to point out that, in the mesoderm, ANTP accumulates at relatively lower levels and in a smaller number of cells in T2 than in T3. This difference in Antp expression may explain the different T2 and T3 muscle patterns: the presence of three ventral muscles in T3 that are absent in T2 (muscles VA1, VO3 and VL4). This hypothesis is consistent with the observation that homogeneous high levels of ANTP lead to a T3-type of muscle pattern in T2 following Antp overexpression (Roy et al., 1997).

Role of Hox genes in determining muscle identity

Our results are in agreement with the phenotypes observed after mesodermal overexpression of other Hox genes. Ectopic expression of Ubx or abd-A throughout the mesoderm causes muscle transformations characterized by the appearance of abdominal-like muscles in the thorax and of posterior abdominal muscles in A1 (Michelson, 1994). This gain of muscles was often associated with a loss of segment-specific muscles, indicating that the observed effects are bona fide transformations. These muscle transformations occurred in the absence of corresponding effects on the overlying cuticle. Ectopic abd-A expression in the thoracic mesoderm eliminates twist-positive adult muscle precursors from the thorax (Greig and Akam, 1993). In addition, an abdominal-type pattern of nau expression appears in the thorax after Ubx or abd-A mesodermal overexpression (Michelson, 1994). These results suggest that Hox genes act at early steps of muscle formation to establish the identity of muscle progenitors in each segment. According to the idea that Hox genes function cell-autonomously in the mesoderm, this is likely to be accomplished through the direct regulation of genes involved in the formation of specific muscles. In the case of muscles LT1-4, in addition to ap, one likely candidate Hox target is the muscle-identity gene msh which is also required for the formation of these muscles (Nose et al., 1998), although its regulation by Hox proteins has not been investigated yet.

In this report, we show that segments differ not only in the presence or absence of specific muscles, but also in the size and shape of some homologous muscles. The apME680lacZ construct allowed us to determine that ap is expressed in more LT1-4 progenitor cells in the thorax than in the abdomen. In light of the current view on muscle development, stating that each muscle derives from a single founder cell (reviewed by Bate et al., 1999), it is somewhat surprising that T2 and T3 do not contain a higher number of LT muscles than the abdominal segments. Nevertheless, our findings suggest that thoracic muscles LT1-4 do derive from more myoblasts than the abdominal ones, as we observed that mature LT1-4 muscles contain more nuclei in the thorax than in the abdomen.

What controls this segmental difference? ANTP is normally present in the progenitors of LT1-4 muscles in T2 and T3 (see Fig. 3) and Antp mutants lack these muscles in the thorax (Roy et al., 1997). Thus, one role of Antp may be to allocate a specific number of mesodermal cells to the formation of muscles LT1-4 in T2 and T3. Our observations indicate that Antp activates ap in more cells in the thorax than other Hox genes do in the abdomen. Ubx is expressed at its highest levels in the A1 and A2 mesoderm (White and Wilcox, 1985; data not shown). In Ubx mutants, Antp is derepressed in A1 and A2 and thoracic-like LT1-4 muscles form in place of abdominal-like LT1-4 muscles in A1 and A2 (Hooper, 1986). In Antp, Ubx double mutants, LT1-4 muscles in A1 and A2 do not form (data not shown), indicating that Ubx, like Antp, is required for their formation. UBx is likely to positively regulate ap directly, as it protects in vitro the same sequences on apME680 as ANTP (data not shown) and as the mutation of these sequences abolishes apME680lacZ expression both in the thorax and in the anterior abdomen. Similarly, it will be interesting to investigate if abd-A and possibly Abd-B positively regulate apME680 in the posterior abdomen.

LT1-4 muscles specified by Antp in T2 and T3 contain more nuclei than those specified by Ubx in A1 and A2. Thus, one mechanism through which Antp and Ubx may determine the segment identity of these muscles is by allocating different numbers of progenitors to the formation of the LT1-4 muscles in different segments. This could be accomplished by regulating the target genes in a different manner and/or by controlling the expression of different targets. Consequently, downstream genes required for muscle specification would be expressed in a slightly different pattern in the thorax than in the abdomen and a different muscle pattern would ensue.

ap is a direct target of ANTP in the somatic mesoderm

Hox genes encode homeodomain transcription factors with segment-specific effects on body patterning in all the different tissues. They are likely to regulate the segment and tissue-specific expression of genes responsible for the morphogenesis of each tissue. Although the pattern of expression of a high number of genes is affected in Hox mutants (as expected from the wide phenotypic consequences of Hox mutations), it is not known what percentage of these genes is regulated by Hox proteins in a direct manner. It has been proposed that homeodomain-containing proteins regulate directly a large number of genes (Liang and Biggin, 1998). Despite this observation, only a limited number of cis-regulatory sequences are known to be direct in vivo targets of homeodomain proteins (see Introduction).
The strategy of introducing compensatory mutations in the binding sites and in the protein in order to restore regulation (Schier and Gehring, 1992) provides the most compelling evidence for the direct regulation of a cis-regulatory element by a transcription factor. Similar experiments have shown that the enhancer that controls dpp expression in the second midgut constriction (dp674) is directly regulated by UBx and by ABD-A (Capovilla et al., 1994; Capovilla and Botas, 1998). This enhancer has been instrumental in demonstrating that one mechanism through which Drosophila Hox proteins acquire specificity is their interaction with a co-factor encoded by extradenticle (exd) (Chan et al., 1994; reviewed in Mann and Chan, 1996). It is possible that exd plays the role of an ANTP co-factor in the regulation of apME680, as consensus binding sites for EXD (Mann and Chan, 1996) are present in the apME680 sequence and since apME680lacZ expression is abolished in embryos maternally and zygotically mutant for exd (data not shown). Nevertheless, to determine unambiguously if exd is involved in the regulation of apME680 it will be necessary to clearly identify and mutagenize the EXD-binding sites on apME680, because exd might be required initially for the formation of the ap-expressing muscle precursors.

Recently, it has been shown that onycophora UBx (OUBX), an evolutionary distant UBx ortholog, positively regulates dp674 in the same manner as Drosophila UBx. Conversely, OUBX is unable to negatively regulate the Dll enhancer, another direct target of Drosophila UBx (Grenier and Carroll, 2000). The residues responsible for these differences map outside the most conserved homeodomain region, suggesting that a new functional domain, possibly allowing interactions with unknown co-factors, has evolved in the Drosophila Ubx protein. Altogether, these results highlight the importance of identifying the enhancers of the direct Hox target genes to unravel the target and transcriptional specificities of Hox proteins and their functional evolution.

Another poorly understood aspect of Hox transcription factors is how these proteins, which are present in all the tissues, regulate their target genes in a tissue-specific manner. Tissue-specific enhancers of direct Hox targets are essential tools for addressing how Hox proteins acquire tissue-specific regulatory functions. apME680 is a somatic mesoderm-specific enhancer of ap, a gene that is structurally and functionally conserved from Drosophila to humans (Rincon-Limas et al., 1999). Its direct regulation by ANTP provides an entry point towards understanding the transcriptional functions of a prototypical Hox protein.

We thank Patrick Lawinger for initial work on the apME680 enhancer, Brenda Lilly for the construction of the pRESET-ANTP plasmid and Maria Jesus Garcia Garcia for valuable help and stimulating discussions. We acknowledge the help and advice of Jun Wu and Christine Brunel for confocal imaging. We thank Steve Cohen, Thom Kaufman, Hahn Nguyen and the Bloomington Stock Center for DNAs, antibodies and fly stocks. M. C. and Z. K. would like to thank Jules Hoffmann for encouragement and support. M. C. is grateful to Angela Giangrande, Filippo Rijli, Marie-Laure Ruhf and Patricia Simpson for their comments on the manuscript. This work was supported by grant n.GM55681 from the National Institute of Health awarded to J. B., and by grants to M. C. from the Fondation pour la Recherche Medicale (Aide à l’Implantation de Nouvelles Equipes) and the Ligue Contre le Cancer.

REFERENCES


