Int. J. Dev. Biol. 37: 499-507 (1993)

Original Article

Protein-DNA interactions at putative regulatory regions of two coordinately expressed genes, msp130 and PM27, during skeletogenesis in sea urchin embryos

VENU RAMAN[1]*, MARY E. ANDREWS[1], MICHAEL A. HARKEY[2] and RUDOLF A. RAFF[1]

[1]*Institute for Molecular and Cellular Biology and Department of Biology, Indiana University, Bloomington and* [2]*Department of Botany, University of Washington, Seattle, USA*

ABSTRACT Development of the primary mesenchyme cells (PMCs) of the sea urchin embryo, which give rise to the larval skeleton, involves the coordinate onset of expression of several structural genes. As part of an effort to identify cis-acting elements that might play a role in this regulatory event, co-regulated genes were examined by two approaches. First, they were compared for conserved sequence elements. Four conserved elements were found as a cluster in all three genes examined, suggesting a regulatory role. Second, as a test for potential function, the putative regulatory regions of two of these genes were examined for protein binding sites. DNase I protection and gel mobility shift assays were used to: 1) identify several nuclear protein binding sites in these regions, two of which correspond to conserved elements among the genes; 2) demonstrate that the developmental time of appearance of the proteins that interact with these sites corresponds to the time of activation of the genes; and 3) show that two of the conserved sequence elements shared by these genes compete for the same binding proteins. These data identify putative regulatory elements, whose specific roles in the coordinate regulation of PMC-expressed genes can now be addressed directly using appropriate transgenic expression constructs.

KEY WORDS: *gel retardation, gene regulation, sea urchin embryo, promoter elements, protein-DNA interactions*

Introduction

The primary mesenchyme cells (PMCs) of the sea urchin embryo are the sole progenitors of the larval skeleton. The PMCs are derived from the four micromeres of the 16-cell stage embryo. At the late blastula stage, 32 daughter cells of the micromeres ingress from the blastula epithelium into the blastocoel space and assume mesenchymal characteristics. These PMCs migrate actively within the blastocoel space, assembling into a pattern that foreshadows the shape of the skeleton. Ultimately these cells lay down a matrix that is mineralized with calcium and magnesium carbonate to form the branched spicular skeleton characteristic of the pluteus larva.

This development is accompanied, at about the time of PMC ingression, by the coordinate onset of expression of several structural genes. Several transcripts (Harkey *et al.*, 1988) and at least 10 major proteins (Harkey and Whiteley, 1983; Benson *et al.*, 1987; Sucov *et al.*, 1987) appear at this time, exclusively in the PMC lineage (Harkey and Whiteley, 1983; Carson *et al.*, 1985; Leaf *et al.*, 1987). The PMC-specific transcripts accumulate to peak levels by gastrula, and remain at high levels through larval development (Harkey *et al.*, 1988; Benson *et al.*, 1987; Leaf *et al.*, 1987). Most or all of the proteins whose synthesis begins at ingression continue to be synthesized through skeletogenesis, suggesting that they play roles in skeleton formation (Anstrom *et al.*, 1987; Benson *et al.*, 1987; Leaf *et al.*, 1987). Several of these proteins have now been identified as spicule matrix proteins (Harkey and Whiteley, 1983; Benson *et al.*, 1987; Sucov *et al.*, 1987).

The co-expression of these genes in skeletogenesis suggests the existence of a mechanism for coordinate gene regulation in PMCs. A simple model of cell-specific coordinate regulation of a set of genes could involve the timely and cell-specific synthesis or activation of a transcription factor that would bind to positive regulatory elements shared by all of the genes in the set. Such a model would include: 1) all of the genes in the set contain one or

Abbreviations used in this paper: ASW, artificial seawater; bp, base pairs; kb, kilobase pairs; PMCs, primary mesenchyme cells.

***Address for reprints**: Institute for Molecular and Cellular Biology and Department of Biology, Indiana University, Bloomington, IN 47405, USA. Fax: 812-855-6082.

0214-6282/93/$03.00

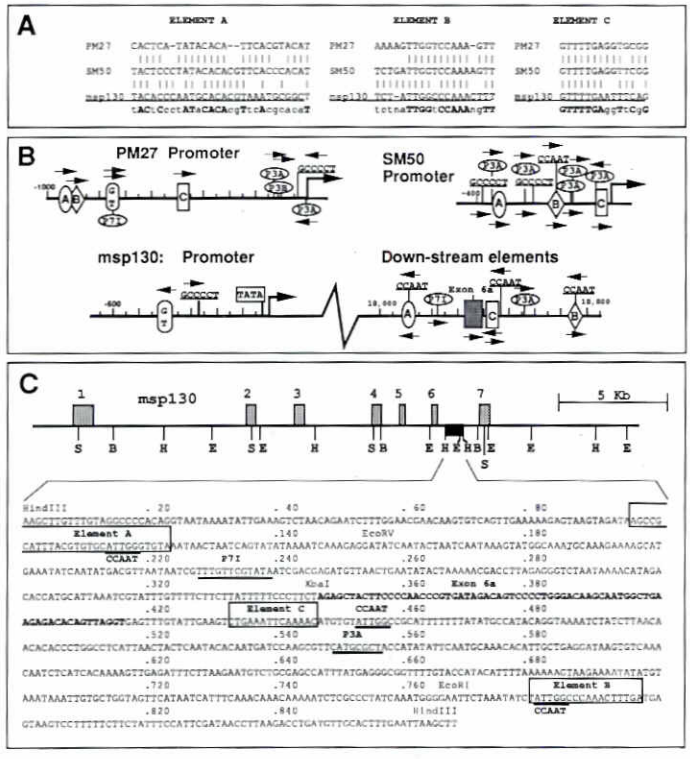

Fig. 1. Conserved elements of the PMC-expressed genes. (A) *Alignment of elements from individual genes. Vertical lines indicate identical nucleotides in a given pair of sequences. A consensus sequence is shown below each set of comparisons, where bold letters indicate fully conserved nucleotides.* **(B)** *Arrangement and orientation of various sequence elements in the three skeletal genes. Scales are in units of nucleotides pairs. Symbols: A=Element 'A', B=Element 'B', C=Element 'C', GT=5' GTbox, P3A, P3B, and P7I are sequence elements homologous to the nuclear protein binding motifs. Relative orientations of the sequence elements are indicated by arrows. Specific positions of elements in the SM50 promoter are as follows: Element 'A' starts at -331; Element 'B' starts at-163; Element 'C' starts at -25; GCCCCT motifs starts at -231 and -378; P3A sites starts at -34, -114, 128, -248, and -354. Specific positions of elements in the msp130 promoter are as follows: 5' GT box starts at -401, and the GCCCCT motif starts at -235. Specific positions of elements within the downstream region in the msp130 are as indicated below.* **(C)** *Location of elements 'A', 'B', and 'C' in the msp130 gene. A map of the msp130 gene is shown at the top. Boxes 1-7 represent the first 7 exons. Restriction sites: S=SalI, B=BamHI, H=HindIII, E=EcoRI. The black box represents the HindIII fragment that hybridizes to element 'A' and 'B' probed from PM27. The sequence of this HindIII fragment is shown below. Sequence motifs homologous to elements 'A', 'B', and 'C' of PM27 are shown in boxes. A small exon of the msp130 gene, designated exon 6a, is indicated by bold type.*

more functionally homologous cis-acting regulatory elements; 2) the embryo contains a nuclear protein that can bind to these sites; and 3) this factor appears developmentally at the time of activation of the genes.

Three of the genes coordinately expressed in the PMCs of *Strongylocentrotus purpuratus* have been isolated and characterized in some detail. SM50 and PM27 encode spicule matrix proteins of 50 kD and 27 kD, respectively (Sucov *et al.*, 1987, 1988; Harkey, in preparation). Msp130 encodes a 130 kD cell surface glycoprotein (Anstrom *et al.*, 1987; Leaf *et al.*, 1987; Parr *et al.*, 1990). The promoters of SM50 and PM27 have both been shown to contain sufficient information to drive PMC-specific expression of reporter genes in transgenic embryos (Sucov *et al.*, 1987, 1988; Harkey, in preparation).

In this report we provide physical evidence consistent with a simple coordinate mechanism for regulation of these genes. Comparison of the putative regulatory regions reveals that each gene contains a cluster of four types of conserved sequence elements. Protein binding data for two of these genes characterize several sites on each gene that exhibit sequence-specific binding to nuclear proteins. The proteins that bind these sites only appear after ingression, when the genes are normally activated. Two of the binding sites correspond to sequences conserved among all the co-expressed genes. These conserved sites exhibit intergenic competition for binding proteins, demonstrating that they are functionally as well as structurally homologous.

Results

Conserved sequence elements among PMC-expressed genes

The similar patterns of temporal and spatial expression of PM27, SM50, and msp130 suggest that these genes might contain similar cis-regulatory elements. Promoter sequences have been published for msp130 (-1 to -666; Parr *et al.*, 1990), SM50 (-1 to -443; Sucov *et al.*, 1987) and PM27 (-1 to -1079; Harkey, in preparation). Sucov *et al.* (1988) have shown that nucleotides +120 to -440 of SM50 are sufficient to drive appropriate expression of a reporter gene in transgenic embryos. Similarly, nucleotides +20 to -1079 are sufficient to drive PMC-specific expression of a PM27 transgene (Harkey, in preparation). As a first approach to identifying potential skeleton-specific regulatory sequences, pair-wise homology matrix comparisons were made between the SM50 and msp130 promoters, and the PM27 gene and flanking sequence (nucleotides -1079 to +1768).

As the stringency of the length and mismatch criteria for the comparisons was increased, three sequences emerged as conserved elements in the PM27 and SM50 promoters. These elements (designated elements 'A', 'B', and 'C' in Fig. 1A) are all located upstream of the transcription start site of the PM27 gene. As shown in Fig. 1A, the sequences of elements 'A', 'B', and 'C' show identity between the two genes in 21/27, 13/14, and 13/14 contiguous base pairs, respectively (allowing for gaps). The three elements are arranged in the same relative positions and orientations with respect to the start sites of these genes (Fig. 1B). No significant sequence identity was observed in the regions between the elements. Note that both promoters also contain multiple P3A sites, as previously described for SM50 by Thiebaud *et al.* (1990).

No sequences similar to elements 'A', 'B', or 'C' were found in the published 666 base pairs of the msp130 promoter, which also lacks a P3A site. The search for sequences homologous to these elements was continued for an additional 2 kb upstream of the msp130 gene. No homologies were found (K. Klueg and R.A. Raff, unpublished data). Homology matrix comparisons revealed only one conserved sequence in the msp130 promoter, a 5'GTbox, which was found in PM27 but not in SM50 (Fig. 2A and 2B). A search for a GCCCCT motif, which has been noted previously to be shared by SM50 and msp130 (Parr *et al.*, 1990), located this sequence in the PM27 promoter (Fig. 1B).

Since msp130 might harbor important regulatory elements downstream of the promoter, a search for 'A'- and 'B'-like sequences was conducted over the entire msp130 gene. Fig. 1C shows the exon map of msp130 (from Parr *et al.*, 1990), and the associated restriction map developed from several overlapping genomic clones identified by hybridization to the SpLM18 cDNA (Harkey *et al.*, 1988). Oligonucleotide probes containing either sequences from element 'A' or element 'B' were hybridized to DNA gel blots containing restriction digests of these genomic clones. Relaxed stringency conditions permitted mismatches as high as 50%. These probes hybridized preferentially to specific *Bam*HI, *Hind*III and *Eco*RI fragments that overlapped a region between exons 6 and 7. An 850bp *Hind*III fragment, which appeared to contain homologies for both elements, was sub-cloned and sequenced (Fig. 1C). All three elements ('A', 'B', and 'C') are represented within this fragment, along with a P3A site. This fragment also contains a small exon designated as Exon 6a.

The sequences for elements 'A', 'B', and 'C' in this fragment are compared with those of the PM27 and SM50 promoters revealing a consensus sequence (Fig. 1A). In element 'A', the CACA tetrad is conserved along with several isolated bases. This tetrad appears twice within the element 'A' motifs in the SM50 and msp130 genes, and at least once again in the adjacent DNA of all three genes. In element 'B', a central palindrome is conserved, along with three additional nucleotides. The arrangement of the various elements in msp130, which differs from that of PM27 or SM50, is illustrated in Fig. 1B.

The EMBL and Gene Bank databases were searched for sequences similar to elements 'A', 'B', and 'C', using relaxed criteria consistent with the mismatch frequencies between the elements in PM27, SM50, and msp130. The results of the search indicated that each of these sequences occurs individually about as often as would be expected by chance in a random base sequence. In no cases, other than the genes discussed here, did two of the three elements' homologies appear in the same gene. Thus, the clustering of these sequence elements is a unique property of the PMC-expressed genes.

Identification of nuclear protein binding sites in PM27 and msp130

Suspected regulatory domains of PM27 and msp130 were surveyed for potential protein binding sites by DNase I protection assay. The domains examined included the 5' flanking regions of msp130 (-550 to +133) and PM27 (-1060 to +158), and the downstream 863bp *Hind*III fragment containing the A-B-C-P3A element cluster of msp130 (Fig. 2A and B). A series of probes was generated from contiguous restriction fragments from each of these regions (shown below the gene maps in Fig. 2). The probes were incubated with nuclear extracts from gastrula stage embryos, which express these genes at peak levels, and assayed for protected protein binding sites.

As summarized in Fig. 2, several sites were reproducibly protected from DNase I by the gastrula extracts. The 5' flanking region of msp130 (Fig. 2A) exhibited a single detectable binding site, which correlated with a 'CCAT' motif. Elements labeled 'A', 'B', 'AR', and 'MV' in the downstream fragment of this gene were also protected. The PM27 5' flanking region exhibited only one detectable binding site corresponding to the closely apposed 'A' and 'B' elements of that gene (Fig. 2B). There was no reproducible protection of the P3A sites, the conserved 'C' elements, or any other region of the sequences examined.

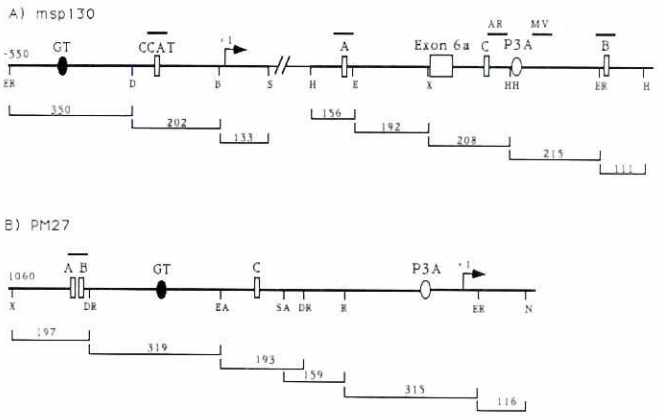

Fig. 2. Schematic representation of the templates and probes used for DNAase I footprinting assays. (Panel A) *The putative regulatory regions of msp130 gene. The region from +133 bp to -550 bp, and the 863 bp sequence (15 Kb downstream of the transcription start site) which contains the conserved elements 'A', 'B', and 'C' were used for the protection assays.* **(Panel B)** *The promoter region of PM27 gene. The region used for protection assays was from +158 bp to -1060 bp. The lines beneath the diagram of the genes represent the different fragments used as probes for footprinting assays; rectangles designate the three conserved boxes ('A', 'B', 'C'); dark ovals represent the conserved sequence called GT (for G+T rich region); open ovals represent the negative spatial regulator site, P3A; the square represents the exon 6a; short solid overhead lines indicate the protected areas. X, Xba I; DR, Dra I; EA, Ear I; SA, Sau 3AI; R, Rsa I; ER, Eco RI; N, Nae I; D, Dde I; B, Bam HI; S, Sca I; H, Hind III; E, Eco RV; HH, Hha I.*

Protein binding activity of these protected regions was examined in more detail by DNase I protection assays, and electrophoretic mobility shift (gel shift) assays.

Embryonic nuclear extracts from three developmental stages were used in order to detect possible developmental changes in populations of binding proteins that might correlate with changes in expression.

Morula stage (8 h) extracts provided nuclear proteins from embryos that did not express these genes. Gastrula extracts (36 h) represented the stage at which the genes had reached peak levels of expression, and were uniformly expressed among PMCs (Harkey *et al.*, 1988). Pluteus extracts represented a stage at which expression of at least one of the genes, msp130, had switched from the uniform cell-autonomous regulation exhibited at gastrula, to position-specific modulation associated with skeletal growth (Harkey *et al.*, 1992).

Characterization of the msp130 binding sites

DNase protection of the element 'A' site of msp130 by embryo nuclear extracts is shown in Fig. 3. Although morula extracts were ineffective, gastrula extracts conferred clear localized protection, and pluteus extracts exhibited an even stronger effect. The protected region corresponded to the most conserved core of the 'A' element. Gel shift assays confirmed the sequence-specific association of the 'A' element with nuclear proteins from gastrulae (Fig. 3A) and plutei (data not shown). At least two size classes of protein-DNA complexes were detectable that could be competed with excess unlabeled element 'A' DNA (lanes 3 and 4), but not with heterologous DNA (lane 5).

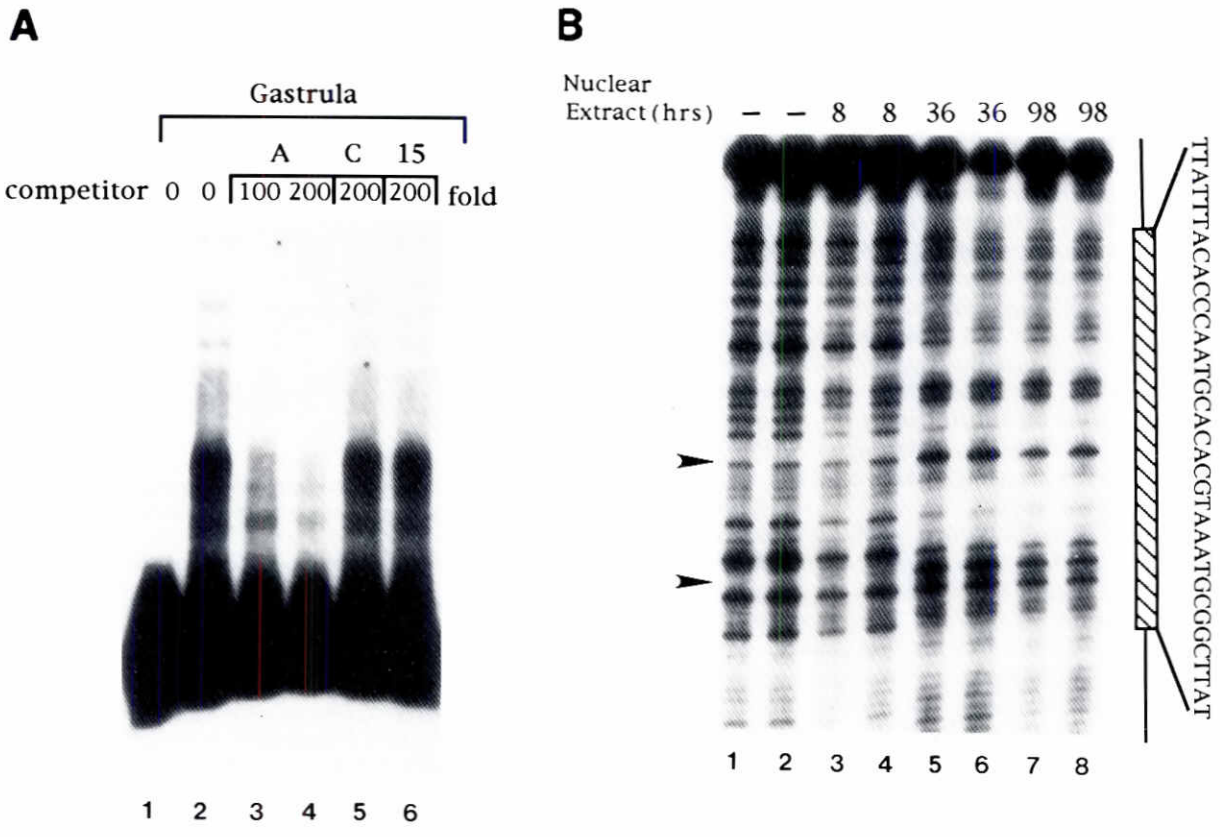

Fig. 3. Gel shift and DNase I footprinting analyses of the 'A' region of msp130 gene with different stage nuclear extracts. *For gel shift analysis, oligomers corresponding to the protected areas (see Materials and Methods) were used to demonstrate the specificity of the nuclear protein binding to the 'A' element. For footprinting, the probe used was 156bp Hind III- Eco RV containing the 'A' element.* **(Panel A)** *Gel-shift competition experiments using the msp130 'A' oligomer and gastrula stage nuclear extract. 5 μg of protein was used per lane. Lane 1 contains no nuclear extract. Lane 2 contains the probe incubated with the nuclear extract. Lanes 3 and 4 contain the indicated unlabeled homologous competitor oligomers. Lane 5 contains unlabeled heterologous oligomers. Lane 6 contains oligomer 15 (corresponding to the 'A' and 'B' regions of the PM27 gene). The band shift complexes are indicated by arrows.* **(Panel B)** *DNase I footprinting of the Hind III- Eco RV fragment containing the 'A' element. Labeled DNA (2ng) was incubated with 20 μg of nuclear extract. Lanes 1 and 2 contain no protein extract. Lanes 3-8 show footprinting reactions with 20 μg of nuclear extract from the different stages of development (indicated above the figure) and 1 unit of DNase I, respectively. Following incubation with the nuclear extract and partial digestion with DNase I, the samples were deproteinized and the nucleic acids resolved by denaturing gel electrophoresis. The DNase I footprint mapped to the same location when footprinted with either labeled strands. The hatched box represents the protected area; arrows indicate DNase I hypersensitive sites. Developmental stages: 8 h, morula; 36 h, gastrula; 98 h, pluteus.*

These complexes were not formed with morula extracts (data not shown), consistent with the inactive state of these genes at this stage. The biological activity of the morula extracts was confirmed in a nuclear run-off transcription assay (data not shown) using an early histone gene, EH4 (Tung *et al.*, 1989).

Similar results were obtained for three sites in the msp130 downstream region (the 'B' element (Fig. 4), the 'AR' site (Fig. 5) and the 'MV' site (data not shown)), as well as for the 'CCAT' motif in the msp130 5' flanking region (Fig. 6). It should be noted that each of these binding sites showed distinct stage-specific patterns of complex formation in the gel shift assay. In each case, no binding was observed in morula extracts, and incubation with the gastrula and pluteus extracts yielded distinct binding patterns. While the 'AR' and 'MV' elements exhibited sequence-specific binding to gastrula extracts, this binding was lost by the pluteus stage. This suggests that the proteins that bind these sites are only transiently present in the active form. In the cases of 'A' and 'B' elements, as well as the 'CCAT' motif, binding occurred in both gastrula and pluteus extracts. However, the gel shift banding patterns were different, suggesting that the number, type, or arrangement of proteins associated with these sites changes during development.

PM27 binding sites

The 5' upstream region (-1060 to +159 bp) of the PM27 gene which contains the three elements ('A', 'B', and 'C'), GT box, P3A and P7I sites was analyzed by DNase I footprinting using the probes shown in Fig. 2B. Protection was observed only in 'A' and 'B' elements bound to gastrula and pluteus stage extracts (data not shown). No binding was observed with morula extracts. Repeated analyses of the other elements as well as the nonconserved regions revealed no protection. The protein binding pattern showed distinct stage-specific patterns of complex formation in gel shift assays. Fig. 7A and B display the gel shift pattern for elements 'A' and 'B' using extracts from the gastrula and pluteus stages. The use of homologous (lanes 3 and 4) and heterologous (lane 5) competitor oligomers further indicates the specificity of the protein-DNA interactions.

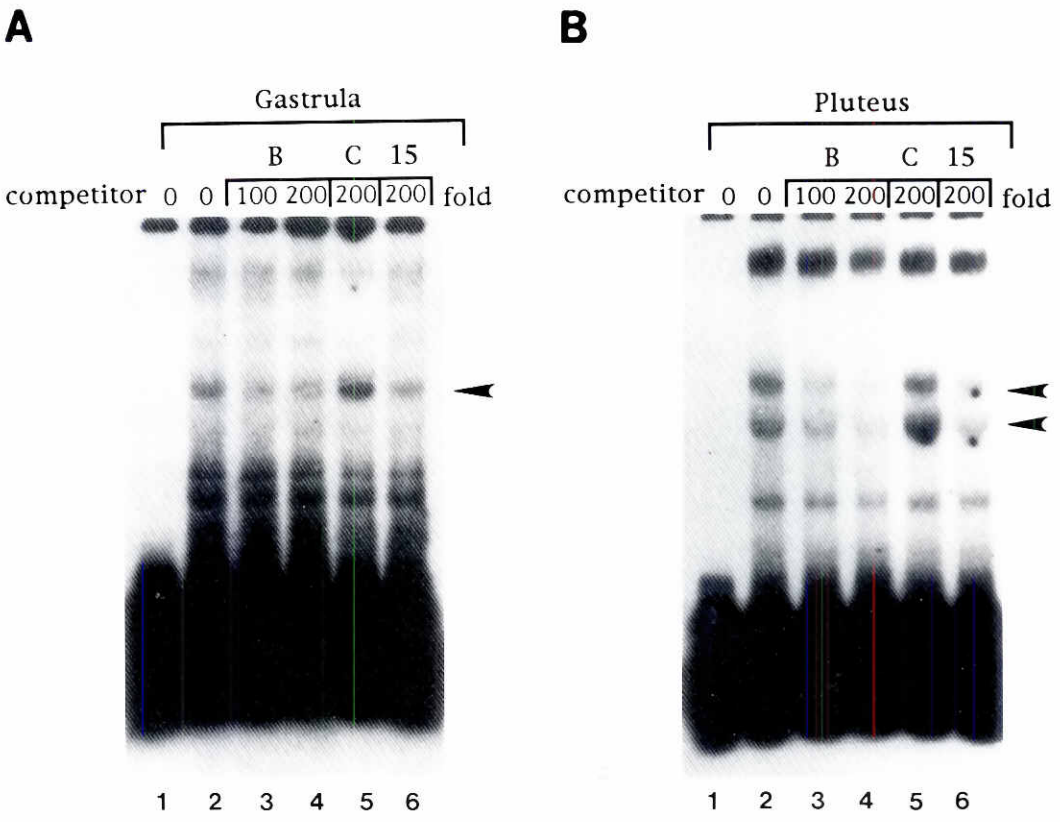

Fig. 4. Gel shift analysis of the 'B' region of msp130 gene using gastrula and pluteus stage nuclear extracts. *Oligomers corresponding to the protected areas were used for gel shift assays.* **(Panel A)** *Gel-shift competition experiment using the msp130 'B' oligomer and gastrula stage nuclear extract;* **(Panel B)** *gel-shift competition experiment using the msp130 'B' oligomer and pluteus stage nuclear extract. 5 μg of protein was used per lane. Lane 1 in panels A and B contains no nuclear extract. Lane 2 in panels A and B contains the probe incubated with the nuclear extract. Lanes 3 and 4 in panels A and B contain the indicated unlabeled homologous competitor oligomer. Lane 5 in panels A and B contains unlabeled heterologous oligomer. Lane 6 contains oligomer 15. The band shift complexes are indicated by arrows.*

Intergenic competition of A and B binding sites

msp130, SM50 and PM 27 genes appear to be coordinately expressed during skeletogenesis and have conserved putative regulatory sites. To test the possibility that these conserved elements might bind the same nuclear factors, we used a competitive gel shift assay (Fig. 3A lane 6). When oligomers containing the 'A' element of PM27 gene were used to compete with msp130 'A' probe, reproducible but incomplete competition was observed, even at high concentrations. In contrast, the msp130 element 'A' oligomer competed strongly for the PM27 element 'A' probe (Fig. 7A and B lanes 6 and 7). The footprint for PM27 element 'A' binding was correspondingly weaker than that of msp130, even though the two patterns were qualitatively similar. These results indicate that the element 'A' motifs for the two genes bind the same nuclear proteins, but with different affinities. The gel shift cross competition assay using the 'B' element oligomer from msp130 and PM27 showed no differences in their binding affinities (Fig. 4A and B lane 6). The specific bases important for strong interactions have not been determined.

Discussion

Embryonic cells appear to display two different strategies for cell fate determination: some cells, such as PMCs, can develop autonomously in isolation (Okazaki, 1975; Harkey and Whiteley, 1980) while others require cell-cell interaction (either suppressive or inductive response) for proper specification of territorial organization (Davidson, 1989; Henry *et al.*, 1989; Livingston and Wilt, 1990; Khaner and Wilt, 1991). Cells in which transcriptional activity is mediated by numerous cell-cell interactions are likely to have multiple interactions with different cis-regulatory sites and trans-acting factors that combinatorially control transcription (Davidson, 1990). The cis-regulatory elements of these genes can be very complex (Cameron *et al.*, 1989; Davidson, 1989) and structurally different (Thiebaud *et al.*, 1990). Such complexity may be required to accommodate the changes that occur during cell-cell interaction. The aboral ectoderm cells in *S. purpuratus* specifically express cytoplasmic actin CyIIIa (Calzone *et al.*, 1988; Theze *et al.*, 1990) and the Spec 1 and 2 gene families (Hardin *et al.*, 1988; Tomlinson *et al.*, 1990), which do not share any common regulatory sites, although they are coordinately expressed (Lynn *et al.*, 1983; Hickey *et al.*, 1987; Tomlinson and Klein, 1990). Therefore, genes expressed in the aboral ectoderm which have different promoter structures can nevertheless have identical spatial and temporal patterns of expression in embryos. A plausible hypothesis is that the activity of transcription factors is mediated by cell-cell interactions via distinct signalling pathways (Horvitz and Sternberg, 1991; Sternberg and Horvitz, 1991). Cells that can develop autonomously

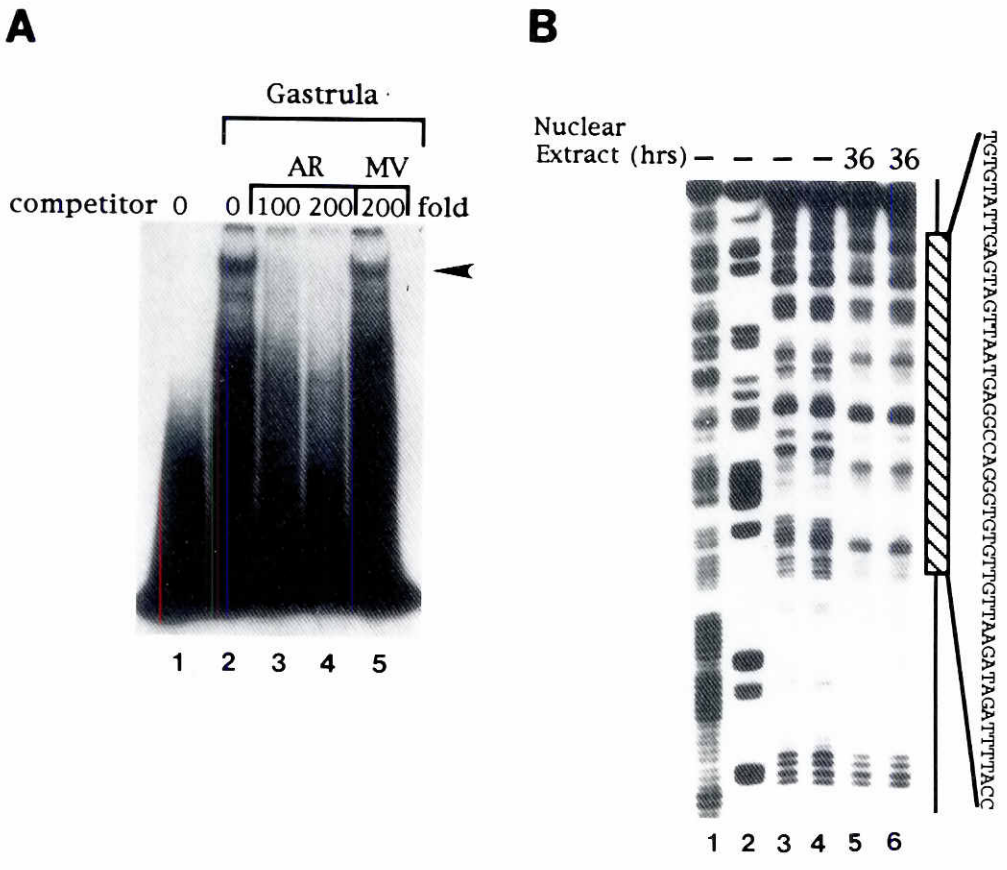

Fig. 5. *In vitro* protein-DNA interaction of the nuclear proteins to the 208 bp *Xba* I-*Hha* I 'AR' fragment of msp130 putative regulatory sites. (Panel A) *Gel shift assay. Each reaction used 5 μg of nuclear extract. Lane 1 contains no protein. Lane 2 contains the probe incubated with the nuclear extract. Lanes 3 and 4 contain the indicated amount of unlabeled homologous oligomers. Lane 5 contains unlabeled heterologous oligomer. The band shift complexes are indicated by arrows.* **(Panel B)** *DNase I footprinting analysis was carried out using the gastrula stage (36 h) nuclear extract. Lanes 1 (G+A) and 2 (G) are the Maxam and Gilbert sequencing products of the probe. Lanes 3 and 4 contain no nuclear extract. Lanes 5 and 6 show footprinting reactions with 20 μg of gastrula stage nuclear extracts and 1 unit of DNase I, respectively. After the incubation reaction, the samples were treated with DNaseI and analyzed by gel electrophoresis. The protected region designated as 'AR' is indicated by the hatched box. The DNase I footprint mapped to the same location when footprinted with either labeled strand.*

in isolation may have a simpler mechanism of regulation of histo-specific genes. The fate of PMCs is determined by the 16-cell stage. The PMCs must carry with them all the information required for the selective expression of the genes involved in skeletogenesis. Thus, a set of genes which appears to be coordinately expressed in early determined cells may share common cis-regulatory sequences and common trans-acting factors.

We found a distinct protein binding pattern for the msp130 gene, using extracts from morula, gastrula and pluteus stages of development in a binding assay. The different extracts represent developmental stages in which the genes are transcriptionally active or repressed. Although the nuclear protein extract from the morula stage supports *in vitro* transcription of an early histone gene, EH4 (Tung *et al.*, 1989), it does not bind to any of the elements. This is consistent with the activities of other PMC-expressed genes *in vivo* (Harkey *et al.*, 1988). The subsequent appearance of proteins that actively bind to the sequence elements suggests that stage-specific nuclear proteins are synthesized, activated or shifted from the cytoplasm to the nucleus during development.

The initial characterization of the putative regulatory sites by DNase I footprint experiments and gel shift assays indicates that the conserved elements bind nuclear proteins. As development progresses, the protein binding pattern to elements of the msp130 gene changes. Of the five elements ('A', 'B', 'CCAT', 'AR' and 'MV') protected by the gastrula stage extract, only three ('A', 'B', and 'CCAT') are protected by the pluteus stage extract. For the PM 27 gene, only the 'A' and 'B' elements are protected using the gastrula and the pluteus stage extracts. It is possible that the msp130 gene has regulatory sites important for its expression that are not used in transcription of the PM27 gene. Also, the trans-acting factors could act either as positive or negative regulatory molecules. A strong similarity to element P3A, which acts as a negative spatial element, exists in the putative regulatory sites of both msp130 and PM27 genes. A second negative spatial element, referred to as P7I, is also found in the regulatory sites of PM27. These sites potentially limit spatial expression in the embryos (Davidson, 1989). We do not detect binding to these sites with our nuclear extracts. This could be due either to the local sequence environment (Thiebaud *et al.*, 1990), or to a low concentration of the proteins that bind to these sites in our total embryo nuclear extracts.

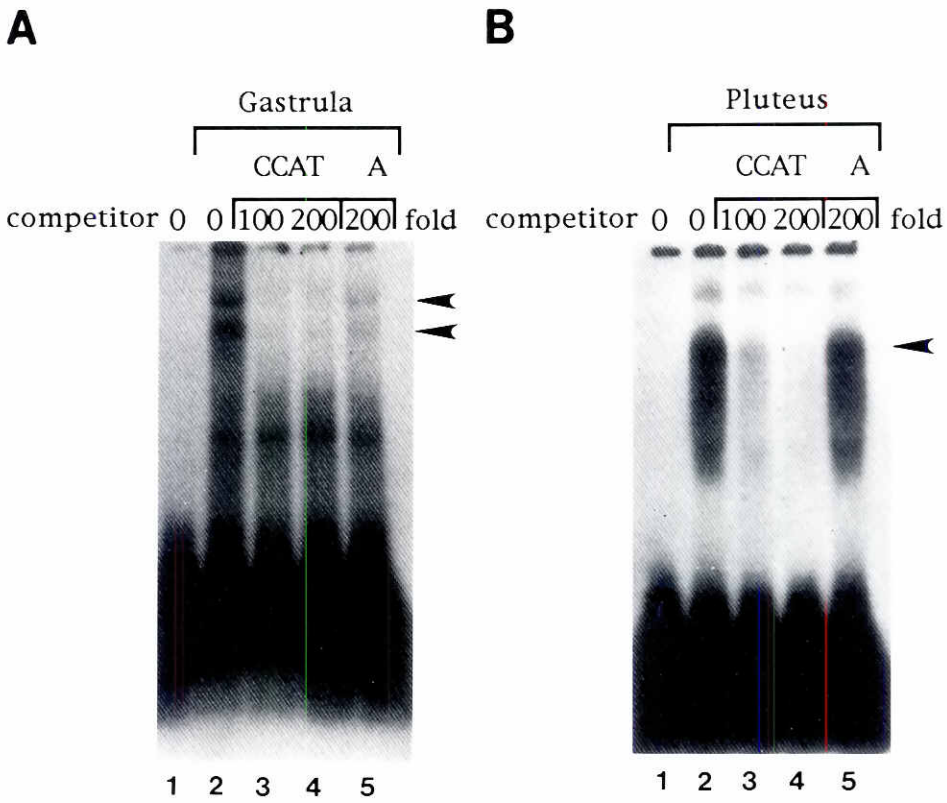

Fig. 6. Protein-DNA interaction studies of the 5' upstream region of msp130 gene (see Fig. 1). *The protected region is designated as 'CCAT' element. 5 μg of protein was used per lane.* **(Panel A)** *Gel shift assay for the interaction of the 'CCAT' element with gastrula stage extract.* **(Panel B)** *Gel shift assay for the interaction of the 'CCAT' element with pluteus stage extract. Lane 1 in panels A and B contains no protein. Lane 2 contains the probe incubated with nuclear extract. Lanes 3 and 4 in panels A and B contain unlabeled homologous oligomer. Lane 5 in panels A and B contains unlabeled heterologous oligomer. The band shift complexes are indicated by arrows.*

If the conserved elements 'A' and 'B' of both msp130 and PM27 genes are involved in coordinate regulation, the same proteins should be bound to both genes, although binding affinities need not be the same. Gel shift cross competition assays demonstrated that the 'A' and 'B' elements of PM27 bind the same proteins as their homolog counterparts in msp130. These experiments further indicate that the element 'A' from msp130 binds proteins more effectively than the PM27 oligomer. As PM27 is expressed at a lower level than msp130, it is possible that strong binding to element 'A' may be necessary for a higher rate of transcription. No difference in protein binding affinity was detected between the 'B' element of the different genes.

These studies of the common putative regulatory sites of msp130 and PM 27 genes suggest several conclusions. First, potential trans-acting factors appear progressively, indicating a possible role in the timing of gene expression. Second, two elements ('A' and 'B') which bind appropriate stage-specific nuclear proteins are found in all three sequences of PMC-expressed genes, consistent with the existence of a coordinate regulatory mechanism for these genes. Third, changes in protein binding patterns of these genes occur between gastrula and pluteus stages, which may be related to the switch from uniform to position-specific expression documented for msp130 for this interval. Finally, the difference in protein binding affinities for certain types of elements may be related to the different levels of expression of the genes. Transgenic experiments (Klueg *et al.*, in progress) have indicated that a 1.5 Kb region upstream from the transcription start site in both msp130 and PM27 genes is sufficient to drive the correct gene expression pattern. However, the role of the conserved cis-elements present among several PMC-expressed genes in controlling spatial and temporal patterns of expression has yet to be fully established.

Materials and Methods

Nuclear protein preparation

Gravid *S. purpuratus* were obtained from Marinus Inc. (Long Beach, CA, USA). Embryos were cultured in artificial seawater (ASW) at 15°C with constant stirring. For morula stage embryos, eggs were fertilized and embryos cultured in 1 mM para-aminobenzoic acid in ASW to allow removal of fertilization membranes from the embryos. Isolation of nuclear proteins followed a procedure similar to the modified protocol of Morris and Marzluff (1985). Total protein was estimated by the method of Bradford (1976) using bovine serum albumin as a standard.

Oligonucleotides

The following oligonucleotides were synthesized (along with the reverse complement) using an Applied Biosystems 394 DNA/RNA synthesizer. The purifications were carried out on a Mono Q column supplied by Pharmacia.

Element A: 5' TTATTTACACCCAATGCACACGTAAATGCGGCTTAT3'

Element B: 5' TACTCATCAAAGTTTGGGCCAATAGATATT 3'

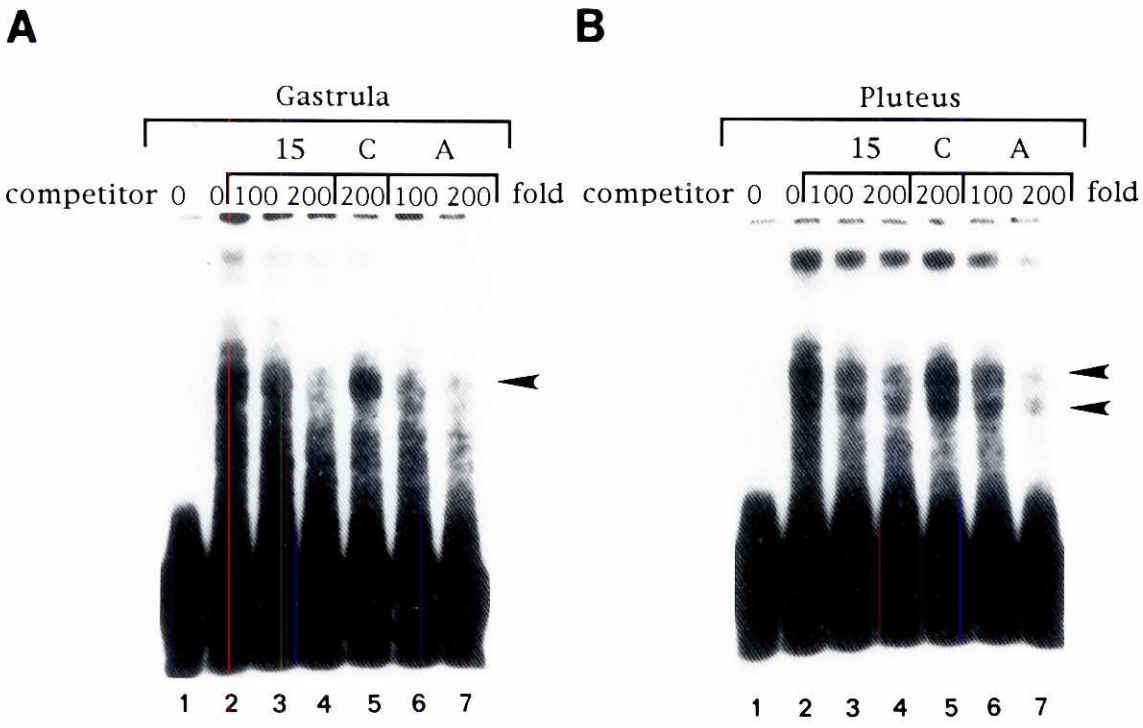

Fig. 7. Gel shift analyses of the 'AB' region of PM27 gene using gastrula and pluteus stage extracts. *5 μg of protein was used per lane. Oligomer 15 (containing both 'A' and 'B' regions of the PM27 gene) corresponding to the protected region was used to demonstrate the binding specificity of nuclear proteins.* **(Panel A)** *Gel-shift competition experiment using gastrula stage nuclear extract;* **(Panel B)** *gel-shift competition experiment using pluteus stage nuclear extract. Lane 1 in panels A and B contains no protein. Lane 2 in panels A and B contains the probe incubated with the nuclear extract. Lanes 3 and 4 in panels A and B contain unlabeled homologous oligomers. Lane 5 in panels A and B contains unlabeled heterologous oligomer. Lanes 6 and 7 in panels A and B contain the indicated amount of unlabeled 'A' oligomer corresponding to the msp130 gene. The band shift complexes are indicated by arrows.*

Element AR: 5' GGTAAAATCTATCTTAACAACACACCCTGG
CCTCATTAACTACTCAATACACA3'
Element MV: 5' TGTTGAGTAGTGTTTTCAACTCTAAAGAATT
CTTACAGACGCTCGGTAAATA3'
Oligomer 15: 5' TAAAACTTTGGACCAACTTTTGCTGCTATGTAC3'

DNase I footprinting assays

Restriction fragments were 5' end-labeled with Klenow enzyme. Markers were prepared by chemical sequencing (Maxam and Gilbert, 1980) of the respective probes. Assays were performed in a total volume of 20 μl. The protein binding reactions were performed on ice for a total of 30 minutes. Nonspecific binding was eliminated by incubating 18-20 μg of crude nuclear extract in 20 mM HEPES (pH 7.9), 50 mM KCl, 1 mM EDTA, 10 mM $MgCl_2$, 1μg of sonicated salmon sperm DNA (200-300 bp) and 6% glycerol for 10 minutes. One to two nanograms of the appropriate radioactively labeled DNA ($4\text{-}5x10^4$ cpm) were then added and incubated for 20 min. Subsequently, 1 Kunitz unit of DNase I (Promega, RQ1) was added and the reactions continued for a further 45 sec at room temperature. The reactions were stopped by adding 100 μl of 250 mM NaCl, 10 mM EDTA, 100 mM Tris (pH 8.0), 0.4% SDS, 100 μg/ml Proteinase K and incubating at 55°C for 30 min. The reactions were then extracted twice with phenol/chloroform and precipitated with 2 volumes of ethanol, washed twice with 70% ethanol, dried and loaded onto an 8% sequencing gel.

Gel shift assays

Protein binding reactions were carried out as described above for the footprinting assay, except that the buffer contained 4 μg of salmon sperm DNA, 4-5 μg of crude nuclear protein extract and 10 μg of BSA (Demczuk *et al.*, 1990). The oligonucleotide probe concentration was approximately 0.3 ng per reaction. The optimal concentration of extract preparation used was determined by titration experiments. The reactions were loaded onto 5% polyacrylamide gels (40:1 acrylamide/bis) in 0.25x TBE (Tris Borate EDTA buffer). The gels were run at 200 volts for 1-2 h in 0.25x TBE. For competition assays double stranded competitor DNA was added to the nuclear extract in binding buffer (see above) and incubated on ice for 10 min. Subsequently, the probe was added and the incubation continued for a further 20 min, and the samples were loaded onto non-denaturing gels.

Acknowledgments

We are grateful to Judith Jaehning for her suggestions and comments during the course of this work, as well as for critically reading the manuscript. This research was supported by NIH grant HD21337 to R.A.R.

References

ANSTROM, J.A., CHIN, J.E., LEAF, D.S., PARKS, A.L. and RAFF, R.A. (1987). Localization and expression of msp130, a primary mesenchyme lineage-specific surface protein of the sea urchin embryo. *Development 101:* 255-265.

BENSON, H.C., SUCOV, H., STEPHENS, L., DAVIDSON, E. and WILT, F. (1987). Lineage specific gene encoding a major matrix protein of the sea urchin embryo spicule. I. Authentication of the cloned gene and its developmental expression. *Dev. Biol. 120:* 499-506.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72:* 248-254.

CALZONE, F.J., THEZE, N., THIEBAUD, P., HILL, R.L., BRITTEN, R.J. and DAVIDSON, E.H. (1988). Developmental appearance of factors that bind specifically to cis-regulatory sequences of a gene expressed in the sea urchin embryo. *Genes Dev. 2:* 1074-1088.

CAMERON, R.A., BRITTEN, R.J. and DAVIDSON, E.H. (1989). Expression of two actin genes during larval development in the sea urchin *Strongylocentrotus purpuratus*. *Mol. Reprod. Dev. 1:* 149-155.

CARSON, D.D., FARACH, M.C., EARLES, D.S., DECKER, G.L. and LENNARZ, W.J. (1985). A monoclonal antibody inhibits calcium accumulation and skeleton formation in cultured embryonic cells of the sea urchin. *Cell 41:* 639-648.

DAVIDSON, E.H. (1989). Lineage-specific gene expression and the regulative capacities of the sea urchin embryo: a proposed mechanism. *Development 105:* 421-445.

DAVIDSON, E.H. (1990). How embryos work: a comparative view of diverse modes of cell fate specification. *Development 108:* 365-389.

DEMCZUK, S., DONOVAN, M., FRANKLIN, G. and OHLSSON, R. (1990). Order of probe and nuclear protein extract addition can determine specificity of protein-DNA complexes in tested mobility shift assays. *Nucleic Acids Res. 19:* 677-678.

HARDIN, P.E., ANGERER, L.M., HARDIN, S.H., ANGERER, R.C. and KLEIN, W.H. (1988). The Spec2 genes of *Strongylocentrotus purpuratus:* structure and differential expression in embryonic aboral ectoderm cells. *J. Mol. Biol. 202:* 417-431.

HARKEY, M.A. and WHITELEY, A.H. (1980). Isolation, culture and differentiation of echinoid primary mesenchyme cells. *Roux Arch. Dev. Biol. 189:* 111-122.

HARKEY, M.A. and WHITELEY, A.H. (1983). The program of protein synthesis during the development of the micromere-primary mesenchyme cell line in the sea urchin embryo. *Dev. Biol. 100:* 12-28.

HARKEY, M.A., WHITELEY, H.R. and WHITELEY, A.H. (1988). Coordinate accumulation of five transcripts in the primary mesenchyme during skeletogenesis in the sea urchin embryo. *Dev. Biol. 125:* 381-395.

HARKEY, M.A., WHITELEY, H.R. and WHITELEY, A.H. (1992). Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryo: a three dimensional *in situ* hybridization analysis. *Mech. Dev. 37:* 173-184.

HENRY, J.J., AMEMIYA, S., WRAY, G.A. and RAFF, A.R. (1989). Early inductive interactions are involved in restricting cell fates of mesomeres in sea urchin embryos. *Dev. Biol. 136:* 140-153.

HICKEY, R.J., BOSHAR, M.F. and CRAIN, W.R., Jr. (1987). Transcription of three actin genes and a repeated sequence in isolated nuclei of sea urchin embryos. *Dev. Biol. 124:* 215-225.

HORVITZ, H.R. and STERNBERG, P.W. (1991). Multiple intercellular signalling systems control the development of the *C. elegans* vulva. *Nature 351:* 535-541.

KHANER, O. and WILT, F.H. (1991). Interactions of different vegetal cells with mesomeres during early stages of sea urchin development. *Development 112:* 881-890.

LEAF, D.S., ANSTROM, J.A., CHIN, J.E., HARKEY, M.A., SHOWMAN, R.M. and RAFF, R.A. (1987). Antibodies to a fusion protein identify a cDNA clone encoding msp130, a primary mesenchyme cell surface protein of sea urchin embryo. *Dev. Biol. 121:* 29-40.

LIVINGSTON, B.T. and WILT, F.H. (1990). Determination of cell fate in sea urchin embryos. *BioEssays 12:* 115-119.

LYNN, D.A., ANGERER, L.M., BRUSKIN, A.M., KLEIN, W.H. and ANGERER, R.C. (1983). Localization of a family of mRNAs in a single cell type and its precursors in sea urchin embryos. *Proc. Natl. Acad. Sci. USA 80:* 2656-2660.

MAXAM, A. and GILBERT, W. (1980). Sequencing end-labelled DNA with base specific chemical cleavages. *Methods Enzymol. 65:* 499-560.

MORRIS, G.F. and MARZLUFF, W.F. (1985). Synthesis of UI RNA in isolated nuclei from sea urchin embryos: UI RNA is initiated at the first nucleotide of the RNA. *Mol. Cell. Biol. 5:* 1143-1150.

OKAZAKI, K. (1975). Spicule formation by isolated micromeres of the sea urchin embryo. *Am. Zool. 15:* 567-581.

PARR, B.A., PARKS, A.L. and RAFF, R.A. (1990). Promoter structure and protein sequence of msp130, a lipid-anchored sea urchin glycoprotein. *J. Biol. Chem. 265:* 1408-1413.

STERNBERG, P.W. and HORVITZ, H.R. (1991). Signal transduction during *C. elegans* vulval induction. *Trends Genet. 7:* 403-407.

SUCOV, H.M., BENSON, S.C., ROBINSON, J.J., BRITTEN, R.J., WILT, F. and DAVIDSON, E.H. (1987). A lineage-specific gene encoding a major matrix protein of the sea urchin embryo spicule. II. Structure of the gene and derived sequence of the protein. *Dev. Biol. 120:* 507-514.

SUCOV, H.M., HOUGH-EVANS, B.R., FRANKS, R.R., BRITTEN, R.J. and DAVIDSON, E.H. (1988). A regulatory domain that directs lineage-specific expression of a skeletal matrix protein gene in the sea urchin embryo. *Genes Dev. 2:* 1238-1250.

THEZE, N., CALZONE, F.J., THIEBAUD, P., HILL, R.L., BRITTEN, R.J. and DAVIDSON, E.H. (1990). Sequences of the CyIIIa actin gene regulatory domain bound specifically by sea urchin embryo nuclear proteins. *Mol. Reprod. Dev. 25:* 110-122.

THIEBAUD, P., GOODSTEIN, M., CALZONE, R.J., THEZE, N., BRITTEN, R.J. and DAVIDSON, E.H. (1990). Intersecting batteries of differentially expressed genes in the early sea urchin embryo. *Genes Dev. 4:* 1999-2010.

TOMLINSON, C.R. and KLEIN, W.H. (1990). Temporal and spatial transcriptional regulation of the aboral ectoderm-specific Spec genes during sea urchin embryogenesis. *Mol. Reprod. Dev. 25:* 328-338.

TOMLINSON, C.R., KOZLOWSKI, M.T. and KLEIN, W.H. (1990). Ectoderm nuclei from sea urchin embryos contain a Spec-DNA binding protein similar to the vertebrate transcription factor USF. *Development 110:* 259-272.

TUNG, L., MORRIS, G.F., YAGER, L.N. and WEINBERG, E.S. (1989). Sea urchin early and late H4 histone genes bind a specific transcription factor in a stable preinitiation complex. *Mol. Cell. Biol. 9:* 1476-1487.

Accepted for publication: October 1993