Roux's Arch Dev Biol (1996) 205:333–343

ORIGINAL ARTICLE

Marie Thérèse Besson · Pascal Chareyre
Gilbert Deléage · Jocelyne Demont
Brigitte Quennedey · Annie Courrent
Jacques Fourche · Georges Bosquet

Isolation and characterization of *Urbain*, a 20-hydroxyecdysone-inducible gene expressed during morphogenesis of *Bombyx mori* wing imaginal discs

Received: 31 March 1995 / Accepted in revised form: 21 November 1995

Abstract In insects, wing imaginal discs respond to the steroid hormone 20-hydroxyecdysone by initiating morphogenesis leading to the formation of the adult flight appendages. In this work we analyse the expression of a *Bombyx* gene, referred to as *Urbain,* whose cDNA had been previously isolated from wing discs (Chareyre et al. 1993). Accumulation of the 1.8 kb transcript occurs concomitantly with the increase of 20-hydroxyecdysone titer at every stage examined during post-embryonic development. In vitro, its accumulation is delayed 6–9 h after exposure to 20-hydroxyecdysone. Studies in the presence of cycloheximide have established that *Urbain* is a secondary response gene. The sequence of the mRNA contains a putative open reading frame of 1656 nucleotides encoding a secreted hydrophilic protein, and we suggest that the *Urbain* gene product plays a role in cellular mechanisms involved in morphogenesis of the epithelium.

Key words Morphogenesis · Wing discs · 20-hydroxyecdysone

Introduction

Embryonic pattern formation in *Drosophila* has been subjected to extensive studies which have led to the identification of many key genes involved in development (for review see Ingham 1988; Nüsslein-Volhard 1991). By comparison with embryogenesis, relatively little is yet known about genes and processes that regulate the development of the adult, such as the formation of the adult appendages (e.g. legs, wings, antennae). The wing imaginal discs, which form the flight appendages and a part of the thoracic adult structures, derive from restricted areas within the embryonic body. Discs are subdivided into anterior/posterior and dorsal/ventral compartments during embryogenesis and early larval stages; the proximo-distal axis is also established during this time. The elaboration of this patterning necessitates the expression of a hierarchy of developmental genes, many of which are also required for embryonic development (Nijhout 1994; Williams et al. 1993, 1994).

Throughout larval life, wing discs increase in size by cell division and finally, during a brief time prior to pupation, they undergo extensive morphogenesis. During this period, coordinated cell movements involving cell shape changes, cell adhesion and (or) cell-cell signalling modifications mould the adult wing (Fristrom et al. 1993; Blair et al. 1994; Nardi 1994).

In Diptera as well as in Lepidoptera, disc morphogenesis is initiated by the insect steroid hormone, 20-hydroxyecdysone (Natzle et al. 1986); previous studies (Chareyre et al. 1993, *Bombyx;* Natzle 1993, *Drosophila*) showed that imaginal discs respond by activation or repression of several genes. However, the molecular response in discs has not been as extensively studied as it has been in *Drosophila* salivary glands (Ashburner et al. 1974). Since the wing discs of *Bombyx mori* are easily obtainable as pure and precisely staged discs, they offer a good model in which the control of gene expression during morphogenesis can be studied. The morphogenetic processes of wing discs occur over a period of 4 days, providing ample time to study their development in detail.

In this report, we present results on the isolation of a gene expressed in wing discs and the characterization of its encoded product. The *Urbain* cDNA clones were previously isolated by differential screening to characterize genes encoding 20-hydroxyecdysone-induced proteins (Chareyre et al. 1993). *Urbain* is detected in the prepupal

M.T. Besson (✉) · P. Chareyre · J. Demont · J. Fourche
G. Bosquet
Centre de Génétique Moléculaire et Cellulaire-UMR 106,
Université Lyon-1, 43 bd du 11 novembre 1918,
F-69622 Villeurbanne Cedex, France

G. Deléage
Institut de Biologie et de Chimie des Protéines,
7 passage du Vercors, F-69367 Lyon Cedex 07, France

B. Quennedey · A. Courrent
Laboratoire de Zoologie, UA 674, Université de Bourgogne,
6 bd Gabriel, F-21000 Dijon, France

wing discs as a 1.8 kb transcript. We show that *Urbain* expression in wing discs coincides with the peaks of the hormone that occur at the end of each larval stage and in the pupa. In in vitro experiments, accumulation of the transcript is observed 6–9 h after hormonal exposure and is blocked by a protein synthesis inhibitor. This suggests that *Urbain* behaves as a secondary response gene, according to the model proposed for *Drosophila* (Ashburner et al. 1974). Other tissues, such as epidermis, midgut and fat body, express *Urbain* at different levels. The complete nucleotide sequence of the transcription unit reveals that the *Urbain* gene product is secreted and shares some structural features with the *Drosophila* IMP-E2 protein involved in disc morphogenesis (Paine-Saunders et al. 1990).

Materials and methods

Isolation of wing imaginal discs

B. mori larvae of the European strain 200–300, provided by the „Unité Natiolnale Séricicole" (France), were reared at 22° C. The developmental ages of the larvae and pupae were determined by the time elapsed from the last ecdysis (see legend of Fig. 1). Precise staging at the moulting period was obtained using morphological criteria (Kiguchi and Agui 1981; Calvez 1981). The four wing imaginal discs (two forewings, two hindwings) were dissected from precisely staged larvae and pupae, and immediately stored at −80° C until use.

In vitro protocol

The discs from day 6 larvae in the 5th instar were sterilely dissected and rinsed with phosphate-buffered saline (PBS) to which 6 µg/ml phenylthiourea (Merck) had been added to avoid tissue damage by endogenous phenoloxidases. To prevent anaerobic stress, the discs were deposited onto sterile steel grids placed into a tissue culture dish (Falcon) with 2 ml of Grace's medium (GIBCO) flush with the grids. The medium was supplemented with 10% fetal calf serum and 10% *Bombyx* haemolymph. The discs were maintained at 25° C for 24 h maximum. Ethanolic dilutions of 20-hydroxyecdysone were added to the Grace's medium and used at indicated final concentrations. The control discs were treated with equivalent volumes of ethanol solvent alone.

For experiments involving inhibition of protein synthesis, cycloheximide (Sigma) was added at a final concentration of 50 µg/ml 30 min before hormone addition.

Isolation of genomic clones

Genomic clones were isolated from a *Bombyx mori* library (Eickbush and Kafatos 1982), which had been constructed after genomic DNA from strain 703 silkworms had been partially digested with EcoR1 and cloned into the phage lambda Charon 4. The library was screened using restriction fragments of *Urbain* cDNA as radioactive probes. After plaque transfer to Nylon membrane (Hybond N+, Amersham) according to standard procedures (Sambrook et al. 1989), hybridizations were performed at 42° C in 50% formamide, 5× sodium-sodium citrate buffer (SSC), 5× Denhardt's solution, 0.5% sodium dodecyl sulphate (SDS) and 100 µg/ml denatured herring sperm DNA. As a final wash, filters were placed in 0.2× SSC-0.1% SDS at 65° C. Two clones were selected and purified. Genomic DNA subclones were inserted into the EcoR1 restriction site of the Bluescript SK+ vector (Stratagene).

Isolation of RNA

Frozen wing discs were ground in lithium chloride-urea-heparin, and RNAs were isolated using a phenol-chloroform method (Clemens 1984). RNAs were ethanol precipitated. The final pellet was lyophilized, resuspended in RNAase-free water and the RNA concentration was estimated by spectrophotometric measurements before being used.

Preparation of probes

DNA was prepared for labelling by appropriate restriction enzyme digestion. Specific fragments were isolated from 0.75% low melting agarose gels (FMC) in TRIS/acetate/ethylenediamine tetraacetic acid (EDTA) buffer (Bioprobe). The purified 1.6-kb-long cDNA was used as a probe for transcript detection and was labelled by Klenow extension of random-primed DNA with ^{32}P-dCTP (Feinberg and Vogelstein 1984).

RNA analysis

Five micrograms of total RNA were fractionated on denaturating formaldehyde-1% agarose gels and transferred onto a nylon membrane (Hybond N+, Amersham) using a Vacugene (Pharmacia) apparatus. RNA markers (BRL) were stained with methylene blue. RNA blots were hybridized at 65° C in 5× SSC, 5× Denhardt's solution, 0.5% SDS, 100 µg/ml of heat-denatured herring sperm DNA and washed under stringent conditions at 65° C with 0.1× SSC-0.1% SDS final washing. Autoradiography was performed at −80° C using X-Omat AR 5 film (Kodak) and intensifying screens.

DNA sequencing

The restriction fragments from cDNA and genomic DNA were subcloned into the Bluescript SK+ vector (Stratagene). Sequencing was performed by the dideoxy method of Sanger et al. (1977) using Sequenase Version 2.0 (U.S.B.) and T7 polymerase kits (Pharmacia). Oligonucleotides were also used to determine internal sequences. The cDNA and genomic DNA were sequenced on both strands, except for the coding region of the genomic DNA which was sequenced on one strand only. Computer-assisted DNA and protein sequence analyses as well as searches in sequence databases were performed using the Antheprot package developed by Geourjon and Deléage (1993). The secondary structure of the protein was predicted using the self-optimized prediction method (Geourjon and Deléage 1994).

DNA analysis

The 5′-end of the *Urbain* cDNA was cloned after Polymerase Chain Reaction (PCR) amplification of cDNA synthesized from total RNA of day 11 discs. One microgram of total RNA was incubated with Superscript II reverse transcriptase (200 units; GIBCO) at 37° C and 20 pmoles of a 19-mer primer. The primer, 5′-CGATTGTTTCAAAAGAAGC-3′, was complementary to the sequence lying between positions 652 to 670. The first strand cDNA was purified, concentrated and tailed according to the 5′ RACE procedure of Frohman et al. (1988) using the GIBCO-BRL kit. The tailing reaction was performed by adding an oligo-dC sequence to the 3′ end of the cDNA in a PCR-compatible buffer.

Amplification was carried out with one-third of the volume of the dC-tailed cDNA solution using the two primers described below. The 3′ primer, 5′-TTCTTAAATCGACTAGAGGCC-3′, was complementary to nucleotides 317–337 of the sequence shown in Fig. 4 and the 5′ primer is a 29-mer oligonucleotide including 3 endonuclease recognition sites (BamH1, Sal I, Ava I) and ending with 15 dGs. The PCR reaction was carried out on a Perkin-Elmer

apparatus using 2.5 units of Taq DNA polymerase (Boehringer-Mannheim) added to a 50 µl final volume and run to 30 cycles of 94° C for 1 min, 53° C for 30 s, 72° C for 1 min and finished at 72° C for 3 min.

The products were digested, run on a 2% NuSieve gel (FMC) in TRIS/borate/EDTA buffer (Bioprobe), cut out, purified with GeneClean (Bio 101) and subcloned into the pBSK+ vector between the sites Sal I – Cla I prior to sequencing.

In situ hybridization

To improve adherence of tissue to slides, Superfrost slides (Poly-Labo, France) were coated according to the NaOH/poly(L)lysine method described by Wilcox (1993). Cryosections (14 µm) were prepared from tissue after 4% paraformaldehyde-1× PBS fixation at 4° C for 24 h and after embedding tissue in OCT compound (Miles). The Cla1-HindIII cDNA fragment (468 bp) was labelled with ^{32}P-dCTP by Klenow extension of random-primed DNA and used after purification for in situ hybridization. The hybridization procedure was carried out according to Charles et al. (1992) at 37° C.

Results

Temporal expression of *Urbain* mRNAs

To characterize transcripts involved in wing morphogenetic processes, a *B. mori* cDNA library was cloned into lambda-gt10 phage using mRNAs expressed in wing discs after four injections of 20-hydroxyecdysone, before the endogenous rise of hormone (Chareyre et al. 1993). A differential screening using cDNA probes raised against mRNAs prepared from wing discs before or after the onset of differentiation led us to isolate 28 clones, and among them, the Z17 clone. Two related clones, Z721 and Z1021 have been isolated by screening the *Bombyx* cDNA library with Z17 cDNA as a probe. The largest clone, Z1021, is presented below and is hereafter named *Urbain*.

In Northern blot experiments (Fig. 1), the *Urbain* transcript is observed at approximately 1.8 kb. A faint hybridization signal appears at 4 kb. This might represent either a cross-hybridization to a related transcript or to the unprocessed primary transcript. We have followed the expression of the *Urbain* mRNA in the wing discs during the last two larval stages and in the pupa (Fig. 1).

During the second half of the fourth larval stadium, the *Urbain* transcript level increases from day 3.5 to day 4, then its accumulation stops at day 5 just before ecdysis. During the first 8 days of the fifth larval stadium, the *Urbain* transcript level remains undetectable even if the Northern blot films are overexposed. Twenty-four hours after the beginning of cocoon spinning (day 8), the *Urbain* mRNAs start to accumulate. The intensity of the hybridization signal increases markedly and reaches a maximum at day 10.5–11. At day 11.5, when cocoon spinning is finished, the level of *Urbain* transcript decreases sharply so that *Urbain* expression is not detectable at larval-pupal ecdysis (day 12.5). In the pupa, a slight expression of the transcripts occurs at day 1 and 2; they rapidly accumulate on day 4 and continue to increase until day 5. But no hybridization signal was detected during day 3 or day 6 which bracket the accumulation of the *Urbain* transcript on days 4 and 5, despite a longer exposure of films from different Northern blots.

Accumulation of the *Urbain* transcript has also been followed with Northern blots in earlier post-embryonic stages (not shown). Because the wing discs are too small to be dissected, entire animals during the first two larval stages and thoracic parts during the third and fourth larval stadia were used to perform RNA analysis with Northern blots. A burst of *Urbain* expression occurs during the second half of each larval stage and is associated with a 20-hydroxyecdysone peak (Coulon 1977) in every case.

By comparing these results with variations in the titers of 20-hydroxyecdysone in *Bombyx* (Calvez et al.

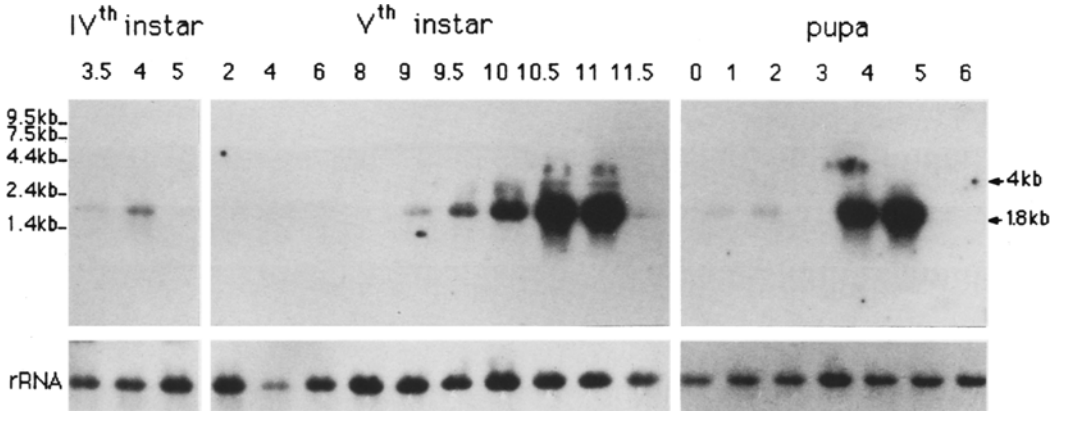

Fig. 1 Developmental profile of *Urbain*. Five micrograms of total wing disc RNA isolated at various times (in days) during the fourth and fifth larval stadia and from pupae are analysed by Northern blots and hybridized with a 1.6 kb cDNA probe. At 22° C, in the fourth larval stadium the feeding period is 3 days long; in the fifth larval stadium the feeding period is 7.5–8 days long, after which spinning begins. Larvae pupate 5 days thereafter and the pupal-adult apolysis starts 4 days after the larval-pupal ecdysis. The *top panel* shows hybridization with the cDNA probe. The RNA size markers are *on the left*. The *bottom panel* (rRNA) shows hybridization with a ribosomal probe, used as a control for loading of RNA

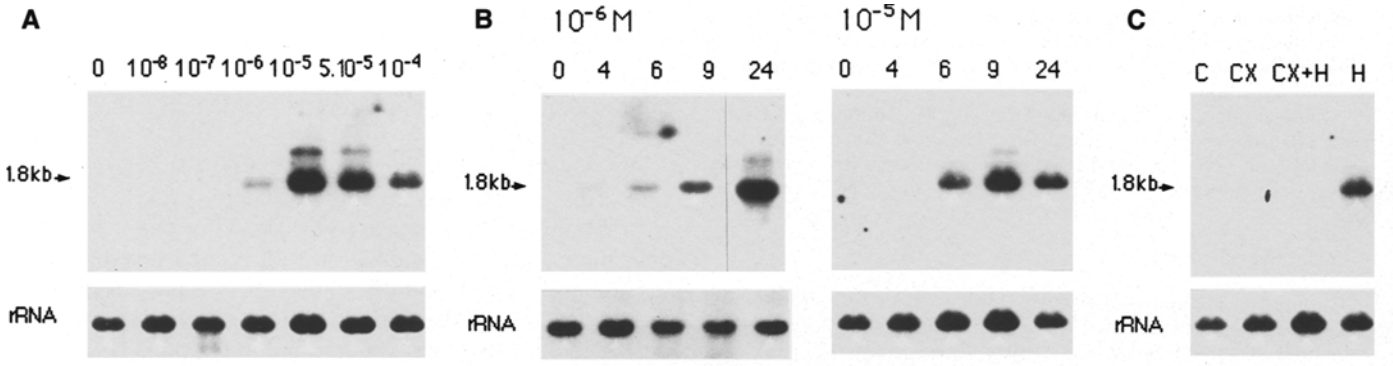

Fig. 2 **A** Dose-response profile for *Urbain* RNA induction by 20-hydroxyecdysone in vitro. Day 6 wing discs from fifth instar larvae are incubated in different molar concentrations of 20-hydroxyecdysone for 9 h, then total RNA (5 µg per lane) is Northern blotted and hybridized with a 1.6 kb cDNA probe. Ribosomal RNA (*rRNA*) is used as a quantitative control. **B** Time-course of *Urbain* expression in cultured wing discs treated with 10^{-6} M or 10^{-5} M 20-hydroxyecdysone. Time is in hours. Total RNA (5 µg per lane) is analysed by Northern blotting and hybridized with a 1.6 kb cDNA probe. Ribosomal RNA (*rRNA*) is used as a quantitative control. **C** Effects of 20-hydroxyecdysone and cycloheximide on *Urbain* expression. Northern blot analysis of total RNA (5 µg per lane) isolated from wing discs of day 6 fifth instar larvae that have been incubated for 9 h in four different culture conditions: in medium alone (control: *C*), in the presence of cycloheximide (50 µm/ml; *CX*), in cycloheximide and 10^{-5} M 20-hydroxyecdysone (*CX+H*) or in 10^{-5} M 20-hydroxyecdysone (*H*). The 1.6 kb cDNA is used as a radioactive probe and ribosomal RNA (*rRNA*) is used as a quantitative control

1976; Coulon 1977), we deduce that *Urbain* expression is associated with pulses of hormone and accumulates when 20-hydroxyecdysone reaches its maximal level during each stage of *B. mori* that was examined. The disappearance of the transcript is always dramatic, since there is no hybridization signal 12–24 h after the peak of mRNA (Fig. 1). This decrease is concomitant with the drop in hormonal titer. The *Urbain* transcript is not detectable during intermoult periods when no 20-hydroxyecdysone is present in the haemolymph. However, the correlation between *Urbain* mRNA accumulation and hormonal titer seems more complex in pupae: there is a high, broad peak of 20-hydroxyecdysone which lasts from day 1 to day 7 and is maximum at days 4–5. This single hormonal peak coincides with two periods of *Urbain* transcript accumulation (days 1–2, days 4–5), separated by a gap at day 3 (Fig. 1), although the hormone titer (6×10^{-6} M at day 3) is still increasing between the titers at day 2 (2×10^{-6} M) and at day 4 (1.2×10^{-5} M).

In vitro accumulation of the *Urbain* mRNAs

Since the above results show that *Urbain* is present during the peaks of 20-hydroxyecdysone, we asked whether *Urbain* accumulation is induced by this hormone. A Northern blot was prepared from wing discs cultured in vitro for 9 h with concentrations of 20-hydroxyecdysone from 10^{-8} M to 10^{-4} M (Fig. 2A). *Urbain* mRNAs show a graded response to 20-hydroxyecdysone over a 100-fold range of concentrations: *Urbain* is very weakly detectable at 10^{-7} M and peaks at 10^{-5} M. When hormonal levels increase to 10^{-4} M, *Urbain* expression decreases.

Cultured wing discs were incubated for increasing periods with10^{-6} M or 10^{-5} M 20-hydroxyecdysone in two independent sets of experiments. Figure 2B shows that *Urbain* mRNAs are not detected in fifth instar larvae at the time of disc explantation (day 6), nor during the first hours of hormonal exposure. In the presence of 10^{-6} M 20-hydroxyecdysone, a very faint hybridization signal appears at 6 h, is readily detectable at 9 h and increases during the 24 h of incubation. With a 10^{-5} M hormone concentration, *Urbain* mRNA accumulation is clearly present at 6 h, peaks at 9 h then decreases, suggesting a reduction in the rate of transcript production or stability.

These patterns show that *Urbain* expression is dependent on hormonal concentration but is delayed with regard to the addition of 20-hydroxyecdysone. This is consistent with the results presented in Fig. 2C: a Northern blot was prepared with RNAs extracted from wing discs incubated in culture medium alone (C), in medium supplemented with 20-hydroxyecdysone (H), with 20-hydroxyecdysone and cycloheximide (CX+H) or with cycloheximide (CX). The transcript accumulated only in the presence of the hormone. No mRNA can be detected when the wing discs are cultured with cycloheximide or 20-hydroxyecdysone and cycloheximide. This indicates that, during the fifth larval stadium, the synthesis of some proteins is required for the accumulation of the *Urbain* mRNA and that this protein synthesis is controlled by 20-hydroxyecdysone.

Tissue distribution of the *Urbain* mRNAs

The tissue distribution of the *Urbain* transcript is examined in fifth instar larvae when *Urbain* mRNA level in the wing discs is maximal (day 11), as shown by Northern blot hybridizations (Fig. 1), and also when there is no *Urbain* expression (day 6). In situ hybridizations are performed on thoracic sections using a *Urbain* cDNA-specific probe. Figure 3A and B shows that *Urbain* mRNAs accumulate when 20-hydroxyecdysone titer is high, but they are barely detected when the hormone is at its basal level. The most intense hybridization signal is

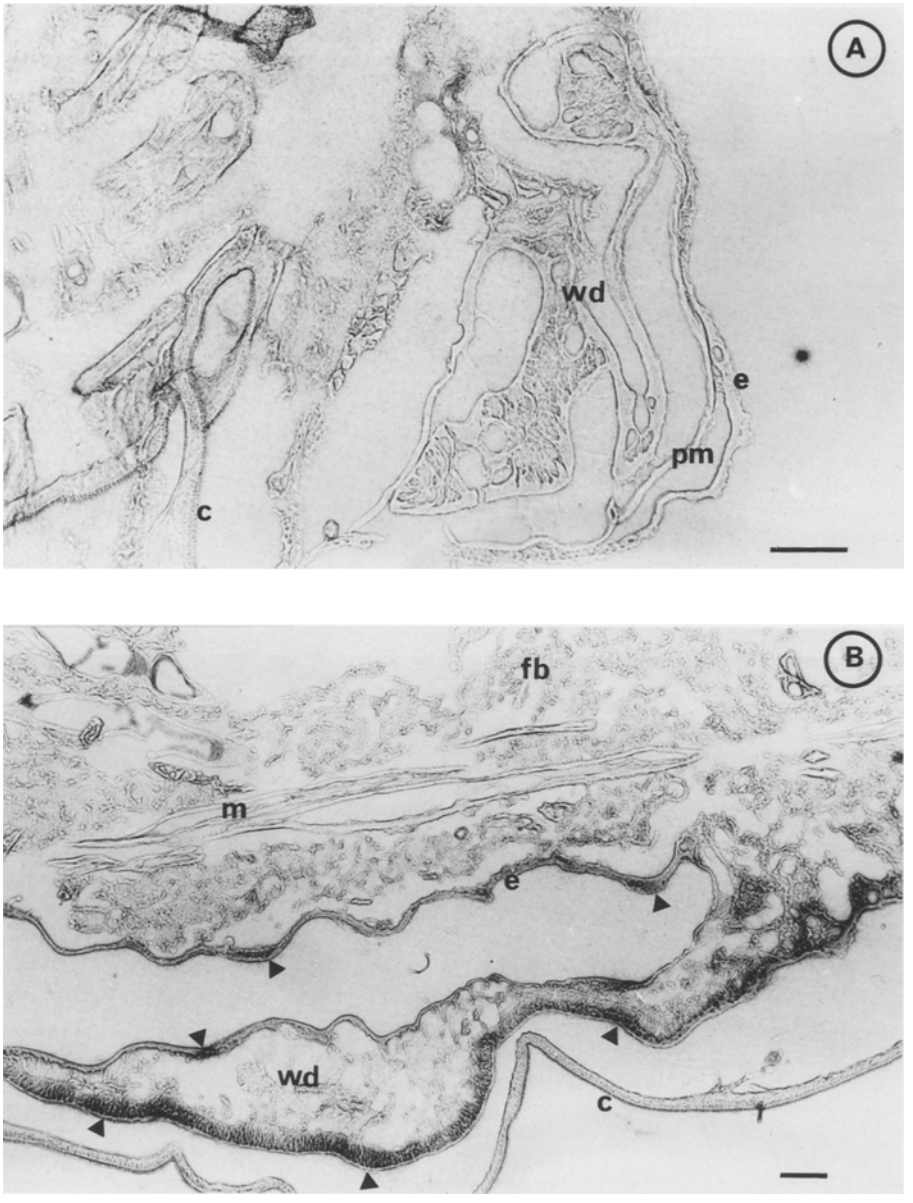

Fig. 3A, B Localization of *Urbain* transcripts by in situ hybridization in day 6 (**A**) and day 11 (**B**) fifth instar larvae. Photographs were taken of frozen cross-sections hybridized with a ^{32}P-c DNA probe and exposed to Ilford K5 autoradiographic emulsion film. **A** The wing disc (*wd*) is enveloped by a peripodial membrane (*pm*). The cuticle (*c*) which normally encloses the animal, was broken during preparation of the section (*e* epidermis, *scale bar* 50 μm). **B** The wing disc is everted. Labelling *(arrowheads)* is concentrated in the wing disc (*wd*) and epidermis (*e*). No signal is found on muscles (*m*). Fat body (*fb*) shows a very slight labelling. A cuticle (*c*) encloses the animal (*Scale bar* 100 μm)

observed in wing discs and epidermis. The intensity of transcript expression is not identical among different tissues; while not all tissues have been examined, it is clear that *Urbain* expression is not restricted to wing discs but occurs also in epidermis, another ectodermal tissue that is a target of 20-hydroxyecdysone. Significantly lower labelling can be seen in other 20-hydroxyecdysone target tissues such as midgut (not shown in Fig. 3) or fat body. No signal is found in the silk glands (not shown in Fig. 3) and muscles.

Sequence analysis of *Urbain*

Figure 4 shows the complete nucleotide sequence for the *Urbain* cDNA which agrees with the size of the transcript determined by Northern blot analysis. This sequence has been determined using several overlapping subclones obtained after restriction analysis of Z17 and Z1021 cDNA clones. All cDNA sequences are determined on both strands, using oligonucleotides as internal primers. In addition, since the cDNAs are not full-length, the 5′ end of the *Urbain* transcript was obtained by RT-PCR. After cloning, PCR fragments of three independent PCR amplification reactions were sequenced and these sequences were found to be identical.

Two genomic clones were isolated from a *B. mori* genomic library (Eickbush and Kafatos 1982), which together include the entire coding sequence of *Urbain* (Fig. 5). The restriction fragments of 3.8, 5 and 2.5 kb were subcloned in the plasmid vector Bluescript SK+ (Stratagene) at the EcoR1 site. Detailed restriction analysis and partial sequencing have revealed that the two genomic clones share a common 3.8 kb region. The sequencing of one strand in the genomic coding region has

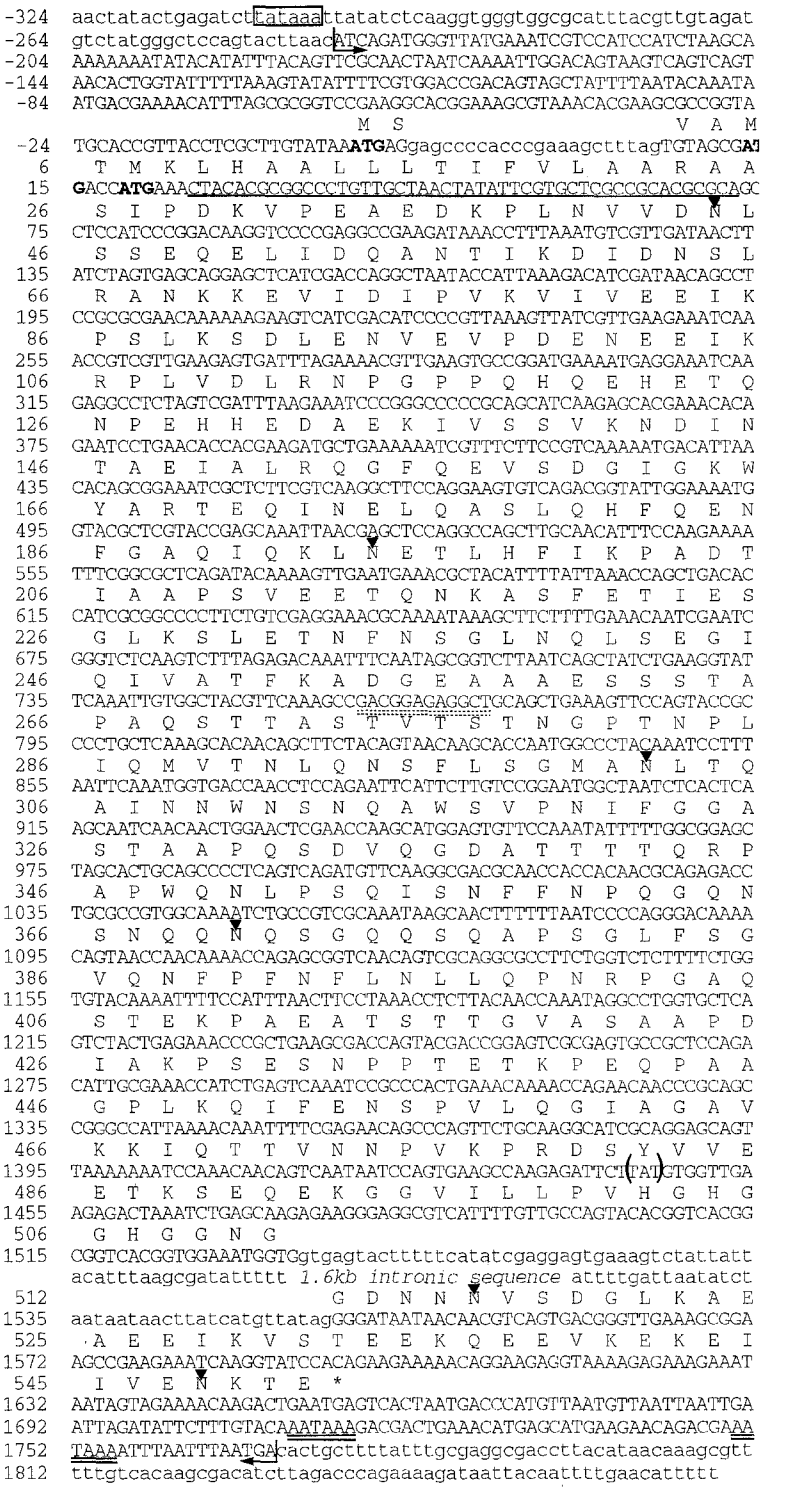

Fig. 4 Composite nucleotide sequence of *Urbain* transcript, its flanking genomic DNA and the predicted amino acid sequence. Nucleotides are *numbered in the left margin.* Amino acid sequence deduced from the putative open reading frame is shown *above the nucleotide sequence.* Nucleotide at position 1 is defined as the first base of the first ATG codon. The three potential translation initiation sites are in *bold type.* The signal peptide is *underlined.* The sites for N-linked glycosylation are indicated by *triangles.* The 2 poly(A) addition signals are indicated by a *double line.* The DGEA sequence is indicated by a *dotted double line.* The putative 5′ end and the 3′ end of the transcript are indicated by *arrows;* the presumptive TATA box is *boxed.* Intron sequence and genomic sequences flanking the transcription unit are in *lowercase.* The TAT triplet, *in parenthesis,* encoding a Tyr residue on the cDNA corresponds to a GAA triplet (Glu residue) of the genomic DNA

shown that the cDNA and genomic DNA are colinear, except for a Glu residue (position 1444–1446) which replaces a tyrosine found on the cDNA (see Fig. 4). The sequence of the mRNA (Fig. 4) contains a putative open reading frame of 1656 nucleotides flanked by a 3′ untranslated region of 113 nucleotides and a 5′ untranslated region of 241 nucleotides. Two polyadenylation consensus sequences (Proudfoot and Brownlee 1976) are defined at positions 1711 and 1750. The second polyadenylation site, located at a functional distance of 15 nucleotides upstream from the poly-A tail, could be the active one.

By comparison of the genomic and cDNA sequences two introns were found. A small putative intron (22 bp) is present downstream from the first ATG codon between nucleotides 5 and 6; however, the splice junctions (ag/GAgc and tttAG/tg) are rather rare. It is more conceivable that the presence of the 22 bp fragment is due to

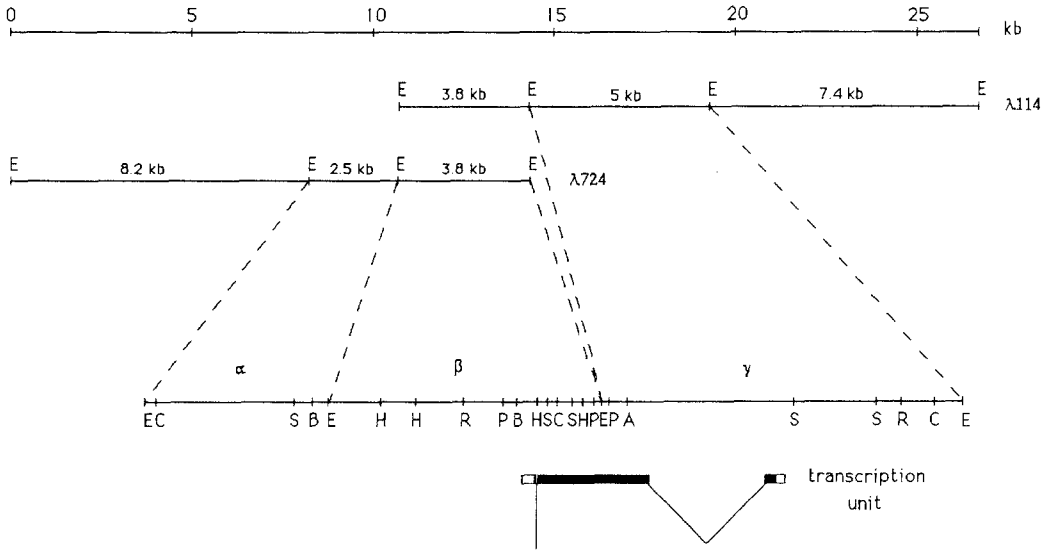

Fig. 5 Schematic organization of *Urbain* gene. The scale is in kilobases *(upper line).* The two genomic clones λ 114 and λ 724 were isolated from the genomic library. Restriction sites are shown for the α,β, γ genomic segments that were cloned independently in the Bluescript plasmid. The restriction enzymes used are: *A* Acc 1, *B* Bgl II, *C* Cla 1, *E* EcoR1, *H* Hind III, *P* Pst 1, *R* EcoRV, *S* Sac I. The *lower line* represents the transcribed region of the *Urbain* gene. *Open boxes* indicate the 5′ and 3′ untranslated regions; the 5′ end of the transcription unit is not precisely determined. The two introns are mapped as *thin lines.* The *filled box* represents the protein-coding region

polymorphisms between the two *Bombyx* strains used for the cDNA study (strain 200–300) and for the genomic library (strain 703). A second intron lies between nucleotides 1534–1535 of the cDNA. By restriction site mapping its length has been estimated to 1.6 kb.

The 5′ end of the *Urbain* transcript has not yet been determined precisely; however preliminary primer extension experiments suggest that it could be located 241 nucleotides upstream of the first translation initiation site (position 1 in Fig. 4). The proposed transcription start site ATCAGAT agrees well with the *Drosophila* start consensus sequence (Cherbas et al. 1986). However, the potential TATA box is located 67 nucleotides upstream from this transcription initiation site, which seems an unusually long distance, and thus suggests that either the effective transcription start site is more upstream than estimated or that *Urbain* does not utilize a TATA box to initiate transcription.

Three in-frame translation start sites are found in closely spaced positions, starting at nucleotide 1, at nucleotide 13 and at nucleotide 19. The nucleotidic context of these ATG triplets does not correspond to the optimum consensus context for the initiation site of translation (Kozak 1989; Cavener and Ray 1991). In the case of *Urbain* mRNA, the three sites could be used for increasing the efficiency of the translation (Kozak 1991).

The *Urbain* gene encodes a putative polypeptide of 551 amino acids. Searches of amino acid sequence data show that *Urbain* protein does not present any structural motif which places it among a particular family of proteins. Nevertheless, a DGEA sequence (nucleotides 760–771) has been found. In vertebrates, the DGEA site is a recognition sequence for collagen I by an integrin (Staatz et al. 1991). The hydropathy analysis (Kyte and Doolittle 1982) presented in Fig. 6A, indicates that the *Urbain* protein is predominantly hydrophilic, except for a prominent hydrophobic segment of 16 amino acids at the N-terminal region of the coding sequence. This latter region corresponds to a peptide signal sequence which suggests that *Urbain* is a secreted protein (Perlman and Halvorson 1983; Heinje 1985). This segment is separated from the methionine codon (nucleotides 19–21) by a lysine one, which is in agreement with the study of Perlman and Halvorson (1983). Six potential N-linked glycosylation sites are found along the protein (Fig. 4).

Discussion

We report here the characterization of a gene encoding a novel protein in *Bombyx.* Indeed, searches in sequence data banks have failed to identify any proteins with significant similarity to the *Urbain* product. Comparison of the predicted amino acid sequence reveals no striking features with catalogued protein structural motifs, except the DGEA sequence. However, the functionality of a DGEA binding site is obscure in *Bombyx,* since the only collagen molecule that has been described in this insect is not collagen I (Chareyre et al., in preparation). At the N-terminal part of *Urbain,* the peptide signal indicates that the protein is probably secreted, this is corroborated by the hydrophilic character of the protein and the presence of potential N-glycosylation sites.

Nevertheless, alignment with the Clustal method (Higgins and Sharp 1988) shows a significant similarity with the IMP-E2 protein: 18% of the amino acids are identical and 42% correspond to conservative substitutions (Table 1). This *Drosophila* protein has been described by Paine-Saunders et al. (1990) as an ecdysone-dependent product, secreted and expressed during wing

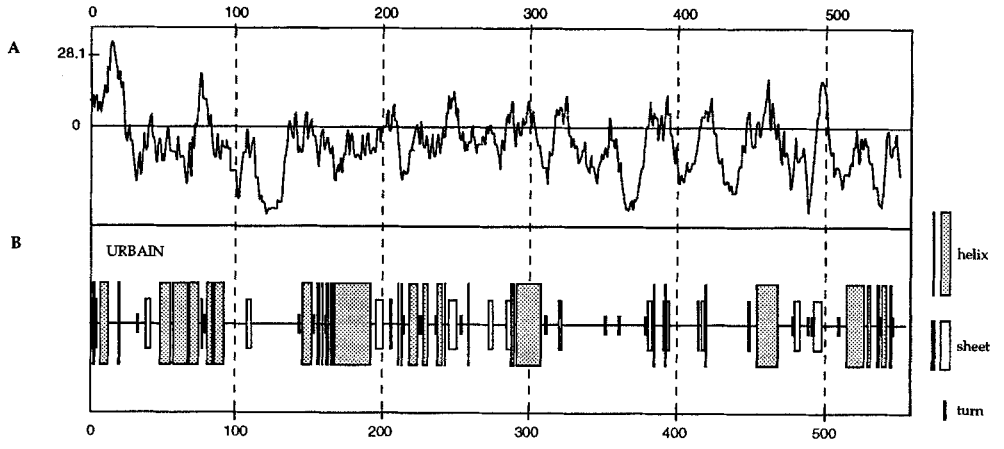

Fig. 6 **A** Hydropathy plot of *Urbain* protein. The hydropathy of the *Urbain* protein is evaluated according to the method of Kyte and Doolittle (1982). Amino acids are numbered from the first translation start site. *Positive values* correspond to hydrophobic regions. **B** Secondary structure predictions of the *Urbain* protein. The amino acids are *numbered on the abscissa axis.* The *dotted boxes* and the *great vertical bars* represent respectively large and narrow regions of α-helices; the *open boxes* and the *middle-sized vertical bars* indicate large and narrow regions of β-sheets; the *small vertical bars* represent turns

disc morphogenesis. The matching residues are distributed throughout the proteins. In the two polypeptides, alanine, glutamic acid and glutamine or asparagine are the most frequent amino acids. Like IMP-E2, *Urbain* contains several EIK-type repeats which consist of short-sequence motifs with an acidic amino acid, an hydrophobic one and then lysine; but for the *Urbain* protein, the motifs are more variable in number and spacing and are located at each end of the protein.

The secondary structure has been predicted according to Geourjon and Deléage (1994) and is presented in Fig. 6B. It reveals that the *Urbain* protein can be divided into three major regions. The first one contains approximately 300 residues at the N-terminus and presents mainly an α-helical structure. A second region (from residues 300 to 460) is poorly organized. The last region lies at the C-terminus of the protein and contains several stretches of α-helix, mixed with β-sheets. The N-terminal region of the IMP-E2 protein also presents long domains with α-helices (Paine-Saunders et al. 1990), whereas the remainder of the protein contains α-helix and β-sheet parts. Thus, both the structural organization (secreted proteins, large domains of similarity, α-helical structures) and the related developmental expression patterns suggest that the two proteins might belong to related families in charge of similar functions outside the cell. This hypothesis must, however, be verified by functional analysis of specific *Urbain* protein domains (work in progress). Immunolocalization results on IMP-E2 protein (Paine-Saunders et al. 1990) showed that the protein is secreted apically and could be essential in remodelling the wing disc cell surface. However, the hormonal regulatory pathways of IMP-E2 and *Urbain* genes are different; the first one is a primary gene induced by 20-hydroxyecdysone whilst *Urbain* shows a secondary response to the same hormone (see below). The significance of such differences in hormonal regulation for genes encoding putatively similar proteins is not known; this might reflect different requirements for development of Diptera and Lepidoptera.

The *Urbain* gene requires the action of 20-hydroxyecdysone to be expressed. The transcript accumulation is induced either in vitro after hormonal stimulation or in vivo at the peak of hormone secretion during each stage that has been examined. Moreover, mRNAs accumulate in response to hormone in a dose-dependent manner: in vitro, the transcript is present at a higher level with a higher 20-hydroxyecdysone concentration; in vivo, the abundance of *Urbain* transcript during the fourth and fifth larval stadia reflects the hormonal level. This suggests that the developmental pattern of *Urbain* transcript accumulation is regulated both by changes in the hormone level and by the duration of the peaks.

The in vitro experiments show that *Urbain* mRNA accumulation is delayed after hormone addition and requires synthesis of regulatory proteins. These proteins could be encoded by 20-hydroxyecdysone-inducible genes, homologous to such *Drosophila* early genes as Broad-Complex (Karim et al. 1993), and could influence *Urbain* gene expression either by acting on the promoter or by affecting some posttranscriptional mechanisms. Thus, temporal expression and in vitro requirements show that *Urbain* behaves as a late or secondary response gene with respect to 20-hydroxyecdysone action, according to the classification proposed by Ashburner et al. (1974) for *Drosophila* salivary gland genes. Late genes have been characterized in *Drosophila* wing discs: IMP-L1 (Natzle et al. 1992), IMP-L2 (Osterbur et al. 1988). Like *Urbain* in *Bombyx,* they are detected in wing discs several hours after the increase of 20-hydroxyecdysone titer and require protein synthesis to be induced. However, there are several differences between *Bombyx* and *Drosophila* wing disc genes.

Whereas the IMP-L2 gene is expressed in prepupae and during embryonic development without exclusive relation to the hormonal level (Garbe et al. 1993) and the

Table 1 The alignment of *Urbain* and IMP-E2 proteins was performed by using the Clustal method (Higgins and Sharp 1988; *Stars and bold letters* indicate identical residues, *italicized letters and dots* indicate conserved residues)

```
URBAIN          MSVAMTMKLHAALLLTIFVLAARAASIPDKVPEAEDKPLNVVDNLSSEQELIDQA   55
IMP2-DROME      MK.......PVALIL.VF.LAISQARVLNLPKEAIDIPVAIVEDKEPPVALSLVK   46
Consensus       *.       ..**.* .* ** ..*.. .   ** * *...*.. ..  .*

URBAIN          NTIKDIDNSLRANKKEVIDIPVKVIVEEIKPSLKSDLENVEVPDENEEIKRPLVD 110
IMP2-DROME      EEVKAEEVKPEEVKPIAQEEKAKDLKEEVKPEIKPEIKEQPKPDIKDEIKEDL..  99
Consensus       ...*. . .  . *  . .  .* . **.**..*.....   **  ..***  *

URBAIN          LRNPGPPQHQEHETQNPEHHEDAEKIVSSVKNDINTAEIALRQGFQEVSDGIGKW 165
IMP2-DROME      .............KADIKEELKEKIEEQINELPNAKPLELKE.........KS 130
Consensus                     ..  ..*  *** . ...  *.  ..*..          *

URBAIN          YARTEQINELQASLQHFQENFGAQIQKLNETLHFIKPADTIAAPSVEETQNKASF 220
IMP2-DROME      LEAEEKPQEIKEEVQQ......PEIKK..................EATEIKE.. 158
Consensus        . .*. .*.....*.       ..*.*                  *.*. *.

URBAIN          ETIESGLKSLETNFNSGLNQLSEGIQIVATFKADGEAAAESSSTAPAQSTTASTV 275
IMP2-DROME      EPAQNILKSLPAEETVVVPAEELSPNPVEQEQSENQDAAHPQVRQATQATPTQQS 213
Consensus       *. .. **** .. .  .  .  .  .  *.  ......**..   ...*.*..

URBAIN          TSTNGPTNPLIQMVTNLQNSFLSGMANLTQAINNWNSNQAWSVPNIFGGASTAAP 330
IMP2-DROME      TTQGNFVQQLIQN.............SPIGQFLNQFQPQPA.........AAAAPA 247
Consensus       *. .. ...***              . ..* .*......*         *..*..

URBAIN          QSDVQGDATTTTQRPAPWQNLPSQISNFFNPQGQNSNQQNQSGQQSQAP....SG 381
IMP2-DROME      AAQVQADDAAAAAAPATP....APTVPGFLNPQAAITSAQQAVQNAAQSAVNATTQ 298
Consensus       ...**.*.........*     .. ...*.***.. ...*..  ...*..     .

URBAIN          LFSGVQNFPFNFLNLLQPNRPGAQSTEKPAEATSTTGVASAAPDIAKP..SESNP 434
IMP2-DROME      AFQGIQQFASNLGNQFQNTLSSLTGQQQQAVSTTPRPPGPIQQFVNNVFGGNNNA 353
Consensus        * *.*.*. *. * .* . ..  . ...* .*..   .. .. ...   ...*.

URBAIN          PTETKPEQPAAGPLKQIFENSPVLQGIAGAVKKIQTTVNNPVKPRDSYVVEETKS 489
IMP2-DROME      TAAAPPAQQSGNPLQGIIN.......FLGGNRPQNAPAAAPATQATEKPAVDDKI 401
Consensus       .... *.*....**. *..       . *. .  ......*... ..  . ..*
```

IMP-L1 gene is expressed during and after pupariation (Natzle et al. 1992), *Urbain* expression is induced at each 20-hydroxyecdysone peak during successive larval stages of *Bombyx* and at metamorphosis. This time course is consistent with a role for the *Urbain* gene in events including the growth and morphogenesis of the wing discs.

However, the dose-dependence for *Urbain* accumulation is not verified when cultured discs are incubated with the extraphysiological concentration of 10^{-4} M 20-hydroxyecdysone (Fig. 2A), and in pupae where *Urbain* is transcribed in two waves during the peak of 20-hydroxyecdysone (Fig. 1). The temporal regulation may therefore be complex and different pathways of several regulatory proteins might be involved according to the developmental stage and/or the hormone level. Thus, the *Urbain* gene can be a good candidate for identifying upstream genes; these unknown genes could interact from hormonal induction to developmental tissue modifications.

Urbain expression has a tissue distribution different from that of the *Drosophila* late genes. Whereas IMP-L1 mRNA is detected only in discs (Natzle et al. 1992), IMP-L2 expression is not restricted to the discs but is implicated in the development of the nervous system (Garbe et al. 1993). In situ hybridizations show that *Urbain* is expressed in wing discs and epidermis. The latter has the same ectodermal origin as the wing primordia and also undergoes complex, hormonally-induced modifications during insect life. In contrast, the silk glands which also originate from the ectoderm, do not express *Urbain;* this might be related to their developmental fate since they will degenerate after cessation of spinning. A slight labelling is seen on fat body, another 20-hydroxyecdysone inducible tissue (Lepesant et al. 1986). These spatial and temporal localizations suggest that *Urbain* is required in different tissues to realize a specific phase of the differentiative processes initiated by the hormone. This specific phase occurs repeatedly throughout larval

life and at metamorphosis -a period characterized by massive tissue reorganization and morphogenesis. Also, the delay between appearance of the hormone peak and accumulation of the *Urbain* transcript suggests that the protein is involved in late events. Identification of these events and analysis of the role and regulation of the production of this protein (work in progress) will help to generate some insight into the mechanisms of tissue reorganization.

In conclusion, we suggest that *Bombyx* wing discs can provide a convenient system in which to study mechanisms underlying wing morphogenesis, and we hypothesize that the *Urbain* gene product may be involved in the process of cell rearrangement leading to changes in tissue size and shape.

Acknowledgements We thank Christian Guillet for the gift of Z17 cDNA clone. We are grateful to Frédéric Humbert for technical assistance. We thank the C.M.A.B. (Université Lyon-1), and particularly Jérôme Briolay for help with the PCR methodology. We thank Madeleine Coulon for critical reading of the manuscript. This research was supported by a grant from the French Ministry for Research and Space (Developmental Biology, grant 92C0452).

References

Ashburner M, Chihara C, Meltzer P, Richards G (1974) Temporal control of puffing activity in polytene chromosomes. Cold Spring Harbor Symp Quant Biol 38:655–662

Blair SS, Brower DL, Thomas JB, Zavortink M (1994) The role of *apterous* in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of *Drosophila*. Development 120:1805–1815

Calvez B, Hirn M, De Reggi M (1976) Ecdysone changes in the haemolymph of two silkworms *(Bombyx mori* and *Philosamia cynthia)* during larval and pupal development. FEBS Lett 71: 57–61

Calvez B (1981) Progress of developmental programme during the last larval instar of *Bombyx mori:* Relationships with food intake, ecdysteroids and juvenile hormone. J Insect Physiol 27: 233–239

Cavener DR, Ray SC (1991) Eukaryotic start and stop translation sites. Nucl Acids Res 19:3185–3192

Chareyre P, Guillet C, Besson MT, Fourche J, Bosquet G (1993) Genes expressed during the differentiation of the pupal wings are also transiently expressed during the larval moult in *Bombyx mori.* Insect Mol Biol 2:239–246

Chareyre P, Besson MT, Fourche J, Bosquet G (1995) Characterization of a cDNA for a protein with multiple domains of Gly-Xaa-Yaa repeats identifies a *Bombyx* collagenous protein with uncommon features. In preparation

Charles JP, Bouhin H, Quennedey B, Courrent A, Delachambre J (1992) cDNA cloning and deduced amino acid sequence of a major, glycine-rich cuticular protein from the coleopteran *Tenebrio molitor.* Temporal and spatial distribution of the transcript during metamorphosis. Eur J Biochem 206:813–819

Cherbas L, Schulz RA, Koehler MM, Savakis C, Cherbas P (1986) Structure of the Eip28/29 gene, an ecdysone-inducible gene from *Drosophila.* J Mol Biol 189:617–631

Clemens MJ (1984) Purification of eukaryotic messenger RNA. In: Hames BD, Higgins SJ (eds) Transcription and translation, a practical approach. IRL Press, Oxford, pp 211–270

Coulon M (1977) Relations between the feeding and radiosensibility cycles and the ecdysone cycle in *Bombyx mori,* at the end of embryonic development and during the first larval instars. Dev Growth Diff 19:181–185

Eickbush TH,Kafatos FC (1982) A walk in the chorion locus of *Bombyx mori.* Cell 29:633–643

Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. Anal Biochem 137:266–267

Fristrom D, Wilcox M, Fristrom J (1993) The distribution of PS integrins, laminin A and F-actin during key stages in *Drosophila* wing development. Development 117:509–523

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85:8998–9002

Garbe JC, Yang E, Fristrom JW (1993) IMP-L2: an essential secreted immunoglobulin family member implicated in neural and ectodermal development in *Drosophila.* Development 119:1237–1250

Geourjon C, Deléage G (1993) Interactive and graphic coupling between multiple alignments, secondary structure predictions and motif/pattern scanning into proteins. CABIOS 9:87–91

Geourjon C, Deléage G (1994) SOMP: a self optimized method for protein secondary structure prediction. Protein Eng 7:157–164

Heinje G von (1985) Signal sequences. The limits of variation. J Mol Biol 184:99–105

Higgins DG, Sharp PM (1988) Clustal: A package for performing multiple sequence alignment on a microcomputer. Gene 73: 237–244

Ingham PW (1988) The molecular genetics of embryonic pattern formation in *Drosophila.* Nature 335:25–33

Karim FD, Guild GM, Thummel CS (1993) The *Drosophila Broad-Complex* plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis. Development 118:977–988

Kiguchi K, Agui N (1981) Ecdysteroid levels and developmental events during larval moulting in the silkworm *Bombyx mori.* J Insect Physiol 27:805–812

Kozak M (1989) The scanning model for translation: An update. J Cell Biol 108:229–241

Kozak M (1991) An analysis of vertebrate mRNA sequences: Intimations of translational control. J Cell Biol 115:887–903

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Lepesant JA, Maschat F, Kejzlarovà-Lepesant J, Benes H, Yanicostas C (1986) Developmental and ecdysteroid regulation of gene expression in the larval fat body of *Drosophila melanogaster.* Arch Insect Biochem Physiol 1:133–141

Nardi JB (1994) Rearrangement of epithelial cell types in an insect wing monolayer is accompanied by differential expression of a cell surface protein. Dev Dynamics 199:315–325

Natzle JE (1993) Temporal regulation of *Drosophila* imaginal disc morphogenesis: a hierarchy of primary and secondary 20-hydroxyecdysone-responsive locus. Dev Biol 155:516–532

Natzle JE, Hammonds AS, Fristrom JW (1986) Isolation of genes active during hormone-induced morphogenesis in *Drosophila* imaginal discs. J Biol Chem 261:5575–5583

Natzle JE, Robertson JP, Majumdar A, Vesenka GD, Endlow B, Clark KE (1992) Sequence and expression of IMP-L1, an ecdysone-inducible gene expressed during *Drosophila* imaginal disc morphogenesis. Dev Genet 13:331–344

Nijhout HF (1994) Genes on the wing. Science 265:44–45

Nüsslein-Volhard C (1991) Determination of the embryonic axes of *Drosophila.* Development Suppl: 1–10

Osterbur DL, Fristrom DK, Natzle JE, Tojo SJ, Fristrom JW (1988) Genes expressed during imaginal discs morphogenesis: IMP-L2, a gene expressed during imaginal disc and imaginal histoblast morphogenesis. Dev Biol 129:439–448

Paine-Saunders S, Fristrom D, Fristrom JW (1990) The *Drosophila* IMP-E2 gene endoces an apically secreted protein expressed during imaginal disc morphogenesis. Dev Biol 140:337–351

Perlman D, Halvorson HO (1983) A putative signal peptidase recognition site and sequence in eucaryotic and procaryotic signal peptides. J Mol Biol 167:391–409

Proudfoot NJ, Brownlee GG (1976) The 3′ non-coding region sequences in eukaryotic messenger RNA. Nature 263:211–214
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
Sanger F, Nicklen S, Coulson A (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
Staatz WD, Fok KF, Zutter MM, Adams SP, Rodriguez BA, Santoro SA (1991) Identification of a tetrapeptide recognition sequence for the $\alpha 2\beta 1$ integrin in collagen. J Biol Chem 266: 7363–7367

Note added in proof

Sequence data described in this paper have been submitted to the EMBL data library under accession number Z47409.

Wilcox JN (1993) Fundamental principles of *in situ* hybridization. J Histochem Cytochem 41:1725–1733
Williams JA, Paddock SW, Carroll SB (1993) Pattern formation in a secondary field: a hierarchy of regulatory genes subdivides the developing *Drosophila* wing disc into discrete subregions. Development 117:571–584
Williams JA, Paddock SW, Vorwerk K, Carroll SB (1994) Organisation of wing formation and induction of a wing-patterning gene at the dorsal/ventral compartment boundary. Nature 368:299–305