IDENTIFICATION AND CHARACTERIZATION OF A NEGATIVE REGULATOR REQUIRED FOR SPATIAL CONTROL OF THE TERRITORY-SPECIFIC *CyIIIa* GENE IN THE SEA URCHIN EMBRYO

Thesis by

David Guo-Wei Wang

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1995

(Submitted September 12, 1994)

This work is dedicated to my parents

who have taught me how to live in an uncertain world without fear.

ACKNOWLEDGMENTS

It is my pleasure to thank all the members of the Davidson lab for making my years at Caltech rewarding.

Particularly, I would like to thank the following people:

Eric Davidson, my adviser, for showing me the path to greater productivity and insightful intellectual inquiries.

Carmen Kirchhamer and Kirsten Rood, "my editors", for lending their creativity.

ABSTRACT

The *CyIIIa* cytoskeletal actin gene of the sea urchin *Strongylocentrotus purpuratus* is activated in late cleavage and expressed exclusively in the aboral ectoderm territory of the embryo. Previous gene transfer studies defined a 2.3 kb *cis*-regulatory domain that is necessary and sufficient for correct temporal and spatial expression of a *CyIIIa·CAT* fusion gene. In this study, a negative regulatory element within this region was identified that is required for repression of the *CyIIIa* gene in skeletogenic mesenchyme cells. The repression mediated by this element takes place after initial territorial specification.

To study negative spatial control of the *CyIIIa* gene at the molecular level, a cDNA clone encoding a DNA-binding protein with twelve Zn fingers (SpZ12-1) was isolated by probing an expression library with this *cis*-regulatory element. Deletion analysis of the SpZ12-1 protein confirmed that a DNA-binding domain is located within the zinc finger region. SpZ12-1 is the only DNA-binding protein in embryo nuclear extract that interacts with the specific *cis* target sites required for repression of *CyIIIa·CAT* in skeletogenic mesenchyme and is likely to be the *trans* factor that mediates this repression.

SpZ12-1 is present in significant quantities even in unfertilized egg cytoplasm, and in similar quantities in mesenchyme blastula-stage embryo cytoplasm. Taken together with earlier measurements of Calzone *et al.* (*Genes & Dev.* **2**,

1074-1088, 1988), our results indicate that SpZ12-1 enters the embryonic nuclei between late cleavage and mesenchyme blastula stages. A low prevalence of SpZ12-1 mRNA is also present throughout development. Translation of this mRNA could easily account for the complete complement of SpZ12-1 protein in the embryo, as estimated from its DNA binding activity. SpZ12-1 probably functions at several developmental stages, and is evidently of both maternal and embryonic provenance.

TABLE OF CONTENTS

INTRODUCTION

One of the central questions of early development is how differential gene activity is established in various regions of the embryo. To answer this question, well-defined embryonic systems are required in which the molecular mechanisms underlying differential gene expression are accessible to experimental manipulation. Since the sea urchin embryo represents a diverse group of invertebrate embryos with a well-characterized spatial organization, it has been an important model for studying basic questions about transcriptional control of spatially regulated genes (Davidson, 1989, 1990, 1991). The picture emerging from ongoing studies suggests that many of the *trans*-regulatory factors required for early specification are of maternal origin.

Embryonic Territories and Territorial Marker Genes

In the sea urchin *Strongylocentrotus purpuratus* the animal-vegetal axis is established in the unfertilized egg (Boveri, 1901), whereas the oral-aboral axis is specified by first cleavage (Cameron *et al.*, 1989). The pattern of cleavage in the undisturbed embryo is invariant and gives rise to five polyclonal territories (Cameron *et al.*, 1987; Cameron and Davidson, 1992). Each of these embryonic territories eventually develops into functional structures of the pluteus larva and can be identified by a unique spatial pattern of gene expression (Davidson, 1989; Cameron and Davidson, 1992). The territories are: the aboral ectoderm, the oral ectoderm, the skeletogenic mesenchyme, the vegetal plate and the small micromeres. A territory can be defined as a region of the early embryo that is derived from a contiguous group of founder cells. A founder cell is the

first cell to arise in any given region of the lineage and gives rise to a single territory. Figure 1 shows precise relationships among cell lineages, founder cells and embryonic territories in the sea urchin embryo. Most importantly, from embryo to embryo any given founder cell occupies a specific position in the cell lineage map and holds a specific location with respect to the axes in the cleavage stage (Cameron *et al.*, 1987, 1989, 1990).

During cleavage the lineage founder cells begin to display differential patterns of territory-specific gene expression. The correlation between the segregation of lineage founder cells and territory-specific gene expression suggests that founder cell specification is a fundamental cellular mechanism which establishes spatially defined gene expression in the sea urchin embryo (Davidson, 1989). Since the expression patterns of territory-specific genes are governed by *trans*-regulatory factors that specifically bind to the corresponding *cis*-regulatory regions (Davidson, 1986, 1989, 1990), the question of how differential gene expression is established at the transcriptional level is a key issue in understanding early development of the sea urchin embryo.

A number of territory-specific genes have been extensively characterized (Coffman and Davidson, 1992), which are activated immediately after the segregation of territorial founder cells from one another at the 3rd and 6th cleavage. For example, the *CyIIIa* and *CyIIIb* actin genes (Cox *et al.*, 1986), the arylsulfatase gene (Yang *et al.*, 1989), the metallothionein A gene (Nemer *et al.*,

1991) and the *Spec1* and *Spec2* genes encoding a family of calcium-binding proteins (Lynn *et al.*, 1983; Hardin *et al.*, 1988) are exclusively expressed in the aboral ectoderm. The *cis*-regulatory regions of some territorial marker genes have been defined *in vivo* and mapped to the level of specific protein-DNA interactions *in vitro*. The *CyIIIa* cytoskeletal actin gene is one of the best characterized aboral ectoderm-specific genes (Shott *et al.*, 1984; Cox *et al.*, 1986; Lee, 1986; Akhurst *et al.*, 1987). At least 10 *trans*-regulatory factors have been identified which participate in 20 different protein-DNA interactions within the *cis*-regulatory region, and they are required for correct temporal and spatial expression of this gene (Calzone *et al.*, 1988; 1991; Franks *et al.*, 1990; Hough-Evans *et al.*, 1990; Höög *et al.*, 1990; Char *et al.*, 1993; Kirchhamer *et al.*, unpublished; Coffman *et al.*, unpublished; Zeller *et al.*, unpublished; Xian *et al.*, unpublished; Chapter 1). For the *CyIIIb* actin gene, about five binding sites were identified in the 2.2 kb flanking sequence, which seem to include all the necessary elements for quantitative expression (Niemeyer and Flytzanis, 1993). Recent studies have revealed similar complexity of transcriptional regulation in two other territorial marker genes, *SM50* and *Endo16*. The *SM50* gene encodes a spicule matrix protein which is an early marker of the skeletogenic mesenchyme territory (Benson *et al.*, 1987; Sucov *et al.*, 1987) This gene is activated during late cleavage and expressed exclusively in the skeletogenic precursors, long before mesenchyme cell differentiation (Killian and Wilt, 1989). At least 20 different protein-DNA interactions within the *SM50 cis*-regulatory region have been mapped in detail by both *in vivo* and *in vitro* methods (Sucov *et*

al., 1988; Makabe *et al.*, unpublished). The *Endo16* gene encodes an extracellular protein which is initially expressed in the vegetal plate territory, and later on in the gut (Nocente-McGrath *et al.*, 1989; Ransick *et al.*, 1993) Highly complex protein-DNA interactions in the *cis*-regulatory region have been defined *in vitro* which include 38 binding sites organized in different modules and 14 *trans*-regulatory factors bound at the corresponding target sites (Yuh *et al.*, 1994). In summary, rapid and significant progress has been made in identifying spatially regulated genes and in defining their *cis*-regulatory regions. These studies have provided mechanistic insights into the problem of how different batteries of genes are regulated in a combinatorial manner (Thiebaud *et al.*, 1990). The challenge now is to dissect and understand the complex regulatory network that is required for establishing initial patterns of gene expression in the sea urchin embryo from these well-characterized *cis*-regulatory sequences.

A Biochemical Approach Towards Understanding Embryonic Gene Expression

Over the years biochemical methods have been developed which allow the isolation and characterization of large numbers of *trans*-regulatory factors required for territorial gene expression. This biochemical approach utilizes sequence-specific DNA affinity chromatography and protein microsequencing (Parker and Topol, 1984; Kadonaga and Tjian, 1986; Calzone *et al.*, 1991). A major experimental barrier for this approach in many systems has been the purification and identification of these DNA-binding proteins, since they are

generally present in extremely low levels in the cell. However, billions of synchronously developing sea urchin embryos are relatively easy to culture and process into large quantities of nuclear extracts which is essential for the purification of *trans*-regulatory factors (Calzone *et al.*, 1991). The recent development of automated sequential affinity chromatography for DNA-binding proteins has further simplified purification (Coffman *et al.*, 1992). An important advantage of DNA affinity chromatography is that multiprotein complexes on *cis*-regulatory elements can also be purified. Several factors have been copurified in this manner (J.A. Coffman, personal communication). At least 20 *trans*-regulatory factors from the sea urchin embryo have been purified and characterized by this method (Coffman and Davidson, 1992; J.A. Coffman, personal communication). In summary, this general approach offers experimental access to purification and cloning of most sea urchin *trans*-regulatory factors for which specific DNA-binding sites can be identified.

A Biochemical Estimation of the Complexity of Embryonic DNA-Binding Proteins

Recently, the complexity of DNA-binding proteins in early sea urchin embryo nuclei was estimated by using a quantitative two-dimensional (2-D) mapping technique and cation-exchange chromatography (Harrington *et al.*, 1992). Blastula stage nuclear extract was fractionated on a weak cation-exchange resin. This column permits the separation of proteins containing basic domains which are not necessarily basic overall, as measured by the isoelectric point (pI).

Basic domains are a general characteristic of specific DNA-binding proteins; it has been suggested that one-dimensional diffusion along DNA molecules may be mediated by the basic domain and greatly enhances the rate for targeting specific binding sites (von Hippel and Berg, 1989). A map of nuclear proteins containing basic domains was generated, as shown in Figure 1; this subset of proteins should include *trans*-regulatory factors. The proteins were selected on the basis that (*i*) they are at least 10-fold enriched in the basic domain fraction; (*ii*) they are present within the concentration range between approximately 250 to 7000 molecules per nucleus. Most known *trans*-regulatory factors which specifically bind to the *CyIIIa cis*-regulatory region have been located within this map (Harrington *et al.*, 1992; M.G. Harrington, personal communication). The map includes 265 spots probably representing about 100 different protein species, as shown in Figure 2. These results suggest that approximately 100 specific DNA-binding proteins are present in the early sea urchin embryo at sufficient concentration for function. This number is surprisingly close to the *Drosophila* estimate that has been implicated in the early embryo by both genetic and molecular methods (Levine and Harding, 1989; Johnston and Nüsslein-Volhard, 1992; Topol *et al.*, 1991; Biggin and Tjian, 1988).

Previous measurements of the total sequence complexity of mRNA populations indicate that approximately 8,500 genes are expressed in the early sea urchin embryo (Davidson, 1986). Thus, the number of expressed genes is about 100-fold higher than the estimated number of the *trans*-regulatory factors. This

strongly implies that transcriptional regulation of gene expression is accomplished through the combinatorial interactions of a relatively small number of *trans*-regulatory factors in the sea urchin embryo. Consistent with this notion, a detailed comparison of *cis*-regulatory regions in a number of territorial marker genes reveals that the same factors are often used in different combinations (Thiebaud *et al.*, 1990).

Trans-Regulatory Factors

Several *trans*-regulatory factors, summarized in Table 1, have been cloned and characterized. As discussed above, these factors each appear to be involved in regulating the expression of multiple territorial marker genes. The following discussions will focus on some of the well-characterized ones.

(1) Trans-regulatory factors that bind to the territory-specific CyIIIa gene SpP3A2 and SpZ2-1

SpP3A2 was purified from blastula stage nuclear extracts by affinity chromatography with P3A sites and partial peptide sequences were used to isolate a cDNA clone that encodes a 62 kD protein (Calzone *et al.*, 1991). Sequence analysis of SpP3A2 reveals that the protein has no significant similarities to any other proteins, suggesting that the SpP3A2 protein contains a novel type of DNA-binding domain (Calzone *et al.*, 1991). The P3A site is also present within the *cis*-regulatory regions of the skeletogenic mesenchyme-specific *SM50* and aboral ectoderm-specific *Spec1* genes (Thézé *et al.*, 1990;

Thiebaud *et al.*, 1990). It is important to note that another *trans*-regulatory factor, SpZ2-1, formerly called P3A1, binds the same target P3A site as the SpP3A2 protein. However, SpZ2-1 was isolated by screening a λgt11 expression library with a probe containing multiple P3A sites (Höög *et al.*, 1991). This 42 kd protein contains two zinc fingers which bind DNA. Careful examination of amino acid sequences of SpZ2-1 and SpP3A2 reveals that five specific amino acids in the first finger region and four in the second finger region of SpZ2-1 are also present in equivalent positions in SpP3A2. Since the two factors recognize the same target sites, it is possible to directly test the significance of these nine amino acids of SpP3A2 for the binding specificity.

A number of *in vivo* observations indicate that P3A sites are required for correct spatial expression of the *CyIIIa* gene in the sea urchin embryo (Hough-Evans *et al.*, 1990; Kirchhamer *et al.*, unpublished). First, *in vivo* competition with excess copies of the P3A sites results in ectopic expression of a *CyIIIa·CAT* fusion gene (Hough-Evans *et al.*, 1990). Second, recent studies confirms a similar ectopic expression pattern when the P3A sites in the 5' *CyIIIa* regulatory region is deleted (Kirchhamer *et al.*, unpublished). At present, it is not clear how the SpP3A2 and SpZ2-1 proteins function to repress *CyIIIa* gene expression in early development. Both SpP3A2 and SpZ2-1 proteins are of maternal origin, but they have different patterns of cytoplasm-nucleus distribution in the early embryo (Zeller *et al.*, unpublished; Calzone *et al.*, unpublished). The level of SpP3A2 in the nucleus remains fairly constant at approximately 1 X 10^4 molecules/nucleus

during early development. Unlike SpP3A2, the level of SpZ2-1 in the nucleus reaches a maximal level at the early cleavage stage with nearly 1 X 10^6 molecules/nucleus, and declines to 1 X 10^3 molecules/nucleus at the blastula stage. The rapid accumulation and disappearance of SpZ2-1 from the nucleus suggest that SpZ2-1 and SpP3A2 function as a regulatory switch in the early embryo (Zeller *et al.*, unpublished).

SpZ12-1

SpZ12-1 was isolated from a λZAP expression library with a DNA probe containing multiple SpZ12-1 binding sites, formerly called P6 sites (Chapter 1). The full-length SpZ12-1 cDNA (2.7 kb) includes an open reading frame of 1779 nucleotides which encods a protein of about 70 kd. A remarkable feature of this protein is that it contains 12 tandem zinc fingers of the TFIIIA type near the carboxyl terminus, suggesting that the zinc finger domain is responsible for the DNA-binding activity. This possibility has been confirmed by deletion experiments which show that the bacterially expressed protein without the zinc finger domain did not bind to the corresponding target sites (Chapter 1).

Gene transfer studies with a *CyIIIa·CAT* fusion gene containing altered SpZ12-1 binding sites demonstrated that SpZ12-1 binding sites are required for repression of the *CyIIIa* gene in skeletogenic mesenchyme cells, and this repression takes place after initial territorial specification (Chapter 1). The SpZ12-1 mRNA and protein are of maternal origin. The prevalence of maternal

SpZ12-1 protein is about 1.8 X 10^6 molecules per egg, whereas the prevalence of cytoplasmic SpZ12-1 protein at the blastula stage is approximately 1.7 X 10^6 molecules per embryo (Chapter 2). The maternal origin and cytoplasm-nucleus distribution of SpZ12-1 protein during early development suggest that SpZ12-1 enters the embryonic nuclei between late cleavage and mesenchyme blastula stages, and the nuclear localization of cytoplasmic SpZ12-1 is a potential control point for regulating its functional activity (Chapter 2).

SpOct-1

The SpOct-1 protein was purified by affinity chromatography from blastula stage nuclear extracts and the gene was identified by screening a λZAP cDNA library with a POU-domain probe that was isolated by PCR homology cloning (Char *et al.*, 1993). The SpOct cDNA encodes a 78 kd POU-domain protein related to the mammalian Oct-1 and Oct-2 proteins (Char *et al.*, 1993). The SpOct protein binds to octamer elements in the *cis*-regulatory regions of the aboral ectoderm-specific *CyIIIa* and α histone *H2B* genes (Thézé *et al.*, 1990; Bell *et al.*, 1992). RNA gel blots reveal four forms of SpOct mRNA, ranging in size from 4 to 12 kb (Char *et al.*, 1993). The temporal profile of SpOct mRNA expression closely resembles that of α histone *H2B* in the early embryo. Interestingly, recent experiments show that single-stranded SpOct-1 antisense oligonucleotides specifically block cleavage, and that injection of SpOct-1 mRNA overcomes this effect (Char *et al.*, 1994). The accumulation of newly

synthesized protein as well as DNA synthesis are significantly reduced in antisense-injected embryos (Char *et al.*, 1994).

Sp(G/C)F-1

Sp(G/C)F-1 was purified from blastula stage nuclear extracts by automated sequential affinity chromatography and a partial peptide sequence was used to isolate the corresponding cDNA (Zeller *et al.*, unpublished). The Sp(G/C)F-1 cDNA contains an open reading frame of 486 amino acids. The Sp(G/C)F-1 protein has no significant sequence similarity to any other known proteins. Sp(G/C)F-1 target sites are present in the *cis*-regulatory regions of several territorial marker genes such as *CyIIIa, SM50* and *Endo16*. It is rather surprising that there are five forms of Sp(G/C)F-1 produced from a single mRNA. All forms of the protein share a common carboxyl-terminal sequence, but differ in their amino-terminal sequences. Since the amino-terminus of Sp(G/C)F-1 is rich in prolines, serines and threonines which are commonly found in activation domains of transcription factors, the multiple forms of Sp(G/C)F-1 may therefore act as modulators for transcriptional activation in the early embryo (Zeller *et al.*, unpublished).

SpRunt

SpRunt was purified from blastula stage nuclear extracts by automated sequential affinity chromatography and partial peptide sequence was used to isolate the cDNA (Coffman *et al.*, unpublished). Sequence analysis of this

protein reveals that SpRunt has significant similarity with the product of the *Drosophila* gene *runt* and the human *AML1* gene which is associated with acute myeloid leukemia.

SpP7II

Peptide sequence from affinity-purified SpP7II was used to derive oligonucleotide probes for screening cDNA libraries (Coffman *et al.*, unpublished data). The SpP7II cDNA encodes a DNA-binding protein that has no significant similarities to any known proteins. Interestingly, there are at least three different forms of SpP7II that are produced by alternatively spliced transcripts from a single gene. They differ in their carboxyl-termini (Coffman *et al.*, unpublished).

Earlier *in vivo* competition experiments demonstrate that the SpP7II binding sites are required for correct spatial expression of the *CyIIIa* gene (Hough-Evans *et al.*, 1990). Recent studies with a *CyIIIa·CAT* fusion gene in which the SpP7II binding site was deleted showed that the SpP7II binding sites are responsible for repression of the *CyIIIa* gene in the oral ectoderm and mesenchyme cells (Kirchhamer *et al.*, unpublished). The SpP7II protein is maternal and present in the cytoplasm of blastula stage embryos (J.A. Coffman, personal communication).

(2) Trans-regulatory factors that bind to the territory-specific Spec genes

SpOtx

SpOtx was identified by screening a cDNA library with an *orthodenticle* (otd) related homeodomain probe that was isolated by PCR homology cloning (W.H. Klein, personal communication). A positive *cis*-element, designated the A/T palindrome, was identified in the 5' *Spec2a* regulatory region (Gan and Klein, 1993). This element contains the core recognition sequence TAATCC which is the high affinity binding site for the SpOtx protein, the sea urchin homologue of the *Drosophila* orthodenticle (Gan *et al.*, 1994). Other proteins interacting with this element have yet to be identified.

SpF1/suUSF

SpF1/suUSF was isolated by cross-hybridization with a human USF cDNA probe which contained the helix-loop-helix (HLH) DNA-binding domain (Kozlowski *et al.*, 1991). Sequence analysis of the SpF1/suUSF protein revealed a high degree of similarity in the basic-HLH domain to the human USF protein and only eight residues were not identical (Kozlowski *et al.*, 1991). The SpF1/suUSF binding sites are present in the 5' regulatory regions of the aboral ectoderm-specific *Spec* genes including *Spec1, Spec2a* and *Spec2c.* SpF1/suUSF binding activity was detectable only in ectoderm cells of the early embryo (Tomlinson *et al.*, 1990). However, mutations in the SpF1/suUSF binding sites within the Spec fusion genes do not alter expression. These sites may be required for proper regulation of the endogenous Spec genes (Tomlinson *et al.*, 1990).

(3) Other trans-regulatory factors of interest

SpCOUP-TF

The *SpCOUP-TF* gene was identified by cross-hybridization with a human COUP-TF cDNA probe (Chan *et al.*, 1992). The SpCOUP-TF binding site, formerly called C1R, is located in the 5' regulatory region of the aboral ectoderm-specific *CyIIIb* gene (Niemeyer *et al.*, 1992). Detailed comparison of the SpCOUP-TF protein sequence with the human COUP-TF revealed some interesting features. Within the DNA-binding and ligand-binding domains the identity of amino acid sequences is 96% and 92%, respectively (Chan *et al.*, 1992). The extensive similarity of the ligand-binding domains strongly suggests that the SpCOUP-TF factor is activated by a similar ligand or a similar signal transduction pathway. The SpCOUP-TF mRNA and protein are maternal and present at low levels during early development (Chan *et al.*, 1992; Calzone *et al.*, unpublished data). Identification of the corresponding ligand or signal transduction pathway would be of importance in elucidating its biological function during sea urchin embryogenesis.

SSAP

SSAP was purified by affinity chromatography from blastula stage nuclear extracts (DeAngelo *et al.*, 1993). This 43 kd protein binds to the stage-specific USE IV enhancer of the *H1*-β histone gene. Before the enhancer is active in early development, SSAP appears as a 43 kd monomer, but it undergoes a change in its molecular weight to approximately 90 kd beginning at the early

blastula stage. This correlates precisely with an increase in *H1*-β gene expression, suggesting that the higher molecular weight form of SSAP activates the blastula stage-specific transcription of the late *H1* gene (DeAngelo *et al.*, 1993).

Maternal Regulatory Factors and Control of Territory-Specific Gene Expression

A substantial number of *trans*-regulatory factors required for expression of territory-specific genes are of maternal origin (Calzone *et al.*, unpublished; Chapter 2). For instance, all the *CylIIa* regulatory factors are present in the egg cytoplasm. Importantly, most of the *CylIIa* factors are also present in large quantities in the cytoplasm of blastula stage embryos (Calzone *et al.*, unpublished; Chapter 2). The maternal origin and cytoplasm-nucleus distribution of these *trans*-regulatory factors imply that relocalization from cytoplasm to nucleus is a major control point for regulating their functional activities. One possible mechanism is that factors are sequestered in the cytoplasm and become translocated into the nucleus upon activation, perhaps mediated by signal transduction (Davidson, 1989). Such a molecular mechanism is well known in a number of different systems, for instance, *dorsal* in *Drosophila* (Rushlow *et al.*, 1989; Steward, 1989), and NF-κB in mammals (Baeuerle and Baltimore, 1988). In the case of NF-κB, the protein factor is present in the cytoplasm of nonstimulated cells that do not express immunoglobulin kappa light chain. Upon stimulation, the factor becomes localized to the nucleus and

exhibits DNA-binding properties. The relocalization is mediated by the action of a protein kinase. This general scheme is consistent with the regulative abilities of the sea urchin embryo that has been extensively documented in the work of Hörstadius and others (Hörstadius, 1971; Davidson, 1989). For instance, recent experiments showed that micromeres transplanted onto the animal cap of 8- or 16-cell embryos induce the mesomeres (which normally give rise to the ectoderm) to develop into gut expressing endoderm-specific genes (Ransick and Davidson, 1993). In summary, development in the undisturbed embryo might initially be controlled by the asymmetric distribution of maternal regulatory factors in the egg, which are partitioned differentially among the cleaving blastomeres, however, the regulative capacities demonstrated by transplantation and other microsurgical experiments suggest that at least some of these factors are present globally in the early blastomeres, but locally activated by short-range intercellular interactions (Davidson, 1989, 1990, 1993). The role of interblastomere interactions in lineage founder cell specification has been clearly demonstrated by physical rearrangment of interblastomere contacts, laser ablation and mutations in *C. elegans*. For instance, specification of the gut founder cell (Goldstein, 1992), certain body wall muscle cells (Schnabel, 1994), and right/left asymmetry (Wood, 1991) is conditional. Similarly, in *C. elegans* the maternal gene *skn-1* encodes a transcription factor that is required for specification of certain posterior founder cells to produce pharyngeal cells (Bowerman *et al.*, 1993). The Skn-1 protein is of maternal origin and the nuclear localization of the protein is distributed unequally in the early embryo. A major

mechanism underlying early spatial specification in *C. elegans* and sea urchins is modification of maternal transcription factors controlled by regional intercellular interactions (Davidson, 1990). Recently, postranslational modification of some *trans*-regulatory factors in the early sea urchin embryo has been demonstrated by using a quantitative two-dimensional electrophoretic method and immunostaining (M.G. Harrington and D.G. Wang, unpublished data), as shown in Figure 3. It is now possible to investigate the developmental regulation of their functional activities *in vivo*.

Figure 1. A cell lineage diagram from the 8-cell stage to the 64-cell stage, illustrating the segregation of founder cells for the five territories. The horizontal axis is divided into cleavage cycles independent of time. The boxes corresponding to blastomeres are gray until segregation to a single territory occurs; then the boxes are coded with the color of the territory. The 11 founder cells for the aboral ectoderm arise at the third cleavage (Na), fifth cleavage (right and left NL1l, right and left NL2l) and the sixth cleavage (VAM1u, VAM2u, right and left VLM1u, and right and left VLM2u). The seven founder cells of the oral ectoderm (No, right and left NL1u, right and left NL2u,VOM1u, VOM2u) originate in a similar way. The founder cells for the vegetal plate territory arise at the sixth cleavage, when they are segregated from ectoderm precursors (VAM1l, VAM2l, right and left VLM1l, right and left VLM2l, VOM1l, VOM2l). The primary mesenchyme founder cells are segregated from the small micromere founder cells at the fifth unequal division in the micromere lineage. Thus the five territories are completely segregated by the sixth cleavage, or about two divisions before the transcriptional activation of many territorial marker genes. The five territories consist of one or more morphologically identifiable structures in the pluteus stage embryo. Each structure in turn consists of one or more cell types, as indicated in the right boxes. For example, the aboral ectoderm consists of only squamous epithelium while the vegetal plate is made up of several cell types.

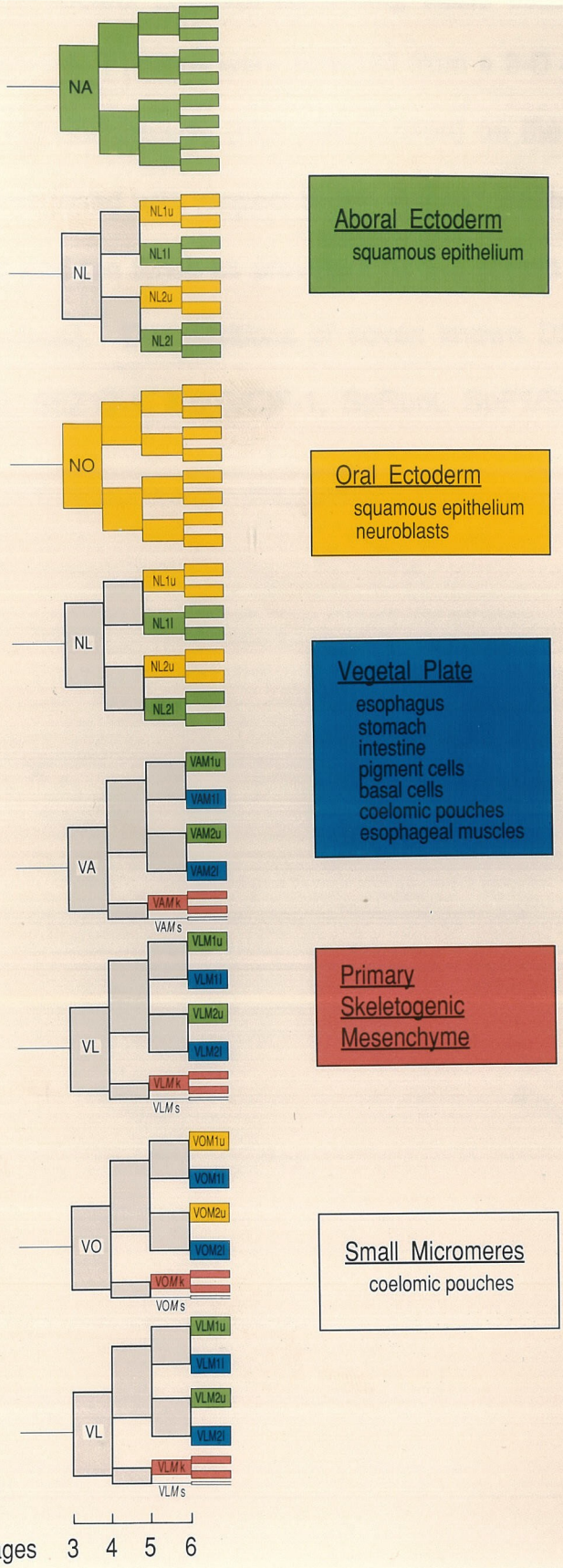

NA
NL
NL1u
NL1l
NL2u
NL2l
NO
NL
NL1u
NL1l
NL2u
NL2l
VA
VAM1u
VAM1l
VAM2u
VAM2l
VAMk
VAMs
VL
VLM1u
VLM1l
VLM2u
VLM2l
VLMk
VLMs
VO
VOM1u
VOM1l
VOM2u
VOM2l
VOMk
VOMs
VL
VLM1u
VLM1l
VLM2u
VLM2l
VLMk
VLMs
Cleavages 3 4 5 6
Aboral Ectoderm
squamous epithelium
Oral Ectoderm
squamous epithelium
neuroblasts
Vegetal Plate
esophagus
stomach
intestine
pigment cells
basal cells
coelomic pouches
esophageal muscles
Primary
Skeletogenic
Mesenchyme
Small Micromeres
coelomic pouches

Figure 2. Map of nuclear proteins containing basic domains (250 to 7,000 molecules per nucleus). Proteins were selected from a 2-D electrophoretic map of blastula stage nuclear extracts (high salt fraction) on the basis that (*i*) they are about 10-fold enriched with respect to the low salt fraction and (*ii*) they are between one third and ten times as prevalent as the known factor SpZ2-1 (700 molecules per nucleus). The positions of seven known DNA-binding proteins (SpZ2-1, SpP3A2, SpZ12-1, Sp(G/C)F-1, SpRunt, SpF1/SuUSF and P5) are indicated.

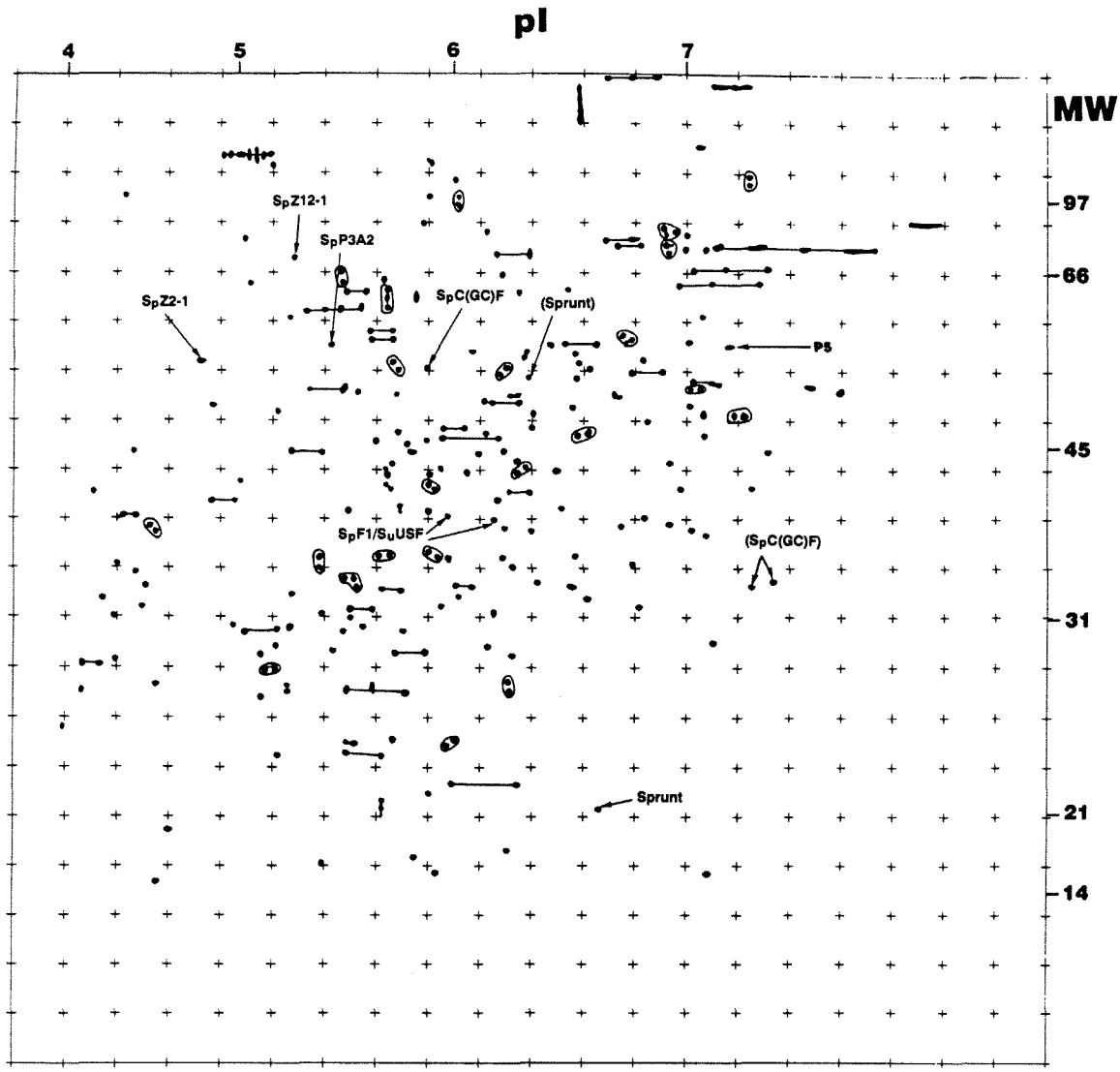
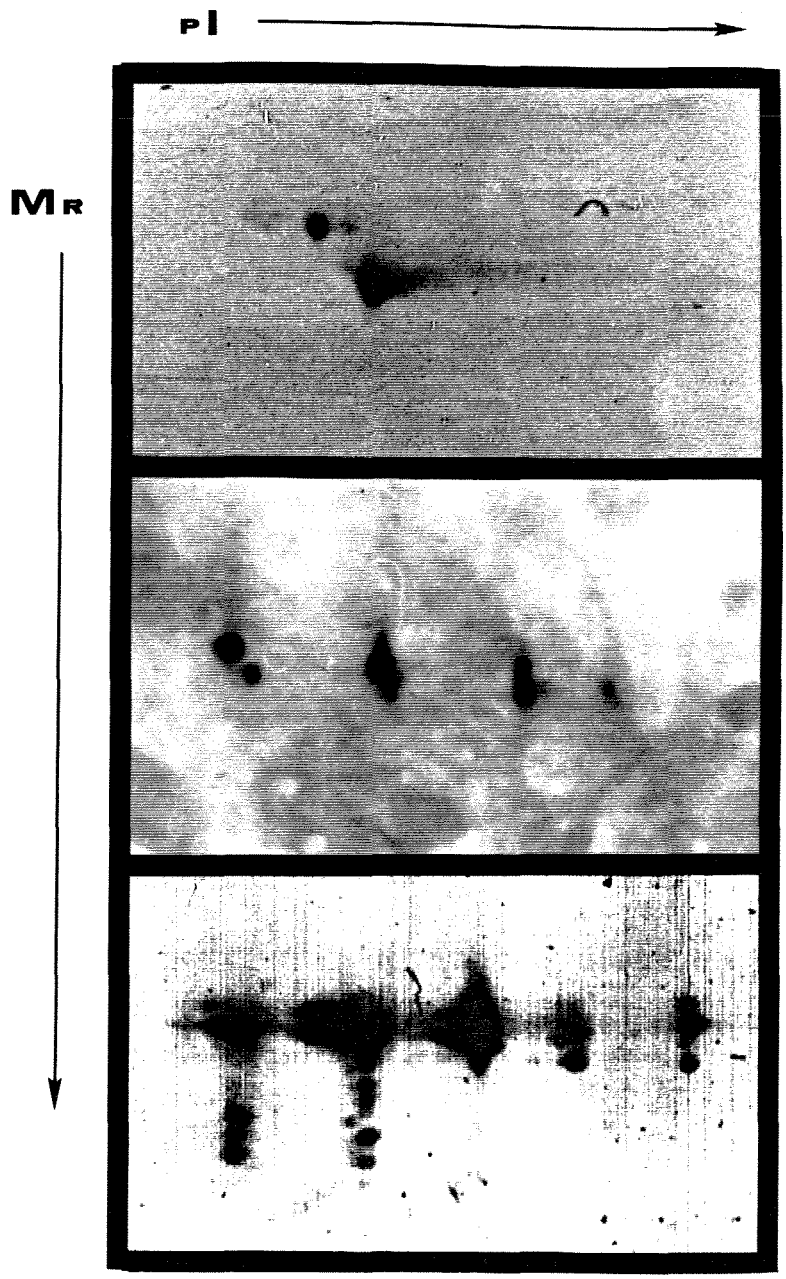

pI
4
5
6
7
MW
97
66
45
31
21
14
SpZ12-1
SpP3A2
SpZ2-1
SpC(GC)F
(Sprunt)
P5
SpF1/SuUSF
(SpC(GC)F)
Sprunt

Table 1. *Trans*-regulatory factors in the sea urchin embryo

Factor	Class	Target promoter(s)	Function(s)
SpZ2-1	ZF	*CyIIIa*	Negative Spatial
SpP3A2	New	*CyIIIa* *SM50* *Spec1*	Negative Spatial
SpZ12-1	ZF	*CyIIIa*	Negative Spatial
SpCOUP-TF	ZF/SR	*CyIIIb*	N.D.
SUM-1	bHLH	N.D.	Activator
SSAP	RNP	Late *H1*	Positive Temporal
ABP-1/*Hbox4*	Hbox	Late *H2B*	Activator
SpF1/*suUSF*	bHLH	*Spec1* *Spec2a* *Spec2c*	N.D.
SpOct-1	POU III	*CyIIIa* Early *H2B*	Activator
Sp(G/C)F-1	New	*CyIIIa* *SM50* *Endo16*	Activator
SpRunt	Runt	*CyIIIa*	Positive Temporal
SpP7II	New	*CyIIIa* *Spec1*	Negative Spatial
SpOtx	Hbox	*Spec1* *Spec2a* *Spec2c*	Activator

Abbreviations include: ZF–zinc finger; SR–steroid finger; bHLH–basic helix-loop-helix; RNP–RNA binding protein; Hbox–homeobox; Runt–runt domain; N.D.–not identified.

Figure 3. Postranslational modification of SpZ12-1, SpF1/*suUSF* and SpP3A2 in the early sea urchin embryo. Gels loaded with whole embryo lysate of the blastula stage were blotted to membranes, stained with colloidal gold to visualize the spot pattern of total protein, and then allowed to react with antibodies specific to SpZ12-1, SpF1/*suUSF* and SpP3A2. The immunoreactive components were located by using a chemiluminescent detection method. A number of variants of SpZ12-1, SpF1/*suUSF* and SpP3A2 are revealed in the upper, middle and lower panels. These variants appear to be the consequence of postranslational modifications (M.G. Harrington and D.G. Wang, unpublished data).


pI
MR

REFERENCES

Akhurst, R.J., F.J. Calzone, J.J. Lee, R.J. Britten and E.H. Davidson. 1987. Structure and organization of the *CylII* actin gene subfamily of the sea urchin *Strongylocentrotus purpuratus*. *J. Mol. Biol.* **194**: 193-203.

Baeuerle, P.A. and D. Baltimore. 1988. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor. *Cell* **53**: 211-217.

Bell, J., R.B. Char and R. Maxson. 1992. An octamer element is required for the expression of the alpha H2B gene during the early development of the sea urchin. *Dev. Biol.* **150**: 363-371.

Benson, S., H. Sucov, L. Stephens, E.H. Davidson and F. Wilt. 1987. A 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.

Biggin, M.D. and R. Tjian. 1988. Transcriptional factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell* **53**: 699-711.

Boveri, T. 1901. Die Polarität von Oocyte, Ei, und Larve des *Strongylocentrotus lividus*. *Zoologische Jahrbücher, Anatomie und Ontogenie der Tiere, Jena.* **14**: 630-653.

Bowerman, B., B.W. Draper, C.C. Mello and J.R. Priess. 1993. The maternal gene *skn-1* encodes a protein that is distributed unequally in early *C. elegans* embryos. *Cell* **74**: 443-452.

Calzone, F.J., N. Thézé, P. Thiebaud, R.L. Hill, R.J. Britten and E.H. Davidson. 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.

Calzone, F.J., C. Höög, D.B. Teplow, A.E. Cutting, R.W. Zeller, R.J. Britten and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo I. Purification by affinity chromatography and cloning of P3A2, a novel DNA-binding protein. *Development* **112**: 335-350.

Cameron, R.A., B.R. Hough-Evans, R.J. Britten and E.H. Davidson. 1987. Lineage and fate of each blastomere of the eight-cell sea urchin embryo. *Genes Dev.* **1**: 75-85.

Cameron, R.A., S.E. Fraser, R.J. Britten and E.H. Davidson. 1989. The oral-aboral axis of a sea urchin embryo is specified by first cleavage. *Development* **106**: 641-647.

Cameron, R.A., S.E. Fraser, R.J. Britten and E.H. Davidson. 1990. Segregation of oral from aboral ectoderm blastomeres is completed at the fifth cleavage in the embryogenesis of *Strongylocentrotus purpuratus. Dev. Biol.* **137**: 77-85.

Cameron, R.A. and E.H. Davidson. 1991. Cell type specification during sea urchin development. *Trends Genet.* **7**: 212-218.

Chan, S.-M., N. Xu, C.C. Niemeyer, J.R. Bone and C.N. Flytzanis. 1992. *SpCOUP-TF*: a sea urchin member of the steroid/thyroid hormone receptor family. *Proc. Natl. Acad. Sci.* **89**: 10568-10572.

Char, B.R., J.R. Bell, J. Dovala, J.A. Coffman, M.G. Harrington, J.C. Becerra, E.H.Davidson, F.J. Calzone and R. Maxson. 1993. *SpOct*, a gene encoding the major octamer-binding protein in sea urchin embryos: expression profile, evolutionary relationships, and DNA binding of expressed protein. *Dev. Biol.* **158**: 350-363.

Char, B.R., H. Tan and R. Maxson. 1994. A POU gene required for early cleavage and protein accumulation in the sea urchin embryo. *Development* **120**: 1929-1935.

Coffman, J.A. and E.H. Davidson. 1992. Expression of spatially regulated genes in the sea urchin embryo. *Current Opinion Genet. & Dev.* **2**: 260-268.

Coffman, J.A., J.G. Moore, F.J. Calzone, R.J. Britten, L.E. Hood and E.H. Davidson. 1992. Automated sequential affinity chromatography of sea urchin embryo DNA binding proteins. *Mol. Marine Biol. Biotech.* **1**: 136-146.

Cox, K.H., L.M. Angerer, J.J. Lee, R.J. Britten, E.H. Davidson and R.C. Angerer. 1986. Cell lineage-specific programs of expression of multiple actin genes during sea urchin embryogenesis. *J. Mol. Biol.* **188**: 159-172.

Cutting, A.E., C. Höög, F.J. Calzone, R.J. Britten and E.H. Davidson. 1990. Rare maternal mRNAs code for regulatory proteins that control lineage-specific gene expression in the sea urchin embryo. *Proc. Natl. Acad. Sci.* **87**: 7953-7957.

Davidson, E.H. 1986. "Gene Activity in Early Development" 3rd ed., Academic Press, Orlando, FL.

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.

Davidson, E.H. 1991. Spatial mechanisms of gene regulation in metazoan embryos. *Development* **113**: 1-26.

DeAngelo, D.J., J. DeFalco and G. Childs. 1993. Purification and characterization of the stage-specific embryonic enhancer-binding protein SSAP-1. *Mol. Cell. Bio.* **13**: 1746-1758.

Franks, R.R., R. Anderson, J.G. Moore, B.R. Hough-Evans, R.J. Britten and E.H. Davidson. 1990. Competitive titration in living sea urchin embryos of regulatory factors required for expression of the CyIIIa actin gene. *Development* **110**: 31-40.

Gan, L. and W.H. Klein. 1993. A positive *cis*-regulatory element with a bicoid target site lies within the sea urchin *Spec2a* enhancer. *Dev. Biol.* **157**: 119-132.

Gan, L., C.-A. Mao., L.M. Angerer, R.A. Angerer, W.M. Perry and W.H. Klein. 1994. Target gene and expression patterns of an orthodenticle-related homeobox protein from *Strongylocentrotus purpuratus*. In press.

Goldstein, B. 1992. Induction of gut in *Caenorhabditis elegans* embryos. *Nature* **357**: 255-257.

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

Harrington, M.G., J.A. Coffman, F.J. Calzone, L.E. Hood, R.J. Britten and E.H. Davidson. 1992. Complexity of sea urchin embryo nuclear proteins that contain basic domains. *Proc. Natl. Acad. Sci.* **89**: 6252-6256.

Höög, C., F.J. Calzone, A.E. Cutting, R.J. Britten and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo II. Two dissimilar proteins, P3A1 and P3A2, bind to the same target sites that are required for early territorial gene expression. *Development* **112**: 351-364.

Hörstadius, S. 1973. Experimental embryology of echinoderms. Oxford: Clarendon Press.

Hough-Evans, B.R., R.R. Franks, R.W. Zeller, R.J. Britten and E.H. Davidson. 1990. Negative spatial regulation of the lineage-specific *CyIIIa* actin gene in the sea urchin embryo. *Development* **110**: 41-50.

Johnston, D. and C. Nüsslein-Volhard. 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**: 201-219.

Kadonaga, J.T. and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci.* **83**: 5889-5893.

Killian, C.E. and F.H. Wilt. 1989. Accumulation and translation of a spicule matrix protein mRNA during sea urchin embryo development. *Dev. Biol.* **133**: 148-156.

Kozlowski, M.T., L. Gan, J.M. Venuti, M. Sawadogo and W.H. Klein. 1991. Sea urchin USF: a helix-loop-helix protein active in embryonic ectoderm cells. *Dev. Biol.* **148**: 625-630.

Lee, J.J. 1986. The genomic organization and expression of the *Strongylocentrotus purpuratus* actin gene family. Ph.D. thesis, California Institute of Technology, Pasadena, California.

Levine, M.S. and K.W. Harding. 1989. *Drosophila*: the zygotic contribution. In: *Genes and Embryos*, (Chapter 2), D.M. Glover and B.D. Hames, Eds., New York: IRL Press, pp. 39-94.

Lynn, D.A., L.M. Angerer, A.M. Bruskin, W.H. Klein and R.C. Angerer. 1983. Localization of a family of mRNAs in a single cell type and its precursors in sea urchin embryos. *Proc. Natl. Acad. Sci.* **80**: 2656-2660.

Nemer, M., R.D. Thornton, E.W. Steabing and P. Harlow. 1991. Structure, spatial and temporal expression of two sea urchin metallothionein genes, *SpMT13* and *SpMTA*. *J. Biol. Chem.* **266**: 6586-6593.

Niemeyer, C.C. 1992. Ph.D. Thesis (Baylor College of Medicine, Houston, Texas).

Niemeyer, C.C. and C.N. Flytzanis. 1993. Upstream elements involved in the embryonic regulation of the sea urchin *CyIIIb* actin gene: temporal and spatial specific interactions at a single *cis*-acting element. *Dev. Biol.* **156**: 293-302.

Nocente-McGrath, C.A. Brenner and S.G. Ernst. 1989. Endo16, a lineage-specific protein of the sea urchin embryo, is first expressed just prior to gastrulation. *Dev. Biol.* **136**: 264-272.

Parker, C.S. and J. Topol. 1984. A *drosophila* RNA polymerase II transcription factor binds to the regulatory site of an *hsp 70* gene. *Cell* **37**: 273-283.

Ransick, A. and E.H. Davidson. 1993. A complete second gut induced by transplanted micromeres in the sea urchin embryo. *Science* **259**: 1134-1138.

Ransick, A., S.G. Ernst, R.J. Britten and E.H. Davidson. 1993. Whole mount *in situ* hybridization shows *Endo16* to be a marker for the vegetal plate territory in sea urchin embryos. *Mech. Dev.* **42**: 117-124.

Rushlow, C.A., K. Han, J.L. Manley and M. Levine. 1989. The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in *Drosophila*. *Cell* **59**: 1165-1177.

Schnabel, R. 1994. Autonomy and nonautonomy in cell fate specification of muscle in the *Caenorhabditis elegans* embryo: a reciprocal induction. *Science* **263**: 1449-1452.

Shott, R.J., J.J. Lee, R.J. Britten and E.H. Davidson. 1984. Differential expression of the actin gene family of *Strongylocentrotus purpuratus*. *Dev. Biol.* **101**: 295-306.

Steward, R. 1989. Relocalization of the dorsal protein from the cytoplasm to the nucleus correlates with its function. *Cell* **59**: 1179-1188.

Sucov, H.M., S. Benson, J.J. Robinson, R.J. Britten, F. Wilt and E.H. Davidson. 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**: 507-519.

Sucov, H.M., B.R. Hough-Evans, R.R. Franks, R.J. Britten and E.H. Davidson. 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.

Thézé, N., F.J. Calzone, P. Thiebaud, R. L. Hill, R.J. Britten and E.H. Davidson. 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., M. Goodstein, F.J. Calzone, N. Thézé., R.J. Britten and E.H. Davidson. 1990. Intersecting batteries of differentially expressed genes in the early sea urchin embryo. *Genes Dev.* **4**: 1999-2010.

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

Topol, J., C.R. Dearolf, K. Prakash and C.S. Parker. 1991. Synthetic oligonucleotides recreate *Drosophila fushi tarazu* zebra-stripe expression. *Genes Dev.* **5**: 855-867.

von Hippel, P.H. and O.G. Berg. 1989. Facilitated target location in biological systems. *J. Biol. Chem.* **264**: 675-678.

Wood, W.B. 1991. Evidence from reversal of handedness in *C. elegans* embryos for early cell interactions determining cell fates. *Nature* **349**: 536-538.

Yang, Q., L.M. Angerer and R.C. Angerer. 1989. Structure and tissue-specific developmental expression of a sea urchin arylsulfatase gene. *Dev. Biol.* **135**: 53-65.

Yuh, C.-H., A. Ransick, P. Martinez, R.J. Britten and E.H. Davidson. 1994. Complexity and organization of DNA-protein interactions in the 5'-regulatory region of an endoderm-specific marker gene in the sea urchin embryo. *Mech. Dev.* in press.

CHAPTER 1

SpZ12-1, a Negative Regulator Required for Spatial Control of the Territory-Specific *CyIIIa* Gene in the Sea Urchin Embryo

SpZ12-1, a negative regulator required for spatial control of the territory-specific *CyIIIa* gene in the sea urchin embryo

David G.-W. Wang, Carmen V. Kirchhamer, Roy J. Britten and Eric H. Davidson

Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125

The *CyIIIa* cytoskeletal actin gene of the sea urchin *Strongylocentrotus purpuratus* is activated in late cleavage and expressed exclusively in the aboral ectoderm territory of the embryo. Previous gene transfer studies defined a 2.3 kb *cis*-regulatory region that is necessary and sufficient for correct temporal and spatial expression of a *CyIIIa·CAT* fusion gene. In this paper, a negative regulatory element within this region was identified that is required for repression of the *CyIIIa* gene in skeletogenic mesenchyme cells. The repression mediated by this regulatory element takes place after initial territorial specification. A cDNA clone encoding a DNA-binding protein with twelve Zn fingers (SpZ12-1) was isolated by probing an expression library with this *cis*-element. Deletion analysis of the SpZ12-1 protein confirmed that a DNA-binding domain is located within the Zn finger region. SpZ12-1 is the only DNA-binding protein in embryo nuclear extract that interacts with the specific *cis* target sites required for repression of *CyIIIa·CAT* in skeletogenic mesenchyme and is likely to be the *trans* factor that mediates this repression.

Introduction

In the sea urchin embryo cleavage of blastomeres is invariant and gives rise to five polyclonal territories. Each of these embryonic territories develops into specific larval structures and those that have been studied are identified by unique spatial patterns of gene expression. The territories are: the aboral ectoderm, the oral ectoderm, the skeletogenic mesenchyme, the vegetal plate and the small micromeres. Territory-specific gene expression begins shortly after the segregation of lineage founder cells from one another at 3rd to 6th cleavage (Davidson, 1986, 1989, Cameron and Davidson, 1991).

A well-characterized molecular marker for the aboral ectoderm is the *CyIIIa* cytoskeletal actin gene, which is transcriptionally activated late in cleavage and is expressed exclusively in the aboral ectoderm throughout embryonic development (Shott et al., 1984; Cox et al., 1986; Lee et al., 1986; Akhurst et al., 1987). It was shown that a 2.3 kb *cis*-regulatory region located upstream of the *CyIIIa* transcription start site is necessary and sufficient to mediate correct temporal and spatial expression of a *CyIIIa·CAT* fusion gene (Flytzanis et al., 1987; Hough-Evans et al., 1987). At least 20 sites within this region where nuclear proteins bind with high specificity in vitro have been mapped in detail, as indicated in Figure 1 (Calzone et al., 1988; Thézé et al., 1990). Several gene regulatory factors that interact at these sites with high specificity have been isolated and characterized (Calzone et al., 1991; Höög et al., 1991; Char et al., 1993; Coffman et al., unpublished; Zeller et al., unpublished, Xian et al., unpublished). This report concerns the regulatory function of interactions that occur at two adjacent sites within the *CyIIIa* regulatory domain, and a DNA-binding factor which binds specifically at these sites (SpZ12-1). By mutating the SpZ12-1 sites, we showed that interactions at these sites are

required specifically for spatial repression of the *CyIIIa* gene expression in skeletogenic mesenchyme cells at the gastrula stage. SpZ12-1 is the only factor interacting with this regulatory element in embryo nuclear extracts, and thus it is likely that it is the mesenchyme repressor for the *CyIIIa* gene.

Results

Expression pattern of the control **CyIIIa·CAT** *fusion gene at late gastrula stage*

Expression of the *CyIIIa* cytoskeletal actin gene is confined to the aboral ectoderm territory of the early embryo (Cox et al., 1986). *CyIIIa* transcripts are first detected in the late cleavage stage (Hickey et al., 1987), immediately after segregation of the aboral ectoderm lineage founder cells (Cameron et al., 1987, 1990). In earlier gene transfer experiments, a 2.3 kb *CyIIIa* regulatory region was found to direct correct spatial and temporal expression of a linked CAT reporter gene in the embryo (Flytzanis et al., 1987; Livant et al., 1988, 1991; Hough-Evans et al., 1987, 1988, 1990; Franks et al., 1990). A detailed description of the 5' *CyIIIa* regulatory region is shown in Figure 1. In previous experiments using this fusion construct the spatial distribution of CAT mRNA had been detected by section in situ hybridization, while in this study we utilized a sensitive whole-mount in situ hybridization procedure to locate CAT transcripts in the embryo(Ransick et al., 1993).

The expression pattern of the complete *CyIIIa·CAT* fusion gene at late gastrula stage (54 hr postfertilization), as visualized by the whole-mount hybridization procedure, is shown in Figure 2. Figure 2A displays an embryo with aboral ectoderm expression close to the lower aboral-oral boundary; in Figures 2B and 2C two embryos are shown with a patch of positive cells toward the upper

aboral-oral boundary. Figures 2D-2F display three embryos with a large portion of the aboral ectoderm stained. The mosaic pattern of CAT expression observed in these stained embryos (Fig. 2A-2F) is due to mosaic incorporation of the fusion gene during the first several cleavages (Hough-Evans et al., 1988; Livant et al., 1991). Of the hundreds of embryos expressing the *CyIIIa·CAT* fusion gene that were examined, approximately 95% showed expression confined to the aboral ectoderm. Quantitative data for four different experiments with the control *CyIIIa·CAT* construct are given in Tables 1 and 2.

Ectopic expression pattern of the **mZ12CyIIIa·CAT** *fusion gene at late gastrula stage*

Two adjacent SpZ12-1 binding sites were initially identified by Calzone et al. (1988; there referred to as "P6" sites). A preliminary experiment showed that deletion of a region of about 200 bp (-1033 to -821), which includes these sites, sharply derepresses spatial regulation of the *CyIIIa* gene, causing ectopic expression in mesenchyme cells (data not shown). To assess the contribution of the SpZ12-1 binding sites to the spatial control of *CyIIIa* gene expression, both sites within the 5' *CyIIIa* regulatory region of a *CyIIIa·CAT* fusion gene were mutated, generating the *mZ12CyIIIa·CAT* fusion gene. As shown in Figure 3A, these sites occur in reverse orientation 13 base pairs apart. They include a 9/10 nucleotide pair reverse repeat of the sequence TGTTGC(T/C)AGGT. By site-directed mutagenesis we altered each of the nucleotides in this core sequence from purine to pyrimidine or vice versa. As shown in Figure 3B, these substitutions actively abolish DNA-binding activity (Fig. 3B).

Patterns of expression of the *mZ12CyIIIa·CAT* fusion gene at late gastrula stage were determined by whole-mount in situ hybridization as above. The mutation

has a remarkable effect on spatial expression of the fusion gene, evidently resulting in its derepression in mesenchyme cells, but only in mesenchyme cells. Examples are shown in Figure 4. Two major phenotypes occur, *viz.* embryos displaying expression in both mesenchyme and aboral ectoderm cells (Fig. 4A, 4B, 4D, 4E, and 4F), and embryos with expression only in mesenchyme cells (Fig. 4C). A detailed analysis of about 200 stained embryos is summarized in Table 1. Approximately 70% of all the injected embryos express CAT transcripts, and this is true for both the *CylIIa·CAT* and *mZ12CylIIa·CAT* fusion genes. About 46% of embryos bearing the *mZ12CylIIa·CAT* fusion display ectopic mesenchyme cell expression (i.e., 8-fold higher than the low level of ectopic expression observed in the control *CylIIa·CAT* embryos). Importantly, ectopic expression is confined to skeletogenic mesenchyme, rather than secondary mesenchyme, as judged from the position of the stained cells at this stage of embryogenesis. These data indicate that a negative regulatory interaction mediated by the SpZ12-1 binding sites is required for repression of the *CylIIa* gene in mesenchyme cells at the late gastrula stage.

Expression pattern of the **mZ12CylIIa·CAT** *fusion gene at mesenchyme blastula stage*

We wished to determine whether the negative control of *CylIIa* expression in mesenchyme cells that is mediated by the SpZ12-1 sites is part of the initial aboral ectoderm specification mechanism. Thus we examined the expression of the *mZ12CylIIa·CAT* construct at earlier stages. At the mesenchyme blastula stage (30 hr postfertilization), the skeletogenic mesenchyme cells have just ingressed into the blastocoel and can be easily identified morphologically, though at this stage it is difficult to distinguish aboral from oral ectoderm. Surprisingly, in these experiments we observed CAT transcripts exclusively in

the ectoderm in almost all of the embryos staining for CAT RNA. No ectopic expression was detected in mesenchyme cells at this stage, in contrast to the late gastrula stage, as just discussed. Three typical embryos displaying CAT expression in the ectoderm but not in the newly ingressed skeletogenic mesenchyme cells are shown in Figure 5. As indicated in Table 2, only 3% of the stained embryos from this series of experiments show ectopic expression in mesenchyme cells, which is indistinguishable from the control *CyIIIa·CAT* level. It should be noted that the possibility of ectopic expression in the oral ectoderm at this stage cannot be excluded, even though it is unlikely (i.e., since no oral expression is seen in gastrula stage embryos bearing the *mZ12CyIIIa·CAT* fusion). Thus the repression of mesenchyme expression mediated by the SpZ12-1 binding sites is a later control function, that is required to maintain ectodermal expression of the *CyIIIa* gene after gastrulation.

Molecular cloning of the SpZ12-1 gene

A cDNA clone, designated SpZ12-1, was isolated by screening a 4 hr sea urchin embryo λZAP expression library with an oligonucleotide probe containing multiple SpZ12-1 binding sites (Vinson et al., 1988; Höög et al., 1991). Nucleotide sequence analysis of a full-length SpZ12-1 cDNA clone indicated an open reading frame of 1779 bp (Fig. 6). This reading frame encodes a 593 amino acid protein, the predicated molecular weight of which is about 70 kd.

The deduced amino acid sequence of the SpZ12-1 protein reveals a potential DNA-binding domain consisting of twelve tandem Zn fingers (Fig. 6A). The finger domain is located near the carboxyl terminus and spans more than 50% of the entire protein. All twelve fingers of the SpZ12-1 protein are variants of the TFIIIA type of Zn finger, the consensus amino acid sequence of which is

Tyr/Phe-X-Cys-$X_{2/4}$-Cys-X_3-Phe-X_5-Leu-X_2-His-X_3-His-X_5, where X represents a relatively variable amino acid, and each finger folds independently (Berg et al., 1990; Pavletich et al., 1991, 1993). In each case, the cysteines are separated by two amino acids and each finger is connected with the next by a Thr-Gly-Glu-Lys-Pro consensus linker. Each of these sequences is capable of binding a zinc ion via the cysteine and histidine side chains to generate structural configurations that are essential for DNA binding (Berg et al., 1990). Figure 6B shows sequence alignment of SpZ12-1 and *Drosophila Krüppel* Zn fingers (Rosenberg et al., 1986). These two finger proteins share all the consensus amino acids of the TFIIIA type of Zn finger, but also display a great deal of similarity in the region represented by the terminal "X_5" in this canonical consensus (see shading in Fig. 6A). Computer searches of Genbank with the protein sequence outside of the Zn finger domain did not reveal significant similarities to any other known proteins.

Characterization of the SpZ12-1 protein

Since the sequence analysis of the SpZ12-1 protein revealed twelve Zn fingers near the carboxyl terminus, it is likely that a potential DNA-binding domain is located within this region. To test this possibility, gel shift assays were performed to compare the DNA-binding activity of the wild-type and two deletion mutants of the SpZ12-1 protein. DNA fragments encoding the full-length sequence (SpZ12-1), the amino-terminal region (SpZ12-1ΔZF), and the Zn finger domain (SpZ12-1ΔN) of the SpZ12-1 protein were cloned into pET expression vectors (Studier et al., 1986), as shown in Figure 7A. The encoded proteins were expressed and purified from *Escherichia coli*, and after renaturation (see Materials and methods) the proteins were tested for their ability to bind to the SpZ12-1 binding sites of the 5' *CyIIIa* regulatory region by gel shift

assay. Results are shown in Figure 7B. The SpZ12-1 and SpZ12-1ΔN proteins bound to the ^{32}P-labeled DNA probe equally well. However, no binding whatsoever was observed when the SpZ12-1ΔZF protein was tested. These results clearly demonstrate that the DNA-binding domain of the SpZ12-1 protein is located within the Zn finger region. Figure 7C shows an experiment demonstrating that Zn ions are required for the DNA-binding activity of the SpZ12-1 protein. DNA binding is abolished by dialyzing purified SpZ12-1ΔN protein with a low concentration of EDTA, and restored when $ZnCl_2$ is added back to the reaction mixture.

Using bacterially expressed, full-length SpZ12-1 protein, we further defined the SpZ12-1 binding sites within the 5' *CyIIIa cis*-regulatory region by DNase I footprinting experiments. As shown in Figure 8A, the SpZ12-1 protein protects the two adjacent regions indicated as SpZ12-1 binding sites I and II. Alignment of the protected sequences of the sites in Figure 8B shows that they include the inversely repeated consensus motif TGTTGC(T/C)AGG detected earlier (Thézé et al., 1990).

Identification of the SpZ12-1 protein in the early embryo

Since the SpZ12-1 clone was originally obtained by direct ligand screening, it is essential to determine whether the protein it encodes is present in the early embryo. A polyclonal antibody was raised against the recombinant SpZ12-1ΔZF protein, which lacks the Zn finger domain (Fig. 7). In Figure 9A the specificity of the antibody was confirmed by reaction with both blastula stage nuclear extracts and the full-length bacterially expressed SpZ12-1 protein. Only one immunoreactive band was detected with blastula stage nuclear extracts, which is identical in size to the SpZ12-1 protein that is encoded by the cloned cDNA. Gel

shift assays were then performed, using blastula stage nuclear extracts and a ^{32}P-labeled SpZ12-1 binding site probe. The specific SpZ12-1 protein-DNA complex was isolated from the gel; the complexed protein was released by treatment with SDS; and it was then displayed by SDS gel electrophoresis. This gel was blotted, and the filter reacted with the anti-SpZ12-1 antibody, as just described above. Figure 9B shows that the authentic immunoreactive component of the DNA-binding complex that derived from the embryo extract is identical in size to the full-length bacterially expressed SpZ12-1 protein. Furthermore, no other proteins binding to these specific sites could be recovered from the embryo extracts. Therefore SpZ12-1 is in all probability the DNA-binding factor that interacts with the target sites in the *CyIIIa* gene that we studied, and that is responsible for the spatial repression function we identified.

Discussion

Identification of a negative cis-element within the 5' CyIIIa regulatory region

Previous studies suggest that the *CyIIIa* gene is under negative spatial control in the early embryo (Franks et al., 1988; 1990; Hough-Evans et al., 1990). The first such evidence came from an interspecific experiment (Franks et al., 1988) in which ectopic spatial expression was seen when a *CyIIIa·CAT* fusion gene was injected into eggs of a distant second species, *L. variegatus.* This phenomenon implies that the normal spatial regulation of the *CyIIIa* gene is at least in part negative. Ectopic spatial expression can also be induced in *S. purpuratus* by in vivo competition. When excess amounts of two particular binding sites, SpP3A2 and SpP7II (see Figure 1), are coinjected with the *CyIIIa·CAT* construct, ectopic expression of CAT transcripts is also detected (Hough-Evans et al., 1990). As expected, if the interactions at these sites are of negative import the competition

for the factors binding these two target sites did not significantly decrease the overall level of expression (Franks et al., 1990). In contrast, competition for the target sites of a number of other positively acting factors depresses the level of expression stoichiometrically, while the normal spatial pattern of expression remains unchanged (Franks et al., 1990; Hough-Evans et al., 1990). Recent studies have confirmed that the SpP3A2 and SpP7II binding sites are in fact required for negative spatial control of the *CyIIIa* gene (C. V. Kirchhamer, unpublished data). However, when the SpZ12-1 binding sites were coinjected in molar excess with the *CyIIIa·CAT* construct in these in vivo competition experiments, neither ectopic expression nor decrease in CAT activity was observed (Franks et al., 1990; Hough-Evans et al., 1990). This could have been due to the relatively high prevalence of SpZ12-1 protein (Calzone et al., 1988; Wang et al., unpublished), which could have made it difficult to obtain an effective excess of competitor sites in vivo. In the present study, when the SpZ12-1 binding sites were altered by site-directed mutagenesis, approximately 46% of embryos bearing this fusion gene showed ectopic expression, specifically in mesenchyme cells, at the late gastrula stage (Table 1). This percentage directly reflects the probability of incorporation of the fusion gene into mesenchyme cell lineages during the first several cleavages (Hough-Evans et al., 1988; Sucov et al., 1988). Thus, other recent experiments carried out with a skeletogenic mesenchyme-specific construct (*SM50·CAT*) demonstrate that maximal percentage of embryos with stained mesenchyme cells for a single injection is about 30-40% (K.W. Makabe, unpublished data). Our results imply that all embryos that have incorporated the mutant fusion gene in skeletogenic mesenchyme cell precursors will express CAT transcripts in these cells. We conclude that the SpZ12-1 binding sites are a third negative spatial control element that is required to repress mesenchyme cell expression of the *CyIIIa* gene.

Late effect: repression of the CyIIIa gene in skeletogenic mesenchyme cells

The results from embryos bearing the *mZ12CyIIIa·CAT* fusion gene at the mesenchyme blastula stage (Fig. 5 and Table 2), together with the data from the late gastrula stage (Fig 4 and Table 1) clearly show that ectopic mesenchyme cell expression of this fusion gene occurs long after mesenchyme cell specification. It is important to note that there are several skeletogenic mesenchyme marker genes, which are activated at late cleavage stage, transcripts of which rapidly accumulate in the primary mesenchyme cells as ingression occurs (Benson et al., 1987; Killian et al., 1989). The skeletogenic mesenchyme is probably autonomously specified during cleavage (Davidson, 1989). Since ectopic expression of CAT transcripts was not detected in primary mesenchyme cells in embryos bearing the *mZ12CyIIIa·CAT* fusion gene, and not even at the stage when the cells invaginate into the blastocoel, the negative function of the SpZ12-1 sites for the *CyIIIa* gene is therefore not part of the initial autonomous specification program and appears to take place only after gastrulation. A plausible interpretation of the function of this negative control system is that during the postgastrula period the *CyIIIa* gene begins to respond to a positive regulator that is active in both aboral ectoderm and skeletogenic mesenchyme, so the role of the repressive interaction we have studied would be to confine the effect of this positive regulator to the aboral ectoderm.

SpZ12-1 is probably the particular transcription factor that exercises the negative control function we have assigned to its target site

We show in Figure 9B that an immunoreactive sea urchin nuclear protein is released from complexes formed with the SpZ12-1 site probe, which is identical in size to the SpZ12-1 protein encoded by our cDNA clone. No other proteins

that react with the anti-SpZ12-1 antibody are seen (Figure 9). Of course we cannot exclude the possibility that some other protein was also present in the complexes formed with the SpZ12-1 target probe, or that different proteins interact at this site in vivo. Nonetheless, it is not unlikely that SpZ12-1 is the factor that services these sites in the *CyIIIa* gene in vivo. The binding activity of the SpZ12-1 protein is highly specific, as shown in the DNase I footprinting experiments (Figure 8). Recently, it has been shown that SpZ12-1 is of maternal origin, and this factor no doubt has other functions in oogenesis or early development (Wang et al., unpublished; Anderson et al., 1994).

The SpZ12-1 protein

Examination of the SpZ12-1 protein sequence reveals twelve tandem Zn fingers of TFIIIA type near the carboxyl terminus (Berg et al., 1990). This type of Zn finger is one of the most common DNA binding motifs among eukaryotic transcription factors, and can recognize a diverse set of DNA sequences (Pellegrino et al., 1991; Pavletich et al., 1993). The similarities between the finger motif of SpZ12-1 and *Drosophila* segmentation gene *Krüppel* is striking (Fig. 6B). *Krüppel* represents a subclass of Zn finger genes of the TFIIIA type that is characterized by multiple Zn fingers in which the canonical sequence motif $CX_2CX_3FX_5LX_2HX_3H$ is separated by highly conserved 7-amino acid linkers of the form TGEKP(Y/F)X (Chowdhury et al., 1987). Recent studies of two crystal structures of the Zn finger proteins, Egr 1 and GLI-1 with their cognate DNA-binding sites, show that each Zn finger domain can independently bind DNA and that each recognizes a 3-nucleotide sequence; in the case of GLI-1, not all of the five Zn fingers contact the DNA; i.e., not all Zn fingers contribute equally to DNA recognition (Pavletich et al., 1991; 1993). Consistent with this observation, our data demonstrate that 12 Zn finger region of SpZ12-1

recognizes only a 23-nucleotide target site (Figure 8B), implying that probably four to five of the Zn fingers are not involved in the interaction displayed by the footprint at each site. The SpZ12-1 protein sequence does not have significant similarities to any other proteins outside the zinc finger domain. For instance, despite the fact that the fingers of SpZ12-1 and *Krüppel* are similar, the amino terminal region of SpZ12-1 has no similarity with the highly conserved *Krüppel*-associated box (KRAB) which is a 75-amino acid domain present in the amino terminal region of more than one-third of all *Krüppel* subclass Zn finger proteins and which is probably responsible for transcriptional repression mediated by this factor (Margolin et al., 1994; Witzgall et al., 1994).

SpZ12-1 contains many potential sites for phosphorylation, but the significance of this is hard to evaluate due to the preponderance of serine, threonine and proline residues in the sequence. Our recent studies have revealed that a number of SpZ12-1 variants are present in the early embryo which can be resolved by two-dimensional gel electrophoresis. These variants appear to be the consequence of postranslational modifications (D.G. Wang and M.G. Harrington, unpublished observations). Since SpZ12-1 is maternal and is present in both cytoplasm and nucleus of the early embryo (Wang et al., unpublished), such postranscriptional modifications could be the means of controlling the nuclear localization of SpZ12-1, or its regulatory function in the nucleus.

In conclusion, we have identified a specific negative spatial *cis*-regulatory function that controls *CyIIIa* expression in the skeletogenic mesenchyme of late embryos. The *trans*-factor which exercises this function is probably SpZ12-1. Since we know the cell type and stage in which this function of SpZ12-1 is

required, we are now in a position to determine the modifications which determine its temporal and spatial activity.

Materials and methods

Embryos

Eggs of *Strongylocentrotus purpuratus* were prepared for injection as previously described (McMahon et al., 1985). The eggs were microinjected with ~2 pl of injection solution after fertilization. The embryos were cultured to mesenchyme blastula stage (30 hr), or late gastrula stage (54 hr) for whole mount in situ hybridization. The injection solution contained 40% glycerol, 0.2 M KCl, 1500-3000 molecules per pl of plasmid, and a 5-fold molar excess of carrier DNA.

Whole mount in situ hybridization

Whole mount in situ hybridizations were carried out according to Ransick et al. (1993), but using a modified fixation procedure. Briefly, embryos were fixed in Streck Tissue Fixative (Streck Laboratories) at 4°C for 2 to 6 days. After a brief wash with filtered sea water and PBST buffer (PBS containing 0.1% Tween-20) the embryos were equilibrated with hybridization buffer (50% formamide, 10% PEG, and 0.6 M NaCl, 5 mM EDTA, 20 mM Tris, pH 7.5, 500 μg/ml yeast tRNA, 0.1% Tween-20, and 2 X Denhardt's solution). It is worth noting that proteinase K digestion is not necessary when Streck fixative is used. Hybridization was carried out at 45°C overnight with a digoxygenin-labeled CAT antisense RNA probe at 0.02 ng/μl. Three washes of 1X SSC at 65°C were followed by incubation with alkaline phosphatase-conjugated anti-digoxygenin F(ab)

(Boehringer Mannheim Biochemicals). The chromogenic reaction was performed with NBT and BCIP (Boehringer Mannheim Biochemicals). After dehydration, the embryos were mounted for observation.

Construction of CyIIIa·CAT and mZ12CyIIIa·CAT

(1) *CyIIIa·CAT*

The *CyIIIa·CAT* construct contains a 4.5 kb genomic DNA fragment that includes the 5' *CyIIIa* regulatory region, transcription start site, and 5' leader sequence of the *CyIIIa* primary transcript (Akhurst et al., 1987; Flytzanis et al., 1987). The *CyIIIa·CAT* construct was linearized for microinjection at a unique *Not*I site (Fig. 3).

(2) *mZ12CyIIIa·CAT*

Site-directed mutagenesis of the SpZ12-1 binding sites was carried out by PCR. A 1.2 kb *Kpn*I fragment from the *CyIIIa·CAT* plasmid that included the two SpZ12-1 binding sites was subcloned into the pBluescript KS (+) vector at its *Kpn*I site, to generate pBlue HE. The site-specific substitutions were made with two overlapping mutagenic oligonucleotide primers (DS-T7 and DS-T3). The sequence of these primers, with the overlapping region underlined, and the substituted nucleotides indicated in boldface, were as follows:

Primer for SpZ12-1 binding site I:

—DS-T7, 5'-<u>ACAAACAATATGGCTTGAC</u>CTA**TTGATTCCTCT**GTACCGTGCATTCTG-3'

Primer for SpZ12-1 binding site II:

—DS T3, 5'-<u>GTCAAGCCATATTGTTTGTC</u>CTA**TTGGTTCCTC**ACTAATGTATGCCGTTC-3'

The first round of PCR was designed to introduce base substitutions into the SpZ12-1 binding sites I and II in two separate fragments using a combination of mutant and non-mutant primers. The reactions were performed in 25 μl containing 100 ng template DNA (pBlue HE), 0.4 mM DS-T7 or DS-T3 primers, 4 mM T7 or T3 primers, 0.4 mM each dNTP, and one unit of Vent DNA polymerase (New England Bioblas), with 15 cycles of amplification (94°C, 30 sec; 55°C, 1 min; 72°C, 1 min). The second round of PCR reactions were carried out to rejoin the two purified PCR fragments (100 ng each) using the external T7 and T3 primers. Following *Kpn*I digestion of the final PCR product, the *mZ12CylIIa·CAT* fusion gene was generated by inserting the 1.2 kb fragment back into the *CylIIa·CAT* fusion gene at the *Kpn*I site (see Fig. 3).

Cloning of the SpZ12-1 gene

The sequence of the SpZ12-1 binding site I was used to generate the complementary oligonucleotide probe shown below (Calzone et al., 1988; Thézé et al., 1990):

5'-ACATGTTGCTAGGTAGGTCAA-3'

3'-CAACGATCCATCCAGTTTGTA-5'

A λZAP (Stratagene) cDNA expression library of 4 hr embryo poly(A) RNA was screened with this probe as described by Vinson et al. (1988) and Höög et al. (1991). One positive clone was isolated and both strands were sequenced using subclones generated by exonuclease III digestion. Sequencing was performed by the dideoxy method, using Sequenase reagents (U.S. Biochemical Corp.).

Expression, purification and renaturation of bacterial proteins

A T7 RNA polymerase based expression system (Studier and Moffat, 1986) was used to prepare recombinant SpZ12-1 protein. Three different expression constructs (SpZ12-1, SpZ12-1ΔZF and SpZ12-1ΔN) were made in pET vectors. The SpZ12-1 construct contains an *Eco*RI-*Nci*I fragment of SpZ12-1 cDNA. The fragment was end-filled and ligated into the pET3c vector at the *Bam*HI site. The SpZ12-1ΔZF construct, which includes an *Eco*RI-*Mlu*I fragment of SpZ12-1 cDNA, was generated by ligating the end-filled fragment into the pET3c vector at the *Bam*HI site. The SpZ12-1ΔN construct contains an *Esp*I-*Nci*I fragment of SpZ12-1 cDNA, prepared by joining the end-filled fragment into the *Bam*HI site of pET3b vector. Expression of SpZ12-1 protein in *E. coli* was performed as described by Studier and Moffat (1986) and Calzone et al. (1991). Briefly, IPTG-induced bacterial cells were harvested by centrifugation, and the pellet was resuspended in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, 0.1% Triton X-100 and 1 mM PMSF. Following a 1 hr incubation with 1 mg/ml lysozyme on ice and brief sonication, 15,000 x g pellets were washed in 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 5 mM EDTA, 1% Triton X-100, 25% sucrose, and 1 mM PMSF. After centrifugation at 15,000 x g, the washed pellets were resuspended in 20 mM HEPES-KOH (pH 7.9), 0.1 mM EDTA, 1 mM DTT, and 5 M guanidine hydrochloride. In order to renature the SpZ12-1 proteins, guanidine hydrochloride was diluted from 5 to 0.5 M with 10 mM HEPES-KOH (pH 7.9), 100 mM KCl, 0.2 mM $ZnCl_2$, 0.1 mM PMSF, 1 mM DTT, 0.1% NP-40 and 20% glycerol in three serial 2-fold dilution steps. At each step the solution was incubated at 4°C for 4 to 6 hr. After renaturation, insoluble protein was eliminated by centrifugation at 15,000 x g and the supernatant was stored at -70°C.

Antibodies and immunoblots

SpZ12-1ΔZF protein (Fig. 8) was expressed from *E. coli* as described above and further purified by SDS gel. The purified protein was used for the immunization of rabbits. The specificity of the antiserum preparations (Cocalico Biological, Inc.) was tested by using bacterially expressed SpZ12-1 protein and whole embryo lysate. Immunoblots were carried out according to Calzone et al. (1991). Immunoreactive proteins were visualized by the ECL detection system (Amersham) following procedures recommended by the supplier.

Acknowledgments

We thank Dr. Christer Höög, Ann E. Cutting and Vikas Duvvuri for technical assistance, and Dr. Kirsten L. Rood for critically editing a draft of the manuscript. This research was supported by NIH Grant HD-05753. D.G.W. was supported by a McCallum Fellowship, California Biochemical Foundation and NIH Grant HD-05753. C.V.K. was supported by a training grant from the Austrian government.

References

Akhurst, R.J., F.J. Calzone, J.J. Lee, R.J. Britten, and E.H. Davidson. 1987. Structure and organization of the *CyIII* actin gene subfamily of the sea urchin *Strongylocentrotus purpuratus*. *J. Mol. Biol.* **194**: 193-203.

Anderson, R., R.J. Britten, and E.H. Davidson. 1994. Repeated sequence target sites for maternal DNA-binding proteins in genes activated in early sea urchin development. *Dev. Biol.* **163**: 11-18.

Berg, J.M. 1990. Zinc finger domains: hypotheses and current knowledge. *Annu. Rev. Biophys. Biophys. Chem.* **19**: 405-421.

Calzone, F.J., N. Thézé, P. Thiebaud, R.L. Hill, R.J. Britten, and E.H. Davidson. 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.

Calzone, F.J., C. Höög, D.B. Teplow, A.E. Cutting, R.W. Zeller, R.J. Britten, and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo I. Purification by affinity chromatography and cloning of P3A2, a novel DNA-binding protein. *Development* **112**: 335-350.

Cameron, R.A., B.R. Hough-Evans, R.J. Britten, and E.H. Davidson. 1987. Lineage and fate of each blastomere of the eight-cell sea urchin embryo. *Genes & Dev.* **1**: 75-85.

Cameron, R.A., S.E. Fraser, R.J. Britten, and E.H. Davidson. 1990. Segregation of oral from aboral ectoderm blastomeres is completed at the fifth cleavage in the embryogenesis of *Strongylocentrotus purpuratus*. *Dev. Biol.* **137**: 77-85.

Cameron, R.A. and E.H. Davidson. 1991. Cell type specification during sea urchin development. *Trends Genet.* **7**: 212-218.

Cameron, R.A., R.J. Britten, and E.H. Davidson. 1993. The embryonic ciliated band of the sea urchin *Strongylocentrotus purpuratus* derives from both oral and aboral ectoderm. *Dev. Biol.* **160**: 369-376.

Char, B.R., J.R. Bell, J. Dovala, J.A. Coffman, M.G. Harrington, J.C. Becerra, E.H. Davidson, F.J. Calzone, and R. Maxson. 1993. *SpOct*, a gene encoding the major octamer-binding protein in sea urchin embryos: expression profile, evolutionary relationships, and DNA binding of expressed protein. *Dev. Biol.* **158**: 350-363.

Chowdhury, K., U. Deutsch, and P. Gruss. 1987. A multigene family encoding several "finger" structures is present and differentially active in mamalian genomes. *Cell.* **48**: 771-778.

Cox, K.H., L.M. Angerer, J.J. Lee, R.J. Britten, E.H. Davidson, and R.C. Angerer. 1986. Cell lineage-specific programs of expression of multiple actin genes during sea urchin embryogenesis. *J. Mol. Biol.* **188**: 159-172.

Davidson, E.H. 1986. *Gene Activity in Early Development* 3rd ed., Academic Press, Orlando, FL.

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

Flytzanis, C.N., R.J. Britten, and E.H. Davidson. 1987. Ontogenic activatin of a fusion gene introduced into the sea urchin egg. *Proc. Natl. Acad. Sci. USA* **84**: 151-155.

Franks, R.R., R. Anderson, J.G. Moore, B.R. Hough-Evans, R.J. Britten, and E.H. Davidson. 1990. Competitive titration in living sea urchin embryos of regulatory factors required for expression of the *CyIIIa* actin gene. *Development* **110**: 31-40.

Hickey, R.J., M.F. Boshar, and W.R. Crain. 1987. Transcription of three actin genes and a repeated sequence in siolated nuclei of sea urchin embryos. *Dev. Biol.* **124**: 215-225.

Höög, C., F.J. Calzone, A.E. Cutting, R.J. Britten, and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo II. Two dissimilar proteins, P3A1 and P3A2, bind to the same target sites that are required for early territorial gene expression. *Development* **112**: 351-364.

Hough-Evans, B.R., R.R. Franks, R.A. Cameron, R.J. Britten, and E.H. Davidson. 1987. Correct cell type-specific expression of a fusion gene injected int sea urchin eggs. *Dev. Biol.* **121**: 576-579.

Hough-Evans, B.R., R.J. Britten, and E.H. Davidson. 1988. Mosaic incorporation and regulated expression of an exogenous gene in the sea urchin embryo. *Dev. Biol.* **129**: 198-208.

Hough-Evans, B.R., R.R. Franks, R.W. Zeller, R.J. Britten, and E.H. Davidson. 1990. Negative spatial regulation of t he lineage-specific *CyIIIa* actin gene in the sea urchin embryo. *Development* **110**: 41-50.

Lee, J.J., F.J. Calzone, and E.H. Davidson. 1992. Modulation of sea urchin actin mRNA prevalence during embryogenesis: nuclear synthesis and decay rate measurements of transcripts from five different genes. *Dev. Biol.* **149**: 415-431.

Livant, D., A. Cutting, R.J. Britten, and E.H. Davidson. 1988. Titration of the activity of a fusion gene in intact sea urchin embryos. *Proc. Natl. Acad. Sci. USA* **85**: 7607-7611.

Livant, D., B.R. Hough-Evans, J.G. Moore, R.J. Britten, and E.H. Davidson. 1991. Differential stability of expression of similarly-specified endogenous and exogenous genes in the sea urchin embryo. *Development* **113**: 385-398.

Margolin, J.F., J.R. Friedman, W.K.-H. Meyer, H. Vissing, H.-J. Thiesen, and F.J. Rauscher III. 1994. *Krüppel*-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. USA* **91**: 4509-4513.

McMahon, A.P., C.N. Flytzanis, B.R. Hough-Evans, K.S. Katula, R.J. Britten, and E.H. Davidson. 1985. Introduction of cloned DNA into sea urchin egg cytoplasm: replication and persistence during embryogenesis. *Dev. Biol.* **108**: 420-430.

Pavletich, N.P. and C.O. Pabo. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**: 809-817.

Pavletich, N.P. and C.O. Pabo. 1993. Crystal structure of a five-finger GLI-DNA complex: new perspective on zinc fingers. *Science* **261**: 1701-1707.

Pellegrino, G.R. and J.M. Berg. 1991. Identification and characterization of "zinc finger" domains by the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **88**: 671-675.

Ransick, R., S.G. Ernst, R.J. Britten, and E.H. Davidson. 1993. Whole mount in situ hybridization shows *Endo 16* to be a marker for the vegetal plate territory in sea urchin embryos. *Mech. Dev.* **42**: 117-124.

Rosenberg, U.B., C. Schröder, A. Preiss, A. Kienlin, S. Côté, L. Riede, and H. Jäckle. 1986. Structural homology of the product of the Drosophila *Krüppel* gene with *Xenopus* transcription factor IIIA. *Nature* **319**: 336-339.

Shott, R.J., J.J. Lee, R.J. Britten, and E.H. Davidson. 1984. Differential expression fo the actin gene family of *Strongylocentrotus purpuratus*. *Dev. Biol.* **101**: 295-306.

Studier, F.W. and B.A Moffat. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**: 113-130.

Thézé, N., F.J. Calzone, P. Thiebaud, R.L. Hill, R.J. Britten, and E.H. Davidson. 1990. Sequences of the *Cyllla* actin gene regulatory domain bound

specifically by sea urchin embryo nuclear proteins. *Mol. Reprod. Dev.* **25**: 110-122.

Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and S.L. McKnight. 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes & Dev.* **2**: 801-806.

Witzgall, R., E. O'Leary, A. Laef, D. Önaldi, and J.V. Bonventre. 1994. The *Krüppl*e-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. USA* **91**: 4514-4518.

Table 1. Ectopic expression pattern of *mZ12CyIIIa·CAT* fusion gene in mesenchyme cells at late gastrula stage (~54 hr)

Construct	Batch	Total embryos stained[a]	Embryos with correct expression[b]	Embryos with ectopic expression in mesenchyme cells	Ectopic as % of stained embryos[c]
CyIIIa·CAT	1	41	40	1	2.4
(control)	2	45	41	4	8.9
	Total	**86**	**81**	**5**	**5.6**
mZ12CyIIIa·CAT	1	48	26	22	46
	2	53	30	23	44
	3	95	50	45	47
	Total	**196**	**106**	**90**	**46**

[a]Embryos were scored if two or more cells were stained.

[b]Embryos with correct expression were scored if two or more cells were stained in a patch within the aboral ectoderm.

[c]Ectopic expression as % of stained embryos = [Σ embryos stained ectopically / Σ stained embryos] x 100.

Table 2. Expression pattern of *mZ12CyIIIa·CAT* fusion gene at mesenchyme blastula stage (~30 hr)

Construct	Batch	Total embryos stained[a]	Embryos with correct expression[b]	Embryos with ectopic expression in mesenchyme cells	Ectopic as % of stained embryos[c]
CyIIIa·CAT	1	18	18	0	0
(control)	2	36	36	0	0
	Total	**54**	**54**	**0**	**0**
mZ12CyIIIa·CAT	1	22	22	0	0
	2	48	46	2	4
	Total	**70**	**68**	**2**	**3**

[a]Embryos were scored if two or more cells were stained.

[b]Embryos with correct expression were scored if two or more cells were stained in a patch within the aboral ectoderm.

[c]Ectopic expression as % of stained embryos = [Σ embryos stained ectopically / Σ stained embryos] x 100.

Figure 1. The 5' regulatory domain of the *CyIIIa* cytoskeletal actin gene. Black boxes represent the specific sites of interaction with nuclear proteins from blastula stage embryos in the 5'-flanking region of the *CyIIIa* gene, which extends from -2280 to +235 bp (Calzone et al., 1988; Thézé et al., 1990; J. A. Coffman, unpublished data). The bent arrow indicates the transcription start site. Note that there are two adjacent SpZ12-1 binding sites in inverse orientation within this region. Restriction sites that flank this region are *Bam*HI (b) and *Pst*I (p). Nuclear factors binding to the sites are indicated: Sp(G/C)F-1 is a factor of novel sequence that recognizes a G/C-rich motif (Coffman and Davidson, 1992, R. W. Zeller and E. H. Davidson, unpublished data); SpRunt is an *S. purpuratus* factor homologous to the *Drosophila* runt protein (J. A. Coffman and E. H. Davidson, unpublished data); SpP3A2 and P7II are both factors of novel sequence, that may also be involved in distinct negative control functions of the *CyIIIa* regulatory system (Hough-Evans et al., 1990, Calzone et al., 1990, C. V. Kirchhamer and E. H. Davidson, unpublished data, J.A. Coffman and E.H. Davidson, unpublished data); SpCTF-1 is a CCAAT binding factor (Calzone et al., 1988, Barberis et al., 1987); SpOct-1 is an octamer factor (Char et al., 1993); SpTEF-1 is an *S. purpuratus* factor homologous to the human enhancer factor TEF-1 (J. Xian and E.H. Davidson, unpublished data). P9 and P1 are not yet characterized. All of these sites of interaction appear to have functional roles in the control of *CyIIIa* expression (with the possible exception of the P9 interaction , which has not yet been studied).

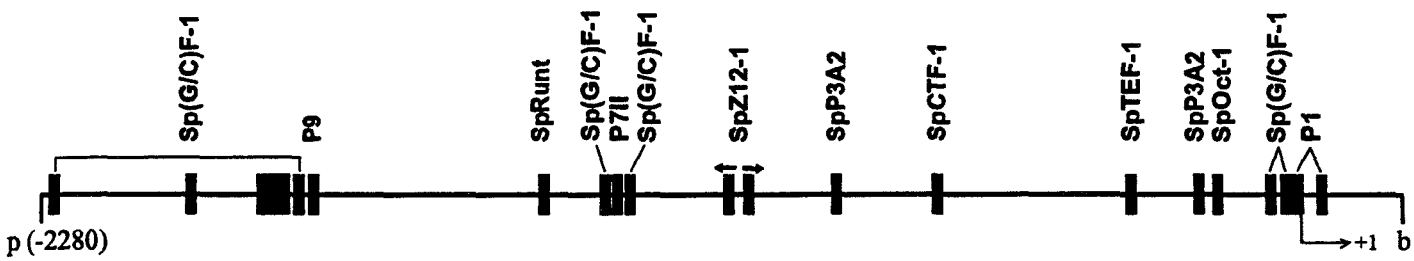

Sp(G/C)F-1
P9
SpRunt
Sp(G/C)F-1
P7II
Sp(G/C)F-1
SpZ12-1
SpP3A2
SpCTF-1
SpTEF-1
SpP3A2
SpOct-1
Sp(G/C)F-1
P1
p (-2280)
+1
b

Figure 2. Expression pattern of the *CyIIIa·CAT* fusion gene at late gastrula stage. The spatial distribution of CAT transcripts in late gastrula stage embryos was detected by whole mount in situ hybridization, using a digoxygenin-labeled antisense CAT RNA probe. Six late gastrula stage embryos with expression restricted to the aboral ectoderm are shown, all oriented with the oral surface to the right. (A) An embryo showing extensive staining along the aboral ectoderm extending to the lower aboral-oral boundary. (B, C) Two embryos displaying a patch of labeled cells toward the upper aboral-oral boundary. (D, E, F) Three embryos expressing CAT transcripts in large regions of the aboral ectoderm.

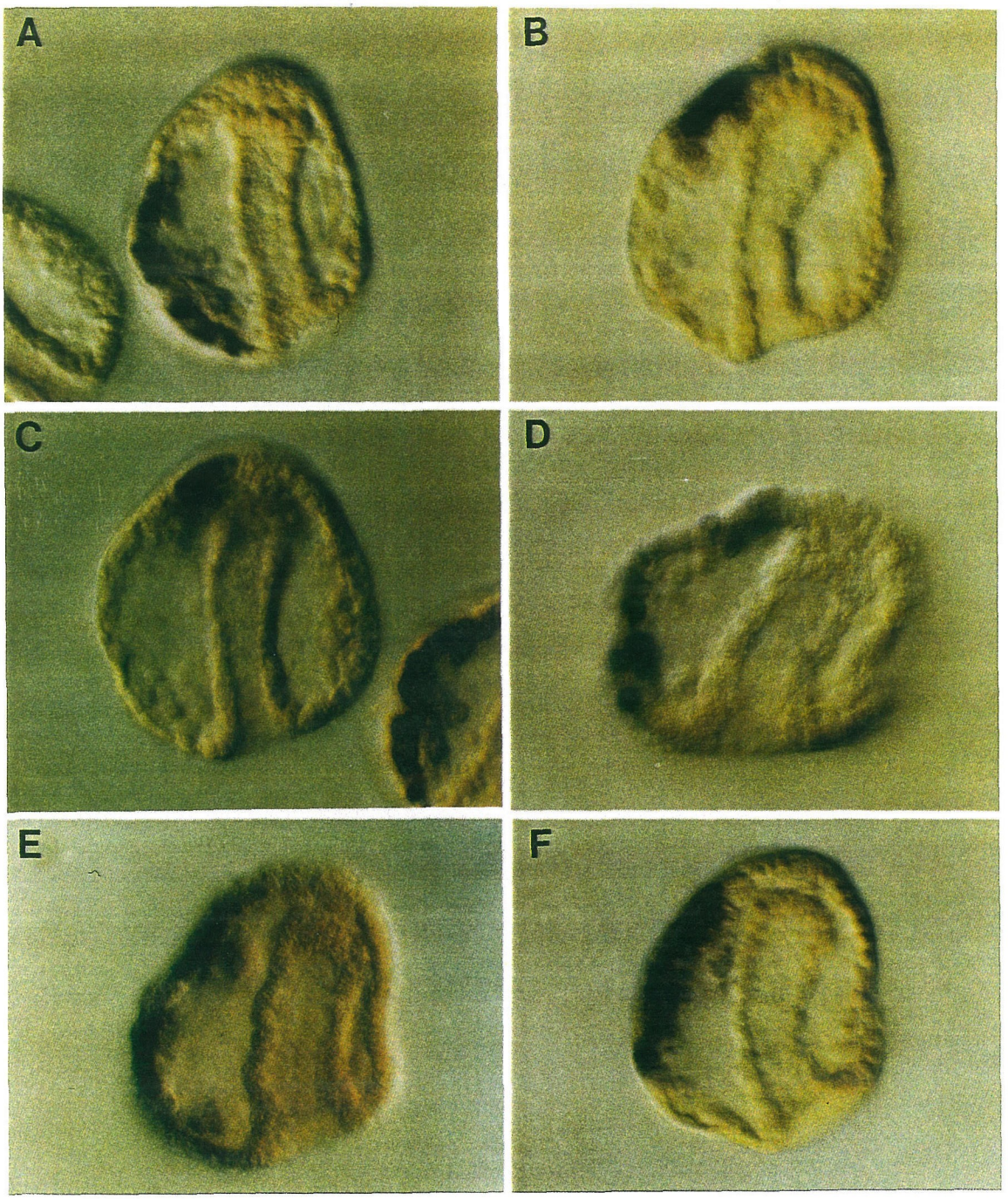

A
B
C
D
E
F

Figure 3. *mZ12CyIIIa·CAT* fusion gene and mutated target sites. (A) A schematic representation of the *mZ12CyIIIa·CAT* fusion gene. The *mZ12CyIIIa·CAT* fusion gene includes the 5' *CyIIIa* regulatory sequence and portions of a non-translated exon, extending from -2280 to +235 bp relative to the transcription start site. The sequences of two adjacent SpZ12-1 binding sites, I and II, in inverse orientation are underlined, at -1030 to -1008 and -994 to -979 bp, respectively. Underlined sequences are derived from the DNase I footprinting experiment shown below. The arrows indicate the inversely repeated core target site sequence noted in text. Substituted nucleotides in the mZ12 mutant gene are indicated below, in boldface. (B) DNA binding activity of the wild-type (Wt) and the mutant (mZ12) SpZ12-1 binding sites. ^{32}P-labeled probes include the two adjacent SpZ12-1 binding sites contained in the DNA fragment that extends from positions -1082 to -894 of the 5' *CyIIIa* regulatory region. Gel shift assays were carried out using blastula stage nuclear extracts as previously described (Calzone et al., 1988). The nine nucleotide substitutions within each of the SpZ12-1 binding sites suffice to abolish all DNA binding activity, and only the free probe is observed.

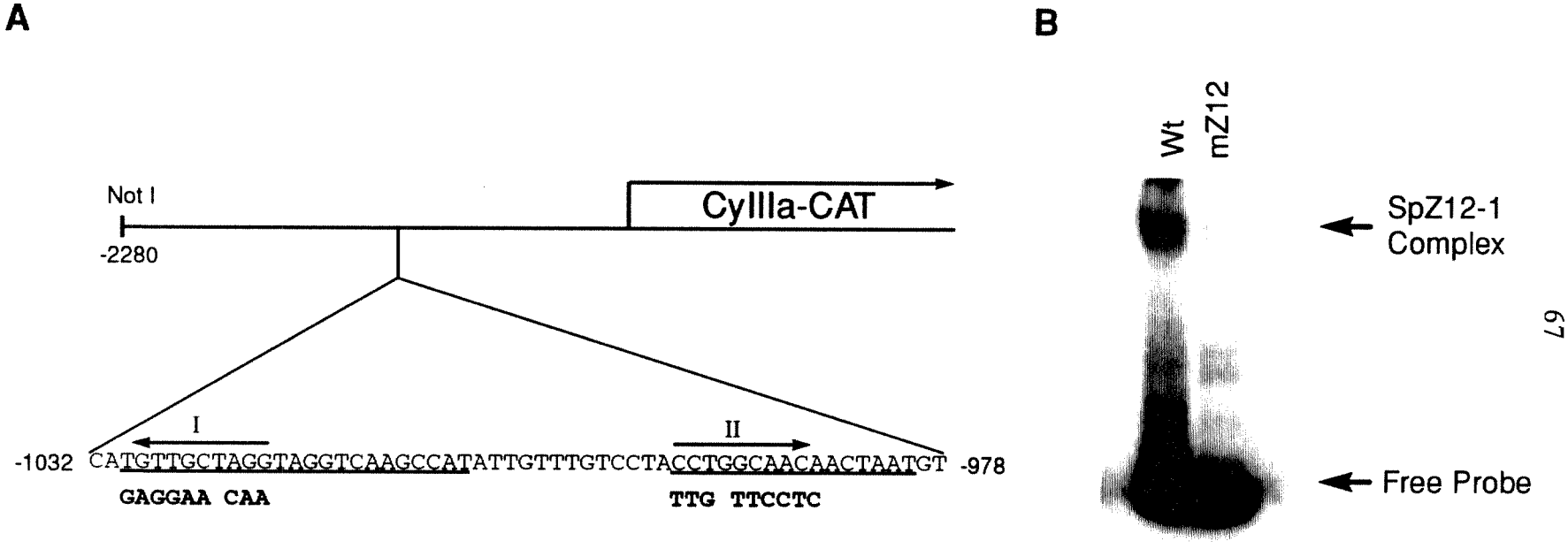

A
Not I
-2280
CyIIIa-CAT
I
II
-1032
CATGTTGCTAGGTAGGTCAAGCCATATTGTTTGTCCTACCTGGCAACAACTAATGT
-978
GAGGAA CAA
TTG TTCCTC
B
Wt
mZ12
SpZ12-1
Complex
Free Probe

Figure 4. Ectopic expression pattern of the *mZ12CylIIa·CAT* fusion gene at gastrula stage. (A) and (B), Embryos expressing CAT transcripts in closely spaced skeletogenic mesenchyme cells surrounding the archenteron, and along the inner ectodermal wall, vegetal view. Note that both embryos show a patch of stained cells in the aboral ectoderm toward the top left sides. (C)-(F) Side views, vegetal pole down. The archenteron, topped with delaminating secondary mesenchyme cells, can be seen in the optical sections shown. (C) An embryo in which expression occurs only in skeletogenic mesenchyme cells. (D, E, F) Three other embryos displaying expression in skeletogenic mesenchyme cells. In other optical sections it can be seen that these particular embryos also express CAT transcripts in the aboral ectoderm.

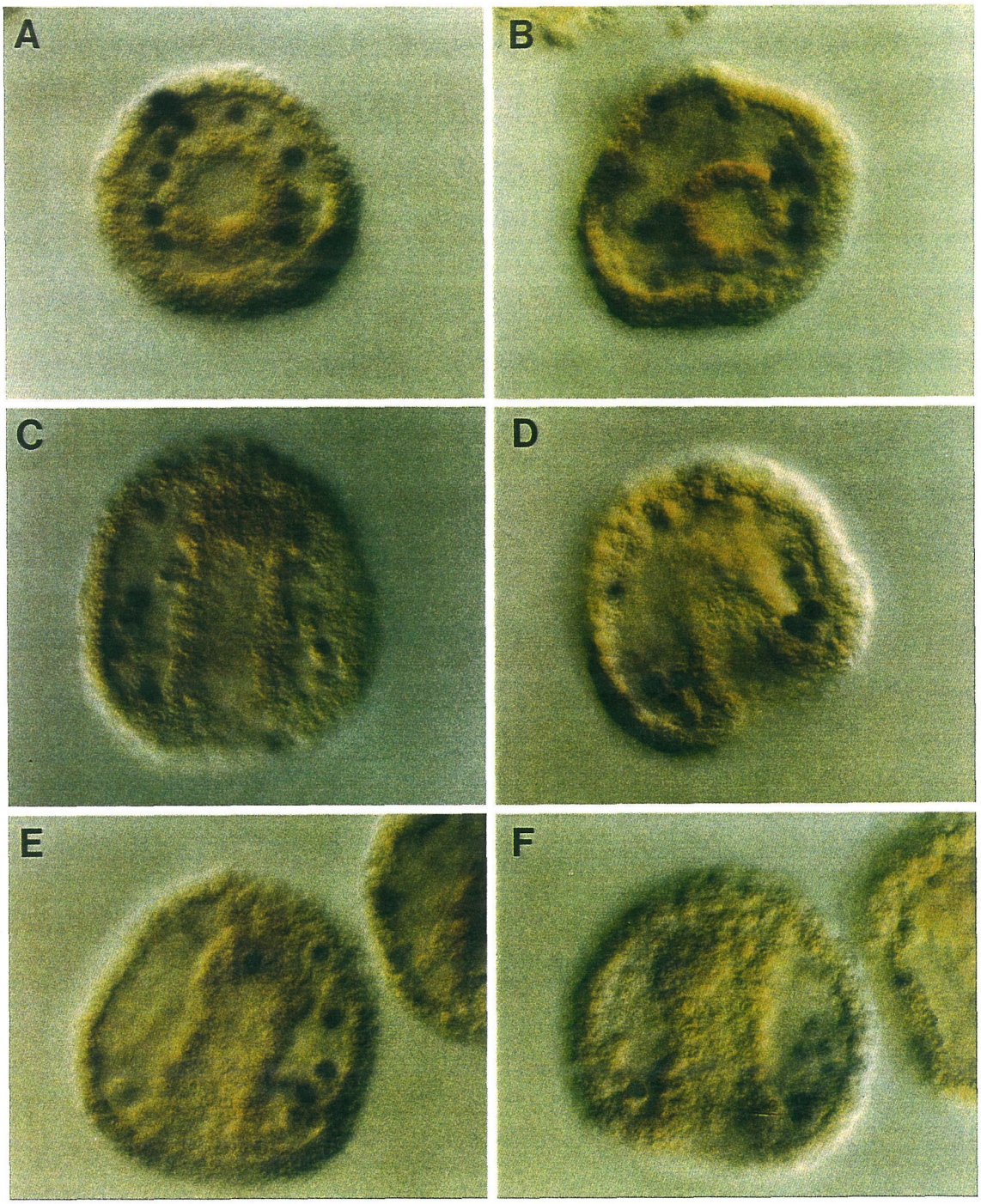

A
B
C
D
E
F

Figure 5. Expression pattern of the *mZ12CylIIa·CAT* fusion gene at mesenchyme blastula stage. (A) and (B) Mesenchyme blastula stage embryos showing staining in the ectoderm close to the vegetal plate, but not in the ingressed skeletogenic mesenchyme cells. (C) An embryo with expression located in the ectoderm closer to the animal pole, but no expression in the newly ingressed skeletogenic mesenchyme.

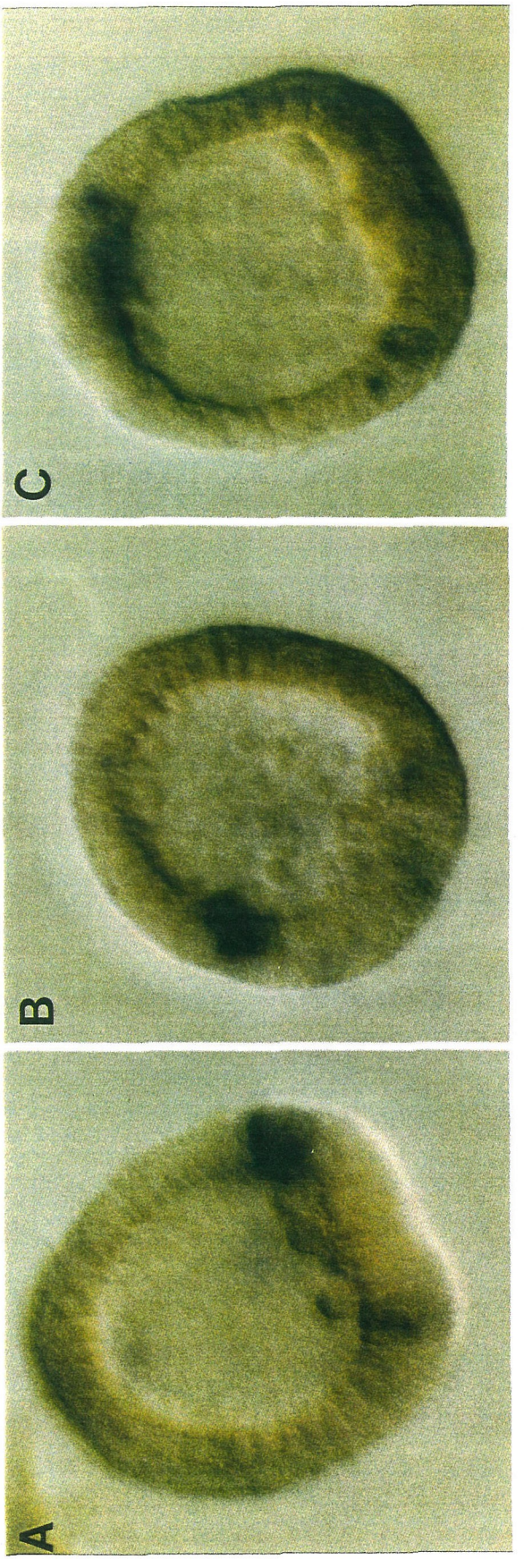

C
B
A

Figure 6. Sequences of SpZ12-1 cDNA and protein, and sequence similarities of the Zn finger domain between the SpZ12-1 and *Krüppel* proteins (Rosenberg et al., 1986). (A) The complete nucleotide sequence and the deduced amino acid sequence of the SpZ12-1 cDNA are shown. The Zn finger region is underlined. The numbered bent arrows below the amino acid sequence mark the positions of the twelve zinc fingers. (B) Alignment of Zn finger sequences of the SpZ12-1 and *Krüppel* proteins. The fingers of both proteins are numbered on the left side. Positions and amino acids that are identical among the Zn fingers are shaded.

A

```
   1 AATTCGGCCGATTCGGGCCCGTGTCACATCCAGGAGAGATCCAGCACAGCAGACGACAAAATGACACAGGCGATGGATCTCTCTAAGAAA
   1                                                                   M  D  L  S  K  K

  91 CGAAAAAGAGATGCACCTATACCGGTCACAGTGCCGGTTATAGCACCGGTTACAGCACAGGTAACCATTAAACGGGAGCCAGTGGAGGAA
   7  R  K  R  D  A  P  I  P  V  T  V  P  V  I  A  P  V  T  A  Q  V  T  I  K  R  E  P  V  E  E

 181 ACTAGTGATGATGGCTACTTTACTAGTCTCGGGAGGGGGAGCGAGAAACATTATGCAAAGTCGTCCAGCCATGTCTTTCAGTACGCTACG
  37  T  S  D  D  G  Y  F  T  S  L  G  R  G  S  E  K  H  Y  A  K  S  S  S  H  V  F  Q  Y  A  T

 271 CAACATCAGGGTATTCCTTTCAGTCCAAGCAGCCCTGTCTATCACCATCCAGTACCAAACAATAATATCGAGATCAAACAAGAGACGGTG
  67  Q  H  Q  G  I  P  F  S  P  S  S  P  V  Y  H  H  P  V  P  N  N  N  I  E  I  K  Q  E  T  V

 361 ACGGACGAAGATAGGTGCGAGGTGTGCTTCAAACCTTACGATAATGCAGCACATGCGAGGACTACTTTCCGTTATCAACTGCATGATCTG
  97  T  D  E  D  R  C  E  V  C  F  K  P  Y  D  N  A  A  H  A  R  T  T  F  R  Y  Q  L  H  D  L

 451 GAGTTGGTGCTGTGTATGAAATCGTGCGCGAGCATCGAGTCGAGGCAGCAGGAGGAGCACTTGGATTCTCGGAGCGATGACATGTCCACG
 127  E  L  V  L  C  M  K  S  C  A  S  I  E  S  R  Q  Q  E  E  H  L  D  S  R  S  D  D  M  S  T

 541 GCTGAAGATGAGACGGACAGCGGGTCGCAGGATACAGGGAGCTCGGAACGACAAGAGAGGTATGAGCAGTTCCGTGAAGAGTATCAGGGA
 157  A  E  D  E  T  D  S  G  S  Q  D  T  G  S  S  E  R  Q  E  R  Y  E  Q  F  R  E  E  Y  Q  G

 631 ATTGGGAATCTAGATGGTGATGCGGATGGTACGGATGAGGAGAACGAGGGGGCGCTGAGCATCGAGGCTGACAGTGAGAGCAGCCAGTCG
 187  I  G  N  L  D  G  D  A  D  G  T  D  E  E  N  E  G  A  L  S  I  E  A  D  S  E  S  S  Q  S

 721 GAGCCATTCACAAGACCATACATTGCCGGTGGGTCGCAGTACATTGACGGGATAGCAACCATCTGGGCCAGATCCAGAGTGTCTTCATGG
 217  E  P  F  T  R  P  Y  I  A  G  G  S  Q  Y  I  D  G  I  A  T  I  W  A  R  S  R  V  S  S  W

 811 AGTGAGGCCCTACGAATGCAAAACTGCGGGAAACGGTTCGCCCAGAAGAGCTGTCTTACGCGTCACACCAAGATCCACACCGGAGAGAGA
 247  S  E  A  L  R  M  Q  N  C  G  K  R  F  A  Q  K  S  C  L  T  R  H  T  K  I  H  T  G  E  R
                 └1→

 901 CTCCATCAATGCGACGAGTGCACCAAGTCTTTCACCCGCGCCAGCAATCTGCGGTCCCACATGCGGACGCACACCGGCGAGAAGTGCTTC
 277  L  H  Q  C  D  E  C  T  K  S  F  T  R  A  S  N  L  R  S  H  M  R  T  H  T  G  E  K  C  F
       └2→                                                                             └3→

 991 CAGTGCCCCATCTGCCAGAAGCTCTTCGGTCAGAGTTCCTGGCTGAAGCGTCACATCCGCACCCACACGGGCGAGAAGGAGGTCAAGTGC
 307  Q  C  P  I  C  Q  K  L  F  G  Q  S  S  W  L  K  R  H  I  R  T  H  T  G  E  K  E  V  K  C
                                                                                 └4→

1081 CCGCTCTGCAGCAAGTGGGTCAGCTCCACGACCAGTCTTCAGTCCCATATGAGAACACACACGGGGGAGAAACCCTTCGCCTGCCTCATA
 337  P  L  C  S  K  W  V  S  S  T  T  S  L  Q  S  H  M  R  T  H  T  G  E  K  P  F  A  C  L  I
                                                                       └5→

1171 TGTGACAAGAGCTTCAGTGTCAAGTGCACCCTGACGGTTCATATGCGTATCCACACAGGCAACAAGCCATACGTGTGTGAGGTGTGTGGT
 367  C  D  K  S  F  S  V  K  C  T  L  T  V  H  M  R  I  H  T  G  N  K  P  Y  V  C  E  V  C  G
                                                                             └6→

1261 GCTGCATTCATACAGGGGACCCAGCTGTCCACTCACATGCGGGTGCATACGGGTGAGAAGCCATACATGTGTGAGATCTGTCATCGGCAG
 397  A  A  F  I  Q  G  T  Q  L  S  T  H  M  R  V  H  T  G  E  K  P  Y  M  C  E  I  C  H  R  Q
                                                                       └7→

1351 TTCGGCAAGAAAGACAACCTCCTCACCCATGTCAAGATTCACACTGAAGGGCAGAATTTTGACTGCAAGTTCTGCAACCTACGTTTCAAC
 427  F  G  K  K  D  N  L  L  T  H  V  K  I  H  T  E  G  Q  N  F  D  C  K  F  C  N  L  R  F  N
                                                                 └8→

1441 AACTCTGAGGAGCTCAAGACCCATACCCTGATCCATATAGAAGAGAGACCATTCCAGTGTGAAGTGTGCAACAAAGCCTTCAAGAGAAAG
 457  N  S  E  E  L  K  T  H  T  L  I  H  I  E  E  R  P  F  Q  C  E  V  C  N  K  A  F  K  R  K
                                                                       └9→

1531 GACAACCTTAAAGAACACATTAAGGTTCACACCAAAGAGAAGCCTTTTAGCTGTGAGGTTTGCAAAGAGGCCTTCTCCAACGGCAGCCAT
 487  D  N  L  K  E  H  I  K  V  H  T  K  E  K  P  F  S  C  E  V  C  K  E  A  F  S  N  G  S  H
                                                          └10→

1621 CTTAAGGTCCACATGCGACGTCACACGGGGGACAAGCCCTACTCCTGTGAGCTCTGCGGGAAGTGCTTCACCCAGAGCAGTAGCTTGAAG
 517  L  K  V  H  M  R  R  H  T  G  D  K  P  Y  S  C  E  L  C  G  K  C  F  T  Q  S  S  S  L  K
                                                             └11→

1711 ACCCACTACCTGATCCACACCGGTGAGAAACCTTACTCCTGTACTATCTGCCATGCACTCTTCACCCGCAATGCCAGCGTCAAGAAGCAC
 547  T  H  Y  L  I  H  T  G  E  K  P  Y  S  C  T  I  C  H  A  L  F  T  R  N  A  S  V  K  K  H
                                                      └12→

1801 ATGAAGAGAGCCCACAATGCCGACATGCCGGGACCAGAGGAGGCGTCTGAATGAGTGGGCCTCAGAGCAGTAGCTTAAAGACCCACTACC
 577  M  K  R  A  H  N  A  D  M  P  G  P  E  E  A  S  E  *

1891 TGATCCGCACCGGTGAGAAACCTTATTCCTGGGGCGCGTTTCACAAAACTTGTCATTGAACAAAGGCTTTGTGAAACGGGTCCCTGTACT

1981 ATCTGCCAGGCGCTCTTCACCCGCAACGCACTCTTCAAACACTATGCCACCATGCCGGGACCGGACGATGCGTCTGAGTGAGCAGCCCTC

2071 GAGGAACCACTGGGAGGGAGTGAGTCAGCGGAGGTTAGGTACAGTCAAACCTGCTTTAGTGACCACCTGTCTACAAAGACCGCCTGTCTA

2161 TAAAGACCACATGTTTTGTTTCCCTTGAAAATTGTTTCTCATTGAAACATGTACTATTAAAGGAACCTGTCTACAAAGACAATTTTTGTG

2251 TATCCCTTGGGTGGTCGCTTTAGACAGGTTTGACTGTAGTTGTTGGTGTCACCATCCAACTTTTGTTTTGTATTTTTCTGATATGCACAC

2341 GGGCCTTTGTATGTGTAGATATGATGAGGTGTATTTTATACGTATTTATGTCTTTTCTAAATGATTAACTAAATTGATTCAGATGTTTTT

2431 CTCTGTGAACGATTGACTGATTGTAACTTAAAATCAACCAGTTAATTTGTTTGTTAAATAGATAAACCCACAATTATGAGGGAAATTTAT

2521 GTATACCTTTTGTAAACAATATGAAAAATAGTATGTCTGTTCATGTAGTTTCCATCTTATAAGTGTATAGTTAAATCTATTTGAAAATAT

2611 ATACATTTTAAAACTTTTCTAGCTGGTTTTGGGAAATTTCGAAAAAATATAAATTAAAAAAAAAAAAAACCGAATT
```

B

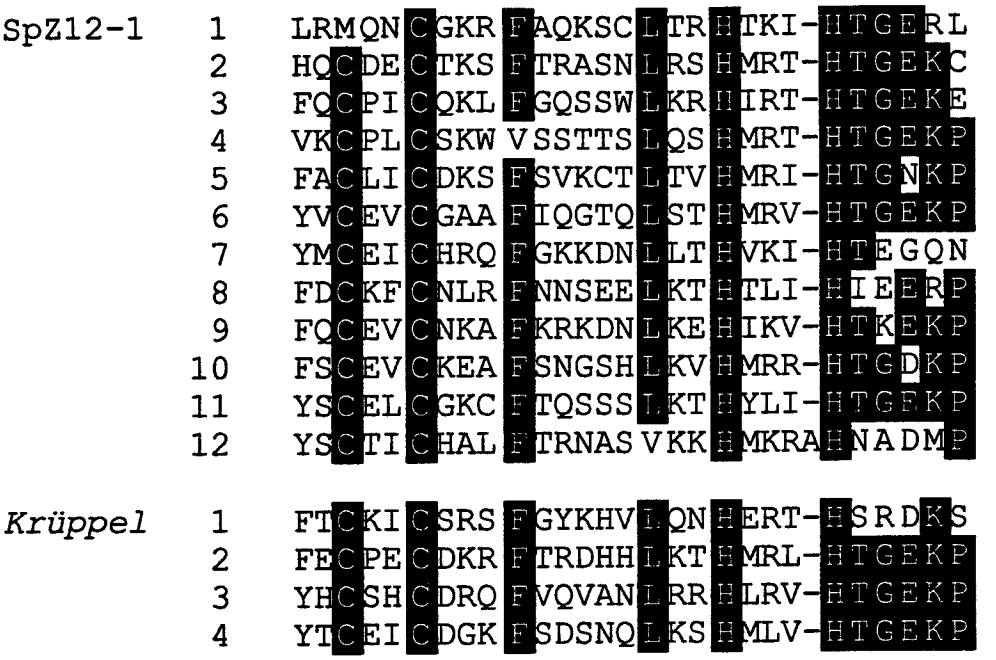
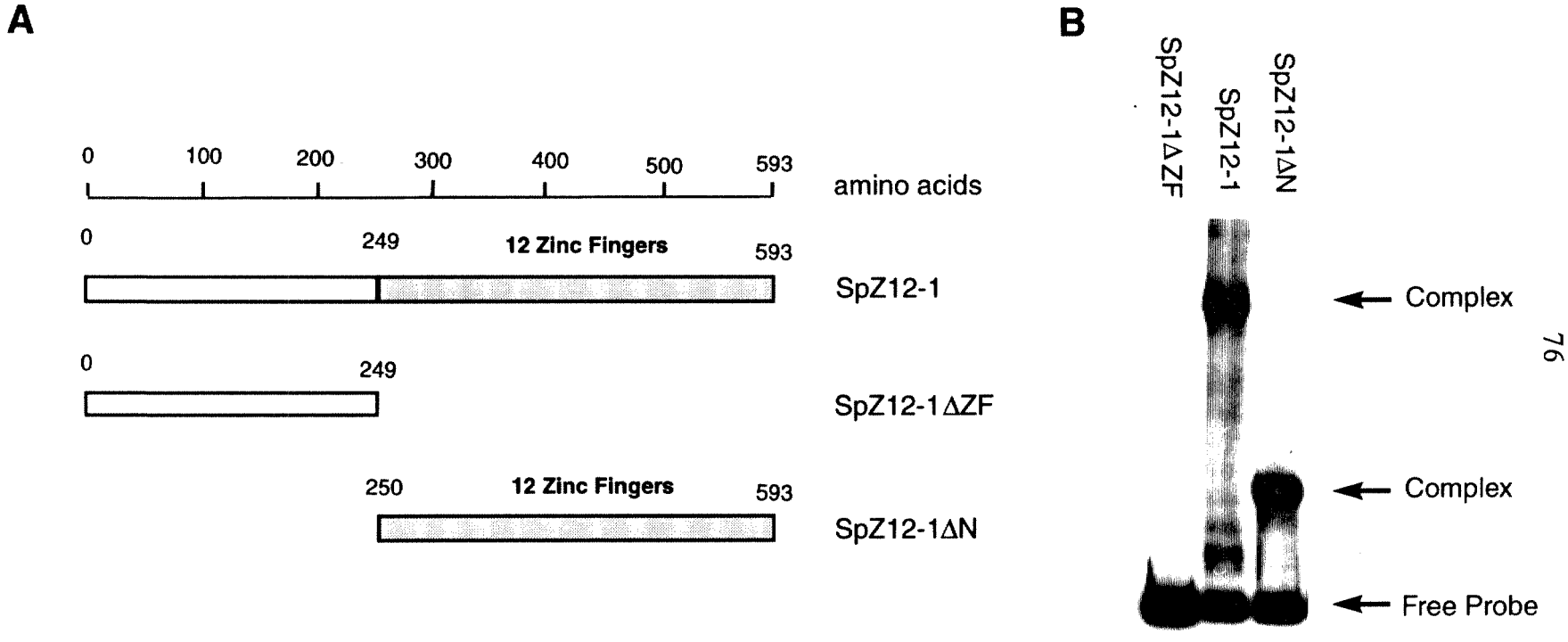

SpZ12-1 1 LRMQNCGKRFAQKSCLTRHTKI-HTGERL
2 HQCDECTKSFTRASNLRSHMRT-HTGEKC
3 FQCPICQKLFGQSSWLKRHIRT-HTGEKE
4 VKCPLCSKWVSSTTSLQSHMRT-HTGEKP
5 FACLICDKSFSVKCTLTVHMRI-HTGNKP
6 YVCEVCGAAFIQGTQLSTHMRV-HTGEKP
7 YMCEICHRQFGKKDNLLTHVKI-HTEGQN
8 FDCKFCNLRFNNSEELKTHTLI-HIEERP
9 FQCEVCNKAFKRKDNLKEHIKV-HTKEKP
10 FSCEVCKEAFSNGSHLKVHMRR-HTGDKP
11 YSCELCGKCFTQSSSLKTHYLI-HTGEKP
12 YSCTICHALFTRNASVKKHMKRAHNADMP
Krüppel 1 FTCKICSRSFGYKHVLQNHERT-HSRDKS
2 FECPECDKRFTRDHHLKTHMRL-HTGEKP
3 YHCSHCDRQFVQVANLRRHLRV-HTGEKP
4 YTCEICDGKFSDSNQLKSHMLV-HTGEKP

Figure 7. Mapping the DNA-binding domain of the SpZ12-1 protein. (A) A schematic diagram of the SpZ12-1 protein and its deletion mutants is shown. The shaded area indicates the zinc finger domain. (B) Gel shift assays were performed with equal amounts of SpZ12-1 and the mutant protein as indicated. The ^{32}P-labeled probe was a fragment of the *CyIIIa* regulatory sequence containing two adjacent SpZ12-1 binding sites extending from nucleotide -1033 to -821. The free probe migrates to the bottom of the gel, the upper and lower band represent the SpZ12-1 and the SpZ12-1ΔN complexes, respectively. The SpZ12-1 complex migrates more slowly because it is larger (~590 amino acids) than the finger domain alone (~350 amino acids). (C) Restoration of specific DNA-binding activity of EDTA-dialyzed SpZ12-1ΔN protein by addition of $ZnCl_2$. SpZ12-1ΔN protein samples were dialyzed against 10 mM EDTA, 10 mM HEPES-KOH (pH 7.9), 100 mM KCl, 0.5 M guanidine hydrochloride, 1 mM DTT, 0.1% NP-40 and 20% glycerol for 12 hr at 4°C. Gel shift assays were carried out with equal amounts (0.5 μl) of control and EDTA-dialyzed SpZ12-1ΔN protein. Lane 1 is a control showing the SpZ12-1ΔN complex formed in the presence of 2 mM $ZnCl_2$. The EDTA-dialyzed SpZ12-1ΔN protein was preincubated in the absence and presence of 2 mM $ZnCl_2$ for 5 min, and the ^{32}P-labeled probe was then added, as shown in lane 2 and 3, respectively.


A
0
100
200
300
400
500
593
amino acids
0
249
12 Zinc Fingers
593
SpZ12-1
0
249
SpZ12-1ΔZF
250
12 Zinc Fingers
593
SpZ12-1ΔN
B
SpZ12-1ΔZF
SpZ12-1
SpZ12-1ΔN
Complex
Complex
Free Probe

C

1 2 3

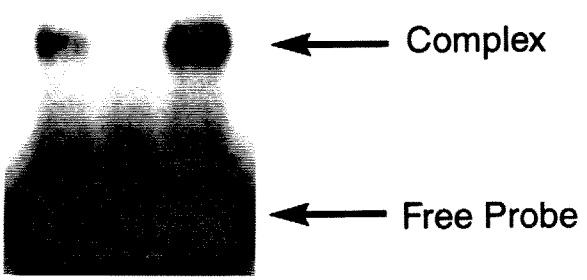
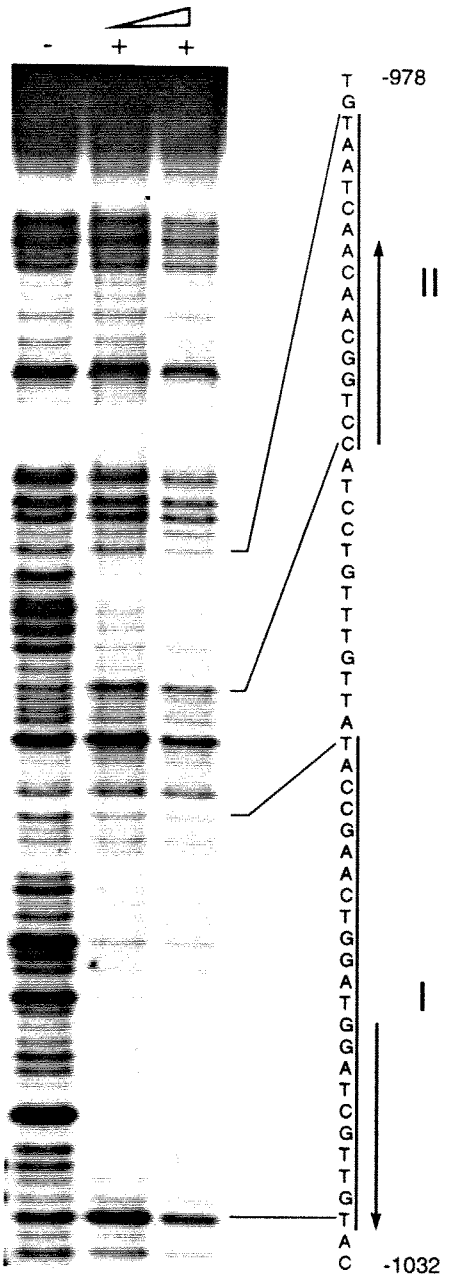


Figure 8. DNase I footprinting using the bacterially expressed, full-length SpZ12-1 protein. (A) DNase I footprint assays were carried out with increasing amounts of bacterially-expressed SpZ12-1 protein as indicated (Calzone et al., 1991). Lane (-) shows DNase I digestion pattern without added protein. The probe is the coding strand of the *CyIIIa cis*-regulatory region containing two SpZ12-1 binding sites (-1033 to -821). Protected regions of sites I and II are bracketed and their sequences are underlined. The inverted repeat sequence element in each site, and their relative orientations, are indicated by the arrows. (B) The nucleotides constituting the inverted repeat present in binding sites I and II are indicated in bold face.

A



B

SpZ12-1 Binding Site I:

-1030 -1008

TGTTGCT**AGG**TAGGTCAAGCCAT

ACAACGA**TCC**ATCCAGTTCGGTA

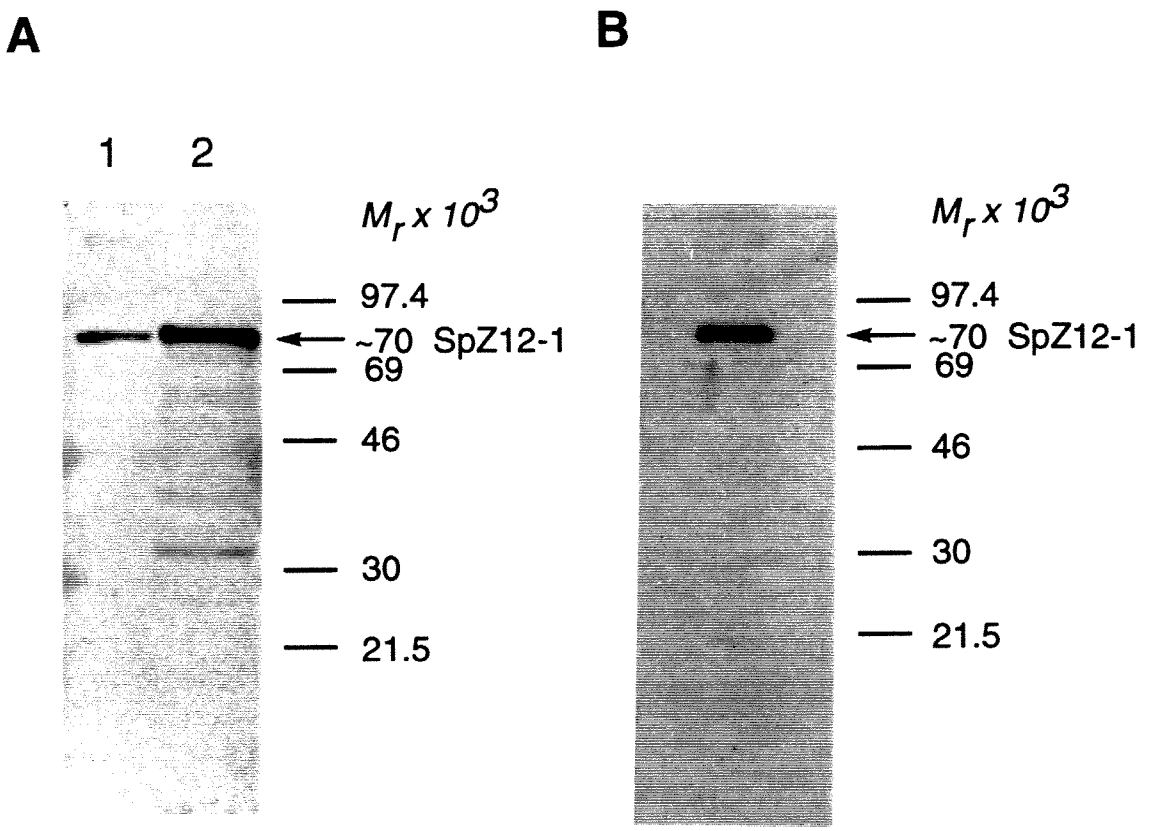
SpZ12-1 Binding Site II:

-994 -979

CCT**GGCAACA**ACTAAT

GGA**CCGTTGT**TGATTA

Figure 9. Identification of the SpZ12-1 protein in blastula stage nuclear extracts. Protein was eletrophoresed on a 7% SDS polyacrylamide gel, transferred to a nitrocellulose filter and reacted with the antibody. The positions of size markers and immunoreactive bands are indicated. (A) Lane 1 and 2 were loaded with bastula stage nuclear extract and full-length bacterially expressed SpZ12-1 protein, respectively. (B) The sample was prepared as follows: the SpZ12-1 binding site probe was mixed with late blastula stage nuclear extracts and the complex displayed by native gel electrophoresis. A gel slice containing the SpZ12-1 protein-DNA complex was excised, minced and incubated in an SDS extraction buffer. The whole mixture was then loaded.


A
1
2
M_r x 10^3
97.4
~70 SpZ12-1
69
46
30
21.5
B
M_r x 10^3
97.4
~70 SpZ12-1
69
46
30
21.5

CHAPTER 2

Maternal and Embryonic Provenance of a Sea Urchin Embryo Transcription Factor, SpZ12-1

Maternal and embryonic provenance of a sea urchin embryo transcription factor, SpZ12-1

David G.-W. Wang, Roy J. Britten and Eric H. Davidson

Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125

SpZ12-1 is a Zn finger transcription factor. Previous work has indicated that this factor functions late in embryogenesis as a spatial transcriptional repressor. We show here that this factor is present in significant quantities even in unfertilized egg cytoplasm, and in similar quantities in mesenchyme blastula-stage embryo cytoplasm. Taken together with earlier measurements of Calzone *et al.*(*Genes & Dev.* 2, 1074-1088, 1988), our observations indicate that SpZ12-1 enters the embryonic nuclei between late cleavage and mesenchyme blastula stages. A low prevalence mRNA encoding SpZ12-1 is also present throughout development. Translation of this mRNA could, however, easily account for the complete complement of SpZ12-1 protein in the embryo, as estimated from its DNA binding activity. SpZ12-1 probably functions at several developmental stages, and is evidently of both maternal and embryonic provenance.

Introduction

SpZ12-1 is a Zn finger transcription factor of *S. purpuratus,* which has been implicated in spatial control of the *CyIIIa* cytoskeletal actin gene. The *CyIIIa* gene is normally expressed in aboral ectoderm, and in studies presented elsewhere (Wang *et al.,* submitted) we showed that a striking and specific ectopic expression of an exogenous *CyIIIa·CAT* fusion gene in skeletogenic mesenchyme cells at late gastrula stage results from deletion or mutation of the two adjacent sites at which the SpZ12-1 factor binds. Ectopic expression of the mutated form of this construct behaves exactly the same as does the wild-type construct, promoting CAT expression exclusively in the ectoderm. We believe that SpZ12-1 is the factor responsible for this late embryonic regulatory effect because it is the only factor present in *S. purpuratus* nuclear extract that binds to the relevant sites in the *CyIIIa* regulatory domain. This was demonstrated by using an antibody raised against a recombinant SpZ12-1 protein (Wang *et al.,* submitted). However, it is likely that SpZ12-1 has maternal and earlier embryonic as well as post-gastrular functions. Thus, as we show in the following, a large amount of this factor is present in the embryo from the beginning of development onward. In this communication we consider parameters relevant to the provenance of SpZ12-1 during development, *viz* the quantity of mRNA encoding this factor, and the distribution of SpZ12-1 DNA binding activity in nuclear and cytoplasmic compartment during early development.

Materials and methods

Probe excess RNA titration

The number of SpZ12-1 mRNA molecules in *S. purpuratus* eggs and embryos was measured by using single-stranded antisense RNA probe excess titration as described by Lee and Costlow (1987). The antisense RNA probe included the sequence from nucleotide 1 to 400 of the SpZ12-1 cDNA (Wang *et al.*, submitted). The probe was prepared with T7 polymerase (Ambion Inc.), and the specific activity of the probe was 1.5×10^9 dpm/µg. Hybridization reactions were performed in probe excess of 10-fold or greater, in 0.4 M NaCl, and 50% formamide at 50°C. The amount of RNA-RNA hybrid was assayed by trichloroacetic acid precipitation after RNase A and RNase T1 digestion. The radioactivity was quantitated in a scintillation counter.

Blastula-stage embryo cytoplasmic extract

Blastula-stage embryos were harvested at 21 hr after fertilization, washed in 1 M dextrose, resuspended in 5X the pellet volume of TEESSD (10 mM Tris, pH 7.4, 1 mM EGTA, 1 mM EDTA, I mM spermidine, 1 mM DTT, 0.36 M sucrose) and frozen in liquid nitrogen (Calzone *et al.*, 1988). The embryos were stored at -70°C. The protocol for preparing cytoplasmic extracts is according to F.J. Calzone (personal communication). Briefly, the embryos were thawed by mixing

with an equal volume of TEESSD at room temperature. Nuclei were removed from the homogenate by centrifugation at 3000 xg. Cytoplasmic DNA-binding proteins were extracted from the post-nuclear supernatants by increasing the KCl concentration to 300 mM, followed by centrifugation at 35,000 rpm for 3 hr at 4°C. The extracts were stored at -70°C after addition of glycerol to a final concentration of 10%. The number of embryos represented in the extract was derived from the amount of ribosomes in the 100,000 xg pellet, taking 60.2 OD_{260} per nanomole of ribosomes and $6.6x10^6$ ribosomes per embryo (Davidson, 1986, p. 160).

DNA binding assays

Gel shift assays were carried out as previously described (Calzone *et al.*, 1988). The probe was an *Rsa*I fragment of the CyIIIa *cis*-regulatory region (-821 to -1033) containing the SpZ12-1 binding sites. The binding reactions contained 20 mM HEPES (pH 7.9), 0.5 mM DTT, 75 mM KCl, 5 mM $MgCl_2$, 5 µg of poly[d(AT)], 3 µl of egg cytoplasmic extract (201 eggs/µl) or 1 µl of blastula-stage cytoplasmic extract (200 embryos/µl), and a one-fifth dilution of buffer C [buffer C is 20 mM HEPES (pH 7.9), 40 mM KCl, 0.1 mM EDTA, 20% glycerol, 0.1% NP-40]. The reactions were incubated for 15 min on ice and separated on 5% acrylamide gels in Tris-borate buffer.

Whole embryo lysate and antibody blots

Gametes of *S. purpuratus* were collected by intracoelomic injection of 0.5 M KCl. After fertilization, embryos were cultured in filtered seawater containing penicillin (20 units/liter) and streptomycin (50 µg/liter) at 15°C, and harvested at different developmental stages. Following a brief wash of ice-cold 1 M glucose, the embryos were pelleted and resuspended in 10 mM Tris-HCl (pH 7.4), 2 mM EDTA, and 1% SDS. The whole embryo lysate was briefly sonicated and stored at -70°C. Antibody blots were carried out according to Calzone *et al.* (1991) and Wang *et al.* (submitted).

Results and discussion

Provenance of SpZ12-1 mRNA

We measured the amount of SpZ12-1 mRNA in egg and embryos, by a single-stranded probe excess procedure (Lee *et al.*, 1986; Lee and Costlow, 1987). These measurements are displayed in Figure 1, and the results are summarized in Table 1. It is apparent that SpZ12-1 mRNA is present at a low and essentially unchanging level from before fertilization to the completion of embryogenesis. The amount of maternal and embryonic message encoding this factor throughout development is calculated to be about $1\text{-}2\text{x}10^3$ molecules/egg or embryo. This is a level typical of the rare message class of *S. purpuratus* embryos, for which the average prevalence is estimated at about $1.6\text{x}10^3$ molecules of each species per embryo (Davidson, 1986, pp. 52-53). If SpZ12-1 message behaves as is typical for messages of this class, the maternal SpZ12-1 message will have been replaced by zygotically transcribed SpZ12-1 mRNA by the blastula stage (Davidson, 1986, pp. 155-160), but at present we have no direct evidence on this point. We note in this connection that the gene encoding SpZ12-1 mRNA is single copy, as shown by genome blot hybridizations carried out on the DNAs of four different individual sea urchins (data not shown).

Provenance of SpZ12-1 protein

An antibody blot of total egg and embryo proteins is shown in Figure 2. It can be seen that the protein species reacting with this antibody displays the same mobility at all stages; and that this protein is present throughout embryogenesis. However, this observation is not quantitative, and no conclusions can be drawn from Figure 2 as to the relative amounts of SpZ12-1 that are present at different stages.

Calzone *et al.* (1988) measured the molecular prevalence of SpZ12-1 in nuclear extracts of late cleavage- and blastula-stage embryos by titration of its DNA binding activity (prior to cloning and characterization this factor was denoted "P6"). We carried out comparable experiments on *cytoplasmic* extracts of egg- and blastula-stage embryos. These experiments are illustrated in Figure 3, which portray measurements of the DNA-SpZ12-1 complex, as assayed in gel shifts, formed at increasing ratio of DNA probe to protein. At completion of the titration the amount of complex formed provides a measure of the molecular content of the specific DNA-binding factor in the aliquot of extract (the data are present in Figure 3 in linear reciprocal plots for ease of obtaining the least-squares solutions). In addition to estimates of the quantities of SpZ12-1 factor active in complex formation (**Po**), we also extracted the relative equilibrium constant (**Kr**) for the reacting species. This parameter is useful as a "signature"

of the SpZ12-1 reaction, as this factor has an unusually high target site specificity (*cf.* Calzone *et al.*, 1988).

Our results and those of Calzone *et al.* (1988) are combined in Table 2, in order to provide a direct comparison of nuclear and cytoplasmic contents of SpZ12-1, i.e., of SpZ12-1 molecules competent to generate DNA-protein complexes. Inactive factor, whether of endogenous or artifactual (i.e., preparative) origin, will of course not be counted. Furthermore, these measurements obviously do not indicate the activity of the factors in respect to transcriptional regulation, since they refer only to its ability to recognize and bind to its specific DNA target site.

We also include in Table 2 a calculation of the amount of newly synthesized SpZ12-1 protein that could have been produced by embryonic translation, based on the average quantity of SpZ12-1 mRNA present over the period 0-24 hr, i.e., ~$2x10^3$ molecules/egg or embryo (Table 1). For this calculation we assume a maximum translation rate of about two molecules of protein produced per mRNA·min^{-1} (i.e., ~1.8 codon·sec^{-1} and 150 nt center-to-center ribosome species; see Davidson, 1986, pp. 77-78); we also assume for the purpose of the calculation complete stability of the protein product over 24 hr. It is of course of no moment for this calculation whether the mRNA is stable or not, or whether it is maternal or zygotic in origin.

Conclusions

Considering the **Kr** data of Table 2 together with the qualitative immunological identification shown in Figure 2, it seems clear that qualitatively the same SpZ12-1 protein with the same specific DNA binding activity is present in the unfertilized egg and throughout early development, in both nuclear and cytoplasmic compartments. Table 2 shows, furthermore, that the amount of SpZ12-1 in the unfertilized egg is significant. Thus the maternal content of this factor, estimated from DNA binding activity, is at least equal to the total complement of SpZ12-1 in the nuclei of the embryo at the 24 hr blastula stage. Assuming equal distribution, there are $>2x10^3$ molecules per nucleus in the 24 hr embryo, more than sufficient to ensure target site occupancy given the high **Kr** characteristic of SpZ12-1 (Calzone *et al.*, 1988, 1991).

The SpZ12-1 in the embryo is unlikely to be of exclusively maternal origin. Table 2 also shows that by 24 hr all of the SpZ12-1 protein in the embryo could easily have been synthesized after fertilization. This mRNA clearly belongs to the rare message class, since in the 24 hr blastula its average prevalence is only about five molecules/cell. Nonetheless it would suffice by 24 hours to generate almost twice the total amount of SpZ12-1 protein estimated from DNA binding to be present in the embryo. Unfortunately, the concentration of this mRNA is too low to detect by whole mount *in situ* hybridization and in respect to its spatial distribution we know only that this message is not concentrated in a small

fraction of embryonic cells (where it would have been differentially detectable). Even as early as 7 hr all the SpZ12-1 DNA binding activity in the embryo could be accounted for as newly translated protein (Table 2). It follows that after the very first few hours post-fertilization, the nuclear SpZ12-1 protein, i.e., the component active in gene regulation, could be entirely either of maternal or zygotic origin, or in part maternal and in part zygotic.

The 7 hr nuclear extract measurement of Calzone *et al.* (1988) combined with the cytoplasmic **Po** measurements of this report provides an interesting conclusion. This is that only a few percent of the SpZ12-1 protein evidently is nuclear in location, while in 24 hr embryos about 50% is nuclear. Though we did not directly measure cytoplasmic SpZ12-1 in 7 hr embryos in this work, a glance at Figure 2 excludes the alternative possibility that at this stage the total SpZ12-1 content is much lower, so that the nuclear compartment would still contain a large fraction of it. Transit of SpZ12-1 to the nucleus is therefore likely to serve as a major determinant of its functional activity. From these data it is evident that the transit process is temporally controlled; it could of course also be spatially regulated. Furthermore, from its low nuclear prevalence at 7 hr post-fertilization, it is unlikely that SpZ12-1 is utilized for gene regulation in cleavage-stage embryos.

Though we are as yet unable to provide data distinguishing this from other scenarios, a reasonable interpretation would be as follows. SpZ12-1 fulfills

some unknown role in oogenesis which requires the significant maternal concentration we observe. It must also be required by mesenchyme blastula stage, since by this point a large fraction of the factor is found in the nuclei. The post-gastrular spatial repression function of SpZ12-1 reported by Wang *et al.* (submitted) would represent a still later phase of its activity, and by this time at least the functional protein is likely to be largely of zygotic origin. In late embryos the spatial activity of SpZ12-1 could be controlled at many different levels, from regulated transcription to regulated post-transcriptional modification. We note in this connection that high resolution 2D gels display at least two different isoforms of SpZ12-1 (Harrington and Wang, unpublished data). The developmental localization may illuminate some of the issues we raise in this report.

Acknowledgments

We thank Carmen V. Kirchhamer and Kirsten L. Rood for critically editing a draft of the manuscript, and Vikas Duvvuri for technical assistance. This research was supported by NIH Grant HD-05753. D.G.W. was supported by a McCallum Fellowship, California Biochemical Foundation and NIH Grant HD-05753.

References

Calzone, F.J., N. Thézé, P. Thiebaud, R.L. Hill, R.L. Britten and E.H. Davidson. 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.

Calzone, F.J., C. Höög, D.B. Teplow, A.E. Cutting, R.W. Zeller, R.J. Britten and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo I. Purification by affinity chromatography and cloning of P3A2, a novel DNA-binding protein. *Development* **112**: 335-350.

Davidson, E.H. 1986. "Gene Activity in Early Development" 3rd ed., Academic Press, Orlando, FL.

Emerson, B.M., C.D. Lewis and G. Felsenfeld. 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β-globin gene: nature of the binding domain. *Cell* **41**: 21-30.

Goustin, A.S. and F.H. Wilt. 1981. Protein synthesis, polyribosomes, and peptide elongation in early development of *Strongylocentrotus purpuratus. Dev. Biol.* **82**: 32-40.

Lee, J.J., F.J. Calzone, R.J. Britten, R.C. Angerer and E.H. Davidson. 1986. Activation of sea urchin actin genes during embryogenesis. Measurement of transcript accumulation from five different genes in *Strongylocentrotus purpuratus. J. Mol. Biol.* 188: 173-183.

Lee, J.J. and N.A. Costlow. 1987. A molecular titration assay to measure transcript prevalence levels. *Meth. Enz.* **152**: 633-648.

Wang, D.G., C.V. Kirchhamer, R.J. Britten and E.H. Davidson. 1994. SpZ12-1, a negative regulator required for spatial control of the territory-specific *CyIIIa* gene in the sea urchin embryo. Submitted for publication.

Table 1. Prevalence of SpZ12-1 transcripts during development

1 Stage	2 cpm hybridization / µg of RNA[a]	3 mRNA per embryo (pg)[b]	4 Number of transcripts per embryo[c]
Egg	947.6 (0.98)	1.1×10^{-2}	2.3×10^{3}
4 hr	882.4 (0.99)	1.1×10^{-2}	2.1×10^{3}
12 hr	643.1 (0.99)	7.5×10^{-3}	1.6×10^{3}
24 hr	798.9 (0.95)	9.4×10^{-3}	1.9×10^{3}
48 hr	482.5 (0.93)	5.7×10^{-3}	1.2×10^{3}
54 hr	467.2 (0.97)	5.5×10^{-3}	1.1×10^{3}
72 hr	674.7 (0.93)	7.9×10^{-3}	1.6×10^{3}

[a]Slopes (column 2) of the titration curves in Figure 1 were determined by a linear least-squares analysis of the data points. The differential accuracy of each titration was estimated by the correlation coefficient as indicated.

[b]Mass of SpZ12-1 mRNA (pg)/embryo = $xy/\alpha\beta$: where x is cpm of probe/µg of total RNA (slope of titration curve); y is mass of total RNA/embryo, here taken as 2.8 ng (Goustin and Wilt, 1981); α is specific activity of the antisense SpZ12-1 RNA probe (dpm/µg); β is fraction of SpZ12-1 mRNA represented on the probe

(440 nt/2700 nt = 0.16). The size of the transcripts is established in RNA gel blot experiments (data not shown) which also demonstrated that only a single transcript, of the same size, is present throughout development.

[c]The number of transcripts/embryo (column 4) is derived by dividing pg of SpZ12-1 mRNA/embryo (column 3) by the molecular weight of the SpZ12-1 mRNA.

Table 2. Provenance of SpZ12-1 protein from DNA-binding measurements

Source	Kr	Po (molecules/egg or embryo)
egg cytoplasm	1.3×10^6 +	1.8×10^6 +
7 hr late cleavage		
nucleus	4.9×10^5 •	4×10^4 •
cytoplasm		(1.8×10^6)[a]
new translation		(1.7×10^6)[b]
24 hr blastula		
nucleus	1.4×10^6 •	1.4×10^6 •
cytoplasm	1.3×10^6 +	1.7×10^6 +
total		3.1×10^6
new translation		(5.7×10^6)[b]

Number in parentheses are calculated or assumed on other bases, as below and in text.

•Data from Calzone *et al.* (1988).

+Data averaged from duplicate experiments such as those in Figure 3.

[a]Assumed to be the same as egg cytoplasm and blastula cytoplasm (*cf.* Figure 2).

[b]Based on parameters given in text: these are *maximum* estimates because maximum translation rate and no protein turnover are assumed.

Figure. 1. Single-stranded probe excess titration of SpZ12-1 transcripts in egg and embryo RNA. Measurements were carried out on total RNA prepared from eggs and embryos (4 hr, 12 hr, 24 hr, 48 hr, 54 hr, and 72 hr) as indicated. Each panel represents probe hybridization (ordinate) as a function of increasing amounts of egg or embryo RNA (abscissa). A ^{32}P-labeled antisense SpZ12-1 RNA probe was generated with T7 polymerase. The lines represent linear least-squares solutions. The reduced data from these measurements are listed in Table 1.

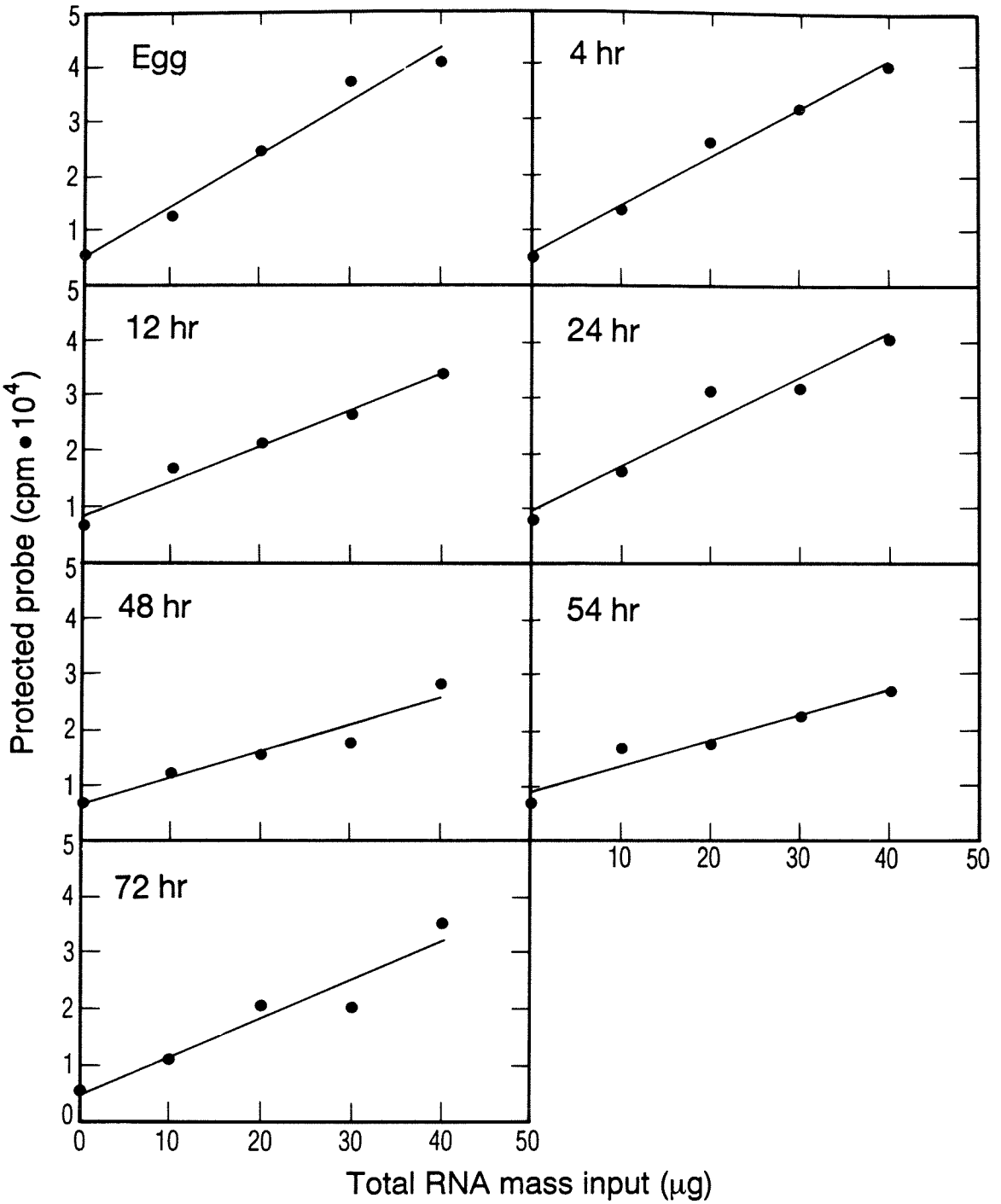

Egg
4 hr
12 hr
24 hr
48 hr
54 hr
72 hr
Protected probe (cpm • 10^4)
Total RNA mass input (μg)
0
1
2
3
4
5
10
20
30
40
50

Figure. 2. Immunological identification of SpZ12-1 protein during development. Lysates of about 1,000 eggs or embryos (per lane) were prepared from egg, 7 hr, 12 hr, 24 hr, 32 hr, 48 hr, 57 hr and 72 hr embryos, displayed by electrophoresis on a 7% SDS-polyacrylamide gel, and transferred to nitrocellulose. As described in Materials and Methods, the filters were then reacted with a polyclonal antibody that had been raised against recombinant SpZ12-1 protein (Wang *et al.,* submitted). The reactant species, as found earlier, is 70 kd, established here by reference to markers (not shown).

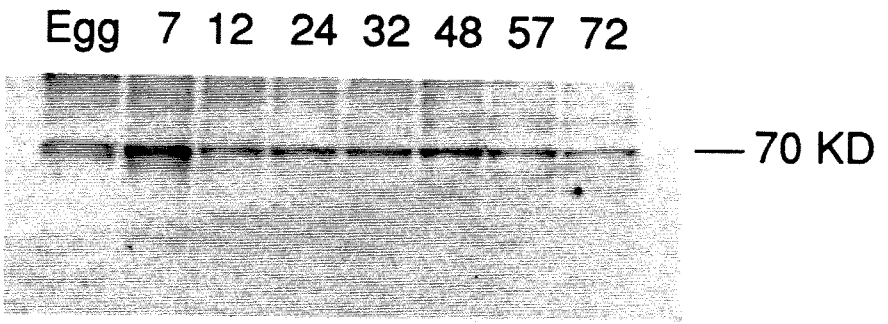
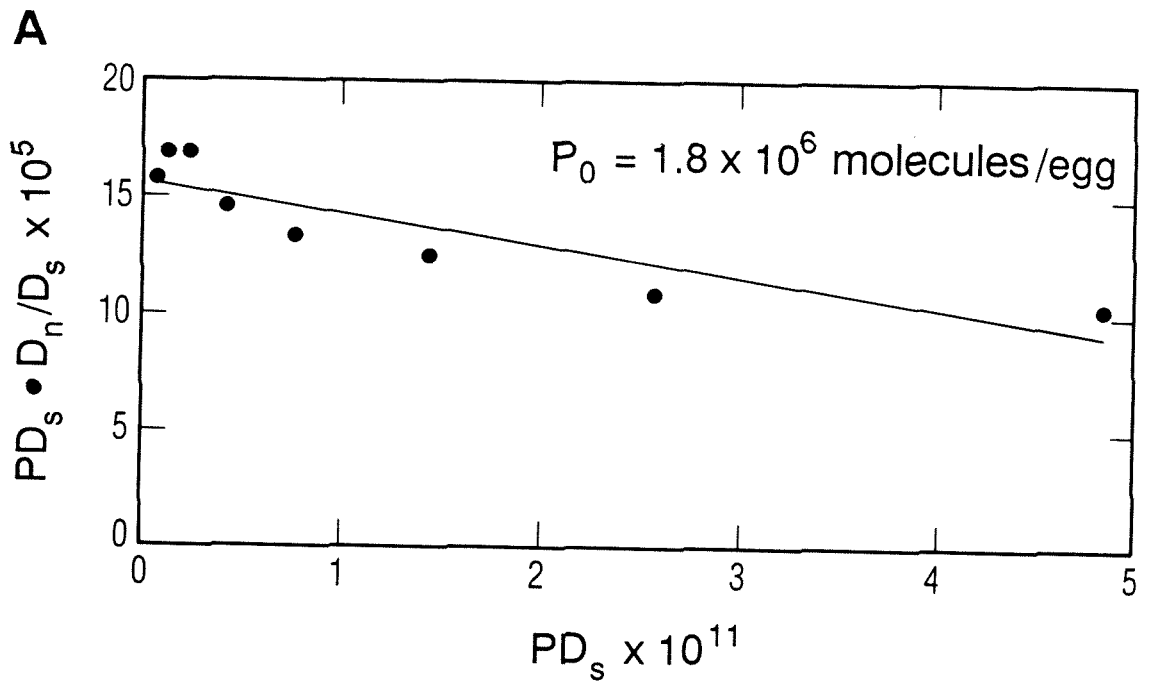
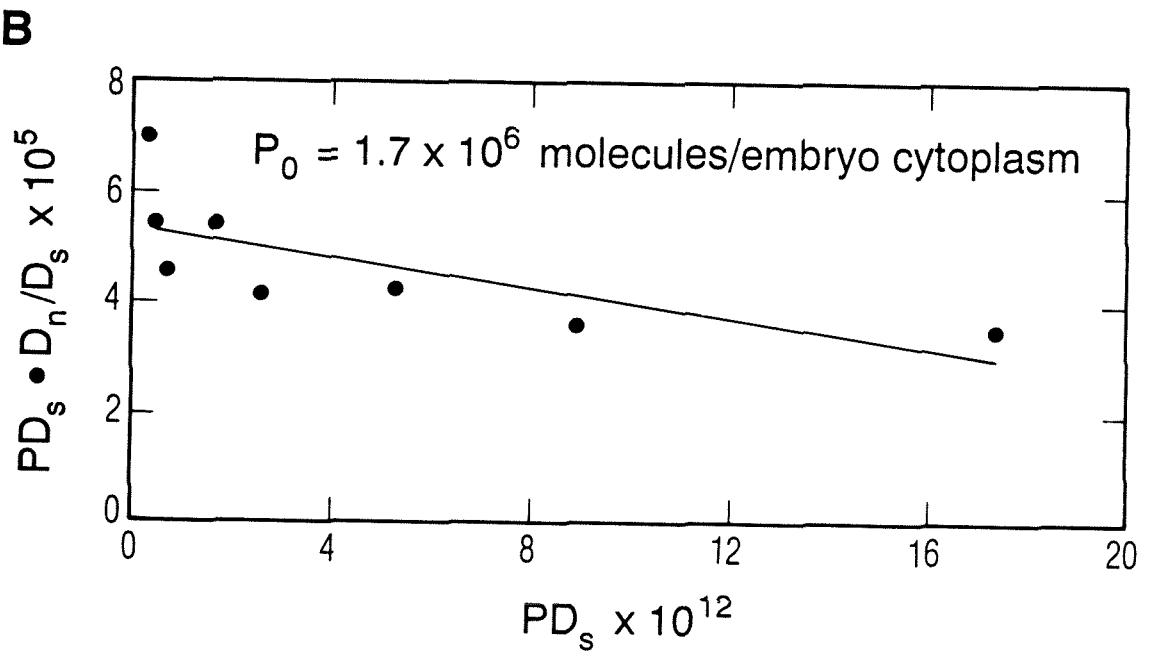

Egg 7 12 24 32 48 57 72
— 70 KD

Figure. 3. Measurements of cytoplasmic SpZ12-1 proteins active in binding a probe containing its specific target site, and of the relative equilibrium constant of the reactive species. All reactions in each experiment contained the same quantity of extract (see Materials and Methods for extracts and procedures). The amount of labeled probe was varied, and the product of the interaction was separated from free probe by native gel electrophoresis (only a single mobility complex is formed by this factor; Calzone *et al.*, 1988; Wang *et al.*, submitted). The complexes were excised and counted, and values of **Po** (molecules of active factor per egg or embryo in the respective cytoplasmic extract) and **Kr** (relative equilibrium constant) were calculated. **Kr** is the ratio of the equilibrium constant for the specific interaction of the factor with its DNA target site to the equilibrium constant for the non-specific interaction of the factor with synthetic double-stranded DNA polymer. These constants were extracted by least squares from the data shown, as earlier (Calzone *et al.*, 1988, 1991), using the form **(PDs X Dn) / Ds = Kr(Po - Pds)** (Emerson *et al.*, 1985), where **Ds** is free probe; **PDs** is complex; **Po** is total active protein in the reaction; and **Dn** is non-specific competitor, all in molar concentration. (**A**) Egg cytoplasmic extract; (**B**) blastula cytoplasmic extract.


A
P_0 = 1.8 x 10^6 molecules/egg
$PD_s \bullet D_n/D_s$ x 10^5
PD_s x 10^{11}
0
5
10
15
20
0
1
2
3
4
5


B
P_0 = 1.7 x 10^6 molecules/embryo cytoplasm
$PD_s \bullet D_n/D_s$ x 10^5
PD_s x 10^{12}
0
2
4
6
8
0
4
8
12
16
20

CHAPTER 3

Conclusions

Brief Summary

A negative *cis*-element within the *CyIIIa* regulatory region was identified that is required for repression of a *CyIIIa·CAT* fusion gene in skeletogenic mesenchyme cells at the late gastrula stage. The repression mediated by this element takes place after initial territorial specification. A full-length cDNA clone encoding the corresponding DNA-binding protein (SpZ12-1) of about 70 kd was isolated from a embryo expression library with a DNA probe containing this *cis*-element. A remarkable feature of the protein is that it contains 12 tandem zinc fingers of the TFIIIA type near the carboxyl terminus which is required for the DNA-binding activity. SpZ12-1 is the only DNA-binding protein in the embryo nuclear extract that specifically interacts with this *cis*-element and is likely to be the *trans* factor responsible for the negative spatial regulatory function *in vivo*. It was shown that both SpZ12-1 mRNA and protein are of maternal origin. The amount of maternal and embryonic SpZ12-1 mRNA remains low and essentially constant at about $1\text{-}2\text{x}10^3$ molecules per egg or embryo during embryogenesis. Translation of this mRNA could easily account for the complete complement of SpZ12-1 protein in the embryo, as estimated from its DNA-binding activity. SpZ12-1 is present in significant quantities in unfertilized egg cytoplasm, and in similar quantities in mesenchyme blastula-stage embryo cytoplasm. Taken together with earlier measurements of nuclear SpZ12-1 (Calzone *et al.*, 1988), our observations indicate that SpZ12-1 enters the embryonic nuclei between late cleavage and mesenchyme blastula stage. The nuclear localization of SpZ12-1 could be a determinant of its functional activity.

SpZ12-1 and the CyIIIa Cis-Regulatory System

The *CyIIIa cis*-regulatory region is organized into three modules and each module has separable regulatory functions during development. An interesting property of the modular organization is that the proximal and middle modules are controlled by both positive and negative spatial *trans* factors possibly interacting with each other locally (Kirchhamer *et al.,* unpublished; Chapter 1). The middle module contains two negative spatial *cis*-elements, the SpZ12-1 and SpP7II target sites, which are required for correct spatial expression of a *CyIIIa·CAT* fusion gene after gastrulation. An important clue to the interpretation of the skeletogenic mesenchyme repression mediated by SpZ12-1 is that a positively-acting transcription factor, SpRunt, binds to a site near the SpZ12-1 sites in the same module. This observation suggests one possible scenario: during gastrulation the *CyIIIa* gene begins to respond to a positive regulator, possibly SpRunt, which functions in a broad spatial region of the late embryo including the skeletogenic mesenchyme, so the late mesenchyme repression mediated by SpZ12-1 might confine the effect of SpRunt to the aboral ectoderm. Since our model predicts that SpRunt is at least present in both aboral ectoderm and skeletogenic mesenchyme after initial territorial specification, one crucial experiment to test this hypothesis is to examine the spatial distribution of SpRunt at relevant developmental stages by using whole-mount antibody staining.

Prospects for Future Research

I) Developmental localization of SpZ12-1

Since in late embryos the spatial activity of SpZ12-1 could be controlled at many different levels, from regulated transcription to regulated post-translational modification, analysis of the spatial expression of *SpZ12-1* would be very useful to eliminate alternative mechanisms underlying the developmental regulation of functional activity of *SpZ12-1*. In particular, this issue can be approached by comparing the expression patterns of SpZ12-1 mRNA and protein at different developmental stages by using whole-mount *in situ* hybridization and antibody staining procedures, respectively.

ii) Biochemical analysis of post-translational modification of SpZ12-1

Another area that needs to be explored is the nucleus-cytoplasm distribution of different isoforms of SpZ12-1 during development and the relevance of its regulatory function. High resolution 2D analysis can be initially used to identify potential post-translational modifications of SpZ12-1 in both nuclear and cytoplasmic compartments. This search will eventually lead us to relate particular isoforms of SpZ12-1 in the stage and region of the embryo to its regulatory function.

iii) Biochemical analysis of the regulated assembly of multiprotein complexes on the CyIIIa cis-regulatory region

A major unresolved question regarding the *CyIIIa cis*-regulatory system is how to relate the interactions among various *trans*-regulatory factors on the *CyIIIa* regulatory region to its temporal and spatial expression in the embryo. Since most of the *CyIIIa trans* factors have been cloned and characterized, it is now possible to search potential partners that interact with each other *in vitro*, particularly partners interacting with SpZ12-1. Recent biochemical studies have shown that several factors were copurified by DNA affinity chromatography, implying that specific protein-protein interactions take place on the *CyIIIa* regulatory region. These studies will provide some mechanistic insights into the problem of how the *CyIIIa trans* factors assemble *in vitro*. It may be premature to conclude that the interactions detected *in vitro* are important for mediating transcriptional regulation in the embryo. However, this search will aid us to formulate hypotheses and to systematically test them *in vivo*.

References

Calzone, F.J., N. Thézé, P. Thiebaud, R.L. Hill, R.J. Britten and E.H. Davidson. 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.

Hough-Evans, B.R., R.R. Franks, R.W. Zeller, R.J. Britten and E.H. Davidson. 1990. Negative spatial regulation of the lineage-specific *CyIIIa* actin gene in the sea urchin embryo. *Development.* **110**: 41-50.