*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 2989–2993, March 1995
Developmental Biology

# A multimerizing transcription factor of sea urchin embryos capable of looping DNA

ROBERT W. ZELLER*†, JACK D. GRIFFITH‡, JAMES G. MOORE*, CARMEN V. KIRCHHAMER*, ROY J. BRITTEN*, AND ERIC H. DAVIDSON*

*Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125; and ‡Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599-7295

*Contributed by Eric H. Davidson, December 20, 1994*

**ABSTRACT** **SpGCF1 is a recently cloned sea urchin transcription factor that recognizes target sites in several different sea urchin genes. We find that in gel-shift experiments this factor is able to multimerize. A quantitative simulation of the gel-shift results suggests that SpGCF1 molecules that are bound to DNA target sites may also bind to one another, thus associating several DNA probe molecules. SpGCF1 might therefore be able to loop DNA molecules bearing its target sites at distant locations. We demonstrate this prediction by electron microscopy, and using the well-characterized cis-regulatory domain of the *CyIIIa* cytoskeletal actin gene, we show that the loop conformations predicted from the known SpGCF1 target site locations are actually formed *in vitro*. We speculate that the multimerization of this factor *in vivo* may function to bring distant regions of extended regulatory domains into immediate proximity so that they can interact with one another.**

SpGCF1 is a sea urchin embryo transcription factor that binds tightly to certain target site sequences surrounding a core CCCCN sequence, where N is usually T. Most of our knowledge of the function of this protein *in vivo* derives from studies on the *Strongylocentrotus purpuratus CyIIIa* cytoskeletal actin gene, which is expressed exclusively in the aboral ectoderm of the embryo and larva. As shown in Fig. 1, the distal end of this regulatory domain (5–8) includes a cluster of SpGCF1 sites (cluster A). If these sites are deleted from a transgene construct consisting of the 5′ leader and upstream sequences of *CyIIIa* ligated to the bacterial chloramphenicol acetyltransferase (CAT) gene (9), expression of the *CyIIIa-CAT* transgene is sharply reduced. The same result is obtained if the factor is sequestered away from the *CyIIIa-CAT* fusion gene by cointroduction of excess target sites—i.e., *in vivo* competition (10, 11)—though this does not affect the spatial domain of the remaining expression. Similar *in vivo* competition results were obtained with the SpGCF1 sites of region E of the CyIIIa regulatory domain (Fig. 1; refs. 10 and 11). In the *CyIIIa* gene, these sites mediate a positive transcriptional activation function. A protein binding specifically to these sites, called SpGCF1, was purified from embryo nuclear extracts and cloned by Zeller *et al.* (12), though it remains possible that other proteins recognizing these sites could also be present in the nuclei. Although in sequence SpGCF1 resembles no other known protein, a specific DNA-binding domain has been identified experimentally. The N-terminal region of SpGCF1 contains a high concentration of proline residues, a feature associated with a transcriptional activation function in some other known transcription factors (12–15). A particularly interesting characteristic of SpGCF1 is that it is present in five different forms, with molecular masses of 37, 40, 43, 50, and 55 kDa. These forms differ in that they include increasing extents

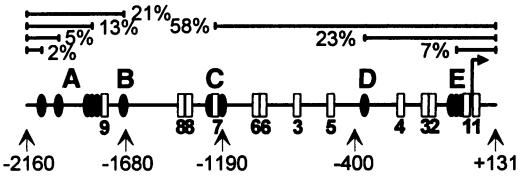


FIG. 1. Locations of SpGCF1 binding sites in the 2.3-kb *CyIIIa* regulatory domain. Binding sites for SpGCF1 protein (from unpublished data and refs. 1–3) are shown as filled ovals. Open boxes indicate the positions of other DNA-binding protein target sites: 1, P1 site; 2, SpOct1 site (4); 3, SpP3A site; 4, SpTEF1 site; 5, SpCTF1 site; 6, the pair of SpZ12-1 sites (5); 7, SpP7II site; 8, Sprunt1 site; and 9, newly discovered site designated the P9 site. The arrow indicates the start of transcription. The regulatory region displayed here extends from −2160 to +131. The percentage values indicate the distance from the various SpGCF1 sites to the ends of the DNA molecule.

of the N-terminal proline-rich domain (12). All five forms are translated from the same mRNA by alternative utilization of one of five in-frame ATG codons (12).

For some time we have been puzzled by the prevalence and distribution of SpGCF1 target sites in the cis-regulatory domain of both *CyIIIa* and another gene that we have been interested in, the *Endo16* gene. *Endo16* is expressed in the vegetal plate and archenteron of the embryo (16, 17). The ≈2500-bp cis-regulatory domain of *Endo16* contains >20 distinct sites where SpGCF1 binds specifically. As in *CyIIIa* (Fig. 1) these sites usually occur in clusters and often lie close to sites where other transcription factors bind. SpGCF1 sites also appear in the regulatory domain of *SM30*, a skeletogenic gene (18). Thus, this factor appears to interact with target sites in genes expressed in at least three different embryonic territories, namely aboral ectoderm, vegetal plate, and skeletogenic territories. We measured the relative equilibrium constants for the interactions of SpGCF1 with three of its *CyIIIa* target sites and found (1, 12) that the preference of this protein for these particular sites [with reference to poly(dA·dT)] ranges from $5 \times 10^4$ to $\approx 10^6$. SpGCF1 is therefore a DNA-binding protein which recognizes its target sites with high specificity. In this communication we describe a new functional characteristic of SpGCF1, namely its ability to multimerize with itself. Perhaps this feature explains both the distribution and prevalence of SpGCF1 sites: interactions between bound SpGCF1 molecules might have the function of bringing into immediate proximity other transcription factors bound at distant locations within the regulatory domain so that these factors are able to form functional complexes with one another. This would imply that the DNA sequence between interacting SpGCF1 molecules can be looped out. Though we have no evidence for such conformations *in vivo*, we show that the expected DNA loops indeed form in a purified system in

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

Abbreviation: CAT, chloramphenicol acetyltransferase.

†Present address: Department of Biology, University of California, San Diego, La Jolla, CA 92093-0347.

which the only macromolecular components are recombinant SpGCF1 and the *CyIIIa* regulatory DNA sequence represented in Fig. 1.

## MATERIALS AND METHODS

**Protein Production and Gel Mobility-Shift Assays.** Recombinant SpGCF1 protein was expressed in *Escherichia coli* and renatured as described (12). Gel mobility-shift assays were performed essentially as described (1). The target-site probe used in the gel mobility shifts was a 44-bp fragment of the *CyIIIa* regulatory DNA located between −2027 and −1983 in the sequence (2). It was labeled by the end–fill reaction (1). The concentration of active SpGCF1, determined by excess probe titration (1, 12) to be 5.8 nM, was divided among the five mass variants. This represents about 1% of the total recombinant SpGCF1 in the preparation. Aliquots (4 μl) containing various amounts of the protein were dispensed into tubes on ice. To each of these tubes was added 16 μl of a premixed solution of the probe (15–50 pM) and 1× binding buffer [83 mM KCl/5 mM $MgCl_2$/20 μM EDTA/700 μM dithiothreitol/14 mM Hepes·KOH, pH 7.9/2% (vol/vol) glycerol]. The complexes were allowed to form on ice for 10–15 min, and were then loaded directly onto an 8% polyacrylamide gel (30:0.8 acrylamide:bisacrylamide) previously prerun for an hour. After electrophoresis at 250 V, the gel was dried, exposed to a Kodak imaging screen, and subsequently analyzed in a Molecular Dynamics PhosphorImager. Data were converted to cpm by reference to a known probe sample that was counted in both the scintillation counter and the PhosphorImager.

**Electron Microscopy of DNA–Protein Complexes.** Samples were prepared for EM by incubating 1 ng of the 2.3-kb *CyIIIa* regulatory region—i.e., positions −2160 to +131—with SpGCF1 protein for 10 min on ice in 1× binding buffer (final volume, 25 μl; final protein concentration, 1.75 nM). Following incubation, the sample was treated with 0.6% glutaraldehyde for 15 min on ice, then chromatographed over 2 ml Bio-Gel A-5m (Bio-Rad) columns equilibrated in 10 mM Tris·HCl/0.1 mM EDTA, pH 7.5. The peak fraction containing the DNA–protein complexes was treated as described (19). In brief, thin carbon foils supported by 400-mesh copper grids were exposed to a glow discharge, and the complexes were adsorbed for 30 sec in a buffer containing 2 mM spermidine and 150 mM NaCl. The grids were then rinsed with sequential washes of water and a graded series of ethanol solutions up to 100% ethanol. The samples were air dried and rotary shadowcast with tungsten at $10^{-7}$ torr (1 torr = 133.3 Pa).

Images were recorded on sheet or 35-mm film by using a Philips CM12 electron microscope. Images for analysis or publication were captured from the film by using a Cohu CCD camera and a Data Translation Quick Capture board in a Macintosh computer. The NIH IMAGE software was used to adjust the contrast in the images for publication. Images were transferred to 35-mm film by using a GCC (San Jose, CA) film recorder. Lengths of DNA segments on the micrographs were

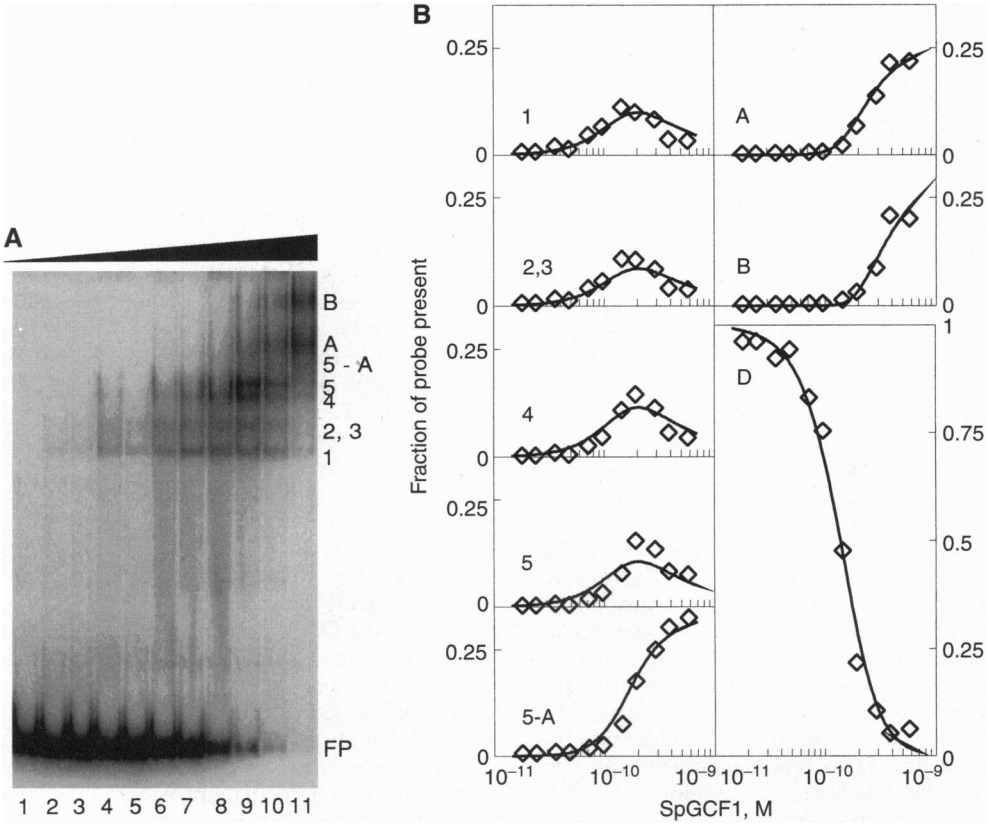


FIG. 2. Quantitative protein excess gel-mobility assay. (*A*) PhosphorImager report of the labeled complexes formed. A synthetic DNA probe bearing a single SpGCF1 target site was incubated with increasing amounts of recombinant SpGCF1 protein. The complexes formed were resolved by electrophoresis through an 8% polyacrylamide gel. The amount of added protein increases from left (lane 1, $1.8 \times 10^{-11}$ M) to right (lane 11, $5.8 \times 10^{-10}$ M). Bands 1–5 (denoted on the right) indicate the complexes formed by the different SpGCF1 variants binding to the probe (see text). Region 5-A and bands A and B are complexes formed by SpGCF1/SpGCF1 interactions, as described in text. FP, free probe. (*B*) Curves generated by the dimer model for SpGCF1 complex formation discussed in text and in *Appendix*. The solid lines show the functions generated by using the least squares solution for the parameters of the model for these data. Each panel represents one of the bands figured in *A*, as indicated. The data from the PhosphorImager analysis are indicated as open symbols after normalization, so that the probe present in each complex is expressed as a fraction of total probe in the experiment. D represents free probe.

measured on the photographic prints by using a Summagraphics (Fairfield, CT) digitizing tablet coupled to a CompuAdd computer programmed with software developed in the laboratory of J.D.G.

## RESULTS

**Multimerization of SpGCF1 on a Single-Site Probe.** An example of a gel-shift experiment that reveals the multimerization of SpGCF1 is shown in Fig. 2*A*. Increasing amounts of renatured recombinant SpGCF1 were mixed with a 44-nt oligonucleotide probe (see *Materials and Methods*) that contained a single target binding site. The bands formed at the lower protein concentrations (bands 1–5; Fig. 2*A*, lanes 1–7) represent complexes formed, respectively, by the 37-, 40- 43-, 50-, and 55-kDa variants of the protein. This was shown earlier (12) by the use of two-dimensional gel electrophoresis and immunological identification of the complexes with a polyclonal antibody that recognizes the C-terminal region of the protein that is common to all five variants. As protein concentration is increased, more slowly migrating complexes appear. A blurred region of the gel is seen above the band containing the complex with the 55-kDa variant in lanes 8–11 (region 5-A), representing a series of complexes that are too close to one another in size to be resolved into discrete bands. Above this region, two more sharply resolved sets of complexes (bands A and B in Fig. 2*A*) appear at the highest protein concentrations. Since there is only a single target site on the probe, and since no proteins other than the purified recombinant SpGCF1 are present in the experiment, all the complexes migrating more slowly than that formed by the 55-kDa variant must be due to SpGCF1/SpGCF1 interactions that involve SpGCF1 molecules that are also bound to the labeled DNA probe. The same set of slowly migrating complexes appears at high protein levels when natural SpGCF1 concentrated from nuclear extract by affinity chromatography is reacted with a *CyIIIa* probe (data not shown). Therefore, the slowly migrating complexes illustrated in Fig. 2*A* are not due to nonrenatured recombinant SpGCF1 that cannot bind DNA, which is present in large amounts in the protein preparations we used for this experiment.

In order to interpret the results of experiments such as shown in Fig. 2*A*, we constructed models of the presumed interactions based on simple equilibrium considerations (see *Appendix*). These models allowed us to test a number of different stoichiometries. The best fits to the data were obtained with what we termed the dimer model and are shown in Fig. 2*B*. The assumptions on which this model is based and the number of SpGCF1 molecules and DNA probe molecules assumed in this model for each of the complexes observed are detailed in *Appendix*. Very briefly, the compositions of the individual bands in the dimer model are as follows: Bands 1–5 are formed of homodimers of the SpGCF1 size variants complexed to a single DNA probe molecule—i.e., the 37-kDa variant dimerizes only with another 37-kDa variant, the 55-kDa variant dimerizes only with itself, and so forth. The steepness with which these complexes accumulate with increasing protein concentration excludes monomeric complexes. (Cooperativity in DNA binding is not an issue since there is only a single target site per probe.) Band 5-A in the dimer model consists of unresolved complexes formed between monomers of any of the SpGCF1 variants and any of the DNA/$(SpGCF1)_2$ complexes represented in bands 1–5. Region 5-A complexes result from SpGCF1/SpGCF1 interactions that occur at higher protein concentrations than does formation of the initial complexes with DNA. Band A in this model consists of dimers of the $(SpGCF1)_3$/DNA complexes represented in Region 5-A—i.e., band A complexes contain six SpGCF1 molecules and two DNA molecules per complex. Band B consists of dimers of the complexes formed in band A—i.e., band B complexes consist of twelve SpGCF1 molecules and four DNA probe molecules per complex. Thus both band A and band B consist of complexes in which DNA probe molecules already complexed to SpGCF1 dimers are brought into a higher order multimeric complex by further SpGCF1/SpGCF1 interactions.

**SpGCF1 Loops DNA That Contains Multiple Target Sites.** If SpGCF1 molecules already bound to their DNA target sites are capable of binding specifically to one another, as in the interpretation of Fig. 2*A* just discussed, then such interactions should occur between SpGCF1 molecules bound at target sites at distant locations on the same DNA molecule. This would result in the formation of a DNA loop held together by SpGCF1 complexes that could be visualized by EM (20–22). We utilized the 2.3-kb *CyIIIa* regulatory domain DNA fragment illustrated in Fig. 1 for a direct test of this prediction and a 600-nt control fragment that included only the "E" SpGCF1 cluster of target sites. The latter should display protein–DNA complexes but no loops, and this is indeed what was observed (data not shown). However, when mixed with SpGCF1 protein, the 2.3-kb *CyIIIa* regulatory DNA fragment formed loops at a frequency of over 60% (see *Materials and Methods* for conditions).

To analyze these loop structures the distances between the ends of the DNA molecules and the bound SpGCF1 protein at the base of the loop(s) were measured. As Fig. 1 indicates, the SpGCF1 target sites and site clusters are located in such a fashion that the loops which could be formed by interactions at any two of these sites will produce ends of different lengths that can be identified unambiguously. Table 1 records the different forms of loops that are thus recognized, and categorizes them according to the site nomenclature of Fig. 1. Table 1 includes data from 59 loop structures that were clearly visualized and that we examined in detail. Examples of many of these categories of loop structure that we saw are shown in Fig. 3. As indicated in Table 1, most of the possible binary interactions that could occur were in fact included in the sample of 59 structures analyzed. The frequency of given loops as opposed to others reflects the relative magnitude of the equilibrium constants for the target sites at which they form, at least in those cases for which we have data. For example,

Table 1. Categories of loops in sample of 59 structures

| No. of loops per molecule | Interaction(s) | No. of occurrences |
|---|---|---|
| 1 (71.2%) | AA | 2 |
| | AB | 4 |
| | AC | 9 |
| | AD | 9 |
| | AE | 3 |
| | BC | 5 |
| | BD | 0 |
| | BE | 1 |
| | CD | 4 |
| | CE | 2 |
| | DE | 3 |
| 2 (22.0%) | AAC | 2 |
| | ABC | 1 |
| | ACD | 2 |
| | ACE | 3 |
| | BCD | 3 |
| | BCE | 1 |
| | CDE | 1 |
| 3 (5.1%) | AABD | 1 |
| | ABCD | 1 |
| | ABCF | 1 |
| 5 (1.7%) | AACDEE | 1 |
| Total | | 59 |

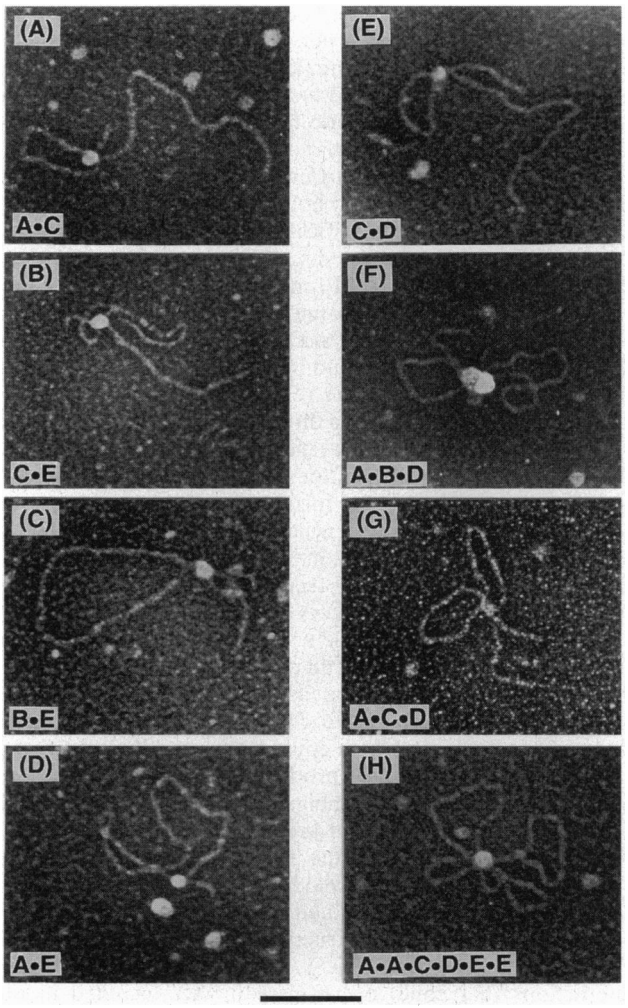


FIG. 3. Visualization of loops formed on *CyIIIa* DNA by interactions between bound SpGCF1 molecules. Locations of SpGCF1 target sites (A–E) are shown in Fig. 1. (*A–H*) A 2.3-kb region of DNA containing SpGCF1 target sites was incubated with purified SpGCF1 protein and prepared for EM by directly mounting onto thin carbon foils, air drying, and rotary shadowcasting with tungsten. The sites where the SpGCF1 molecules whose interactions form the loops are bound are indicated in each panel. Images are shown in reverse contrast. (Bar = 0.2 μm.)

some sites in region A have an apparent equilibrium constant for SpGCF1 interaction that is about 10 times higher than for the other measured *CyIIIa* sites, probably because of the opportunity for cooperative interactions between adjacent, bound SpGCF1 molecules within this cluster. Correspondingly, a major fraction of all the loops observed involved site cluster A (49 of the 59 structures analyzed; Table 1). Single loops formed by SpGCF1 interactions at two sites, as illustrated in Fig. 3 *A–E*, accounted for 71% of the structures we studied (42 of the 59 structures), and multiple loop structures accounted for the remainder. For example, Fig. 3 *F* and *G* show double loops formed by triple interactions and Fig. 3 *H* shows a five-loop structure formed by interactions at six possible sites.

## DISCUSSION

The correlation between the positions of SpGCF1 target sites in the *CyIIIa* regulatory domain and the positions of loop-forming sites, together with the high frequency of looped structures that we encountered in these experiments, confirms that SpGCF1 bound to DNA is capable of forming stable multimeric complexes. In the gel-shift experiment of Fig. 2*A*, this property evidently results in the sequestration of several DNA probe molecules per complex (bands A and B), and in the electron microscope study, it results in the formation of DNA loops. This conclusion stands independently of the exact stoichiometry of the complexes formed in the gel-shift experiment, which will require direct protein cross-linking to establish.

SpGCF1 evidently behaves very much like the mammalian Sp1 transcription factor. Sp1 is a transcriptional activator that binds G+C-rich target sites, which often appear in clusters (23, 24). Just like SpGCF1, Sp1 also forms higher multimeric complexes when bound to DNA probes, as visualized in gel-shift experiments (25). Furthermore, two different groups (21, 22) have shown by EM that *in vitro*, Sp1 loops DNA bearing multiple binding sites in the same way as does SpGCF1. However, SpGCF1 is not a sea urchin version of Sp1, since SpGCF1 has an entirely dissimilar sequence from Sp1. In fact, SpGCF1 displays no sequence similarity with any other known transcription factor (or other protein) (12). Thus, two completely unrelated transcription factors share the specific property of looping DNA by the formation *in vitro* of stable multimers between transcription factor molecules bound to DNA at distant sites. It follows that this property might reflect an important *in vivo* function.

The following is a speculation on the nature of this function that is suggested by the distribution of SpGCF1 sites in the *CyIIIa* regulatory domain. As indicated in Fig. 1, the A–B SpGCF1 site cluster lies at the distal end of the regulatory domain; the C-site cluster occurs toward the middle of the sequence where the SpGCF1 sites are contiguous to sites for two other transcription factors (labeled 7 and 8 in Fig. 1); the proximally located D and E sites bracket target sites for four additional transcription factors (labeled 1–4 in Fig. 1). Earlier work (5, 10, 11) and extensive gene transfer studies to be reported elsewhere show that these three regions of the *CyIIIa* regulatory domain have distinct functions. The distal region, which consists mainly of the SpGCF1 target sites of the A–B cluster, controls the level of expression mediated by the proximal and middle modules of the regulatory domain and has no spatial control function of its own. The proximal and middle modules each include target sites for both negatively and positively acting transcription factors, and they each exercise control over the spatial patterns of *CyIIIa* expression. The proximal module alone (target sites 1–5 plus SpGCF1 sites D and E in Fig. 1) suffices to set expression in the aboral ectoderm early in development. The middle module (target sites 6–8 plus SpGCF1 site C in Fig. 1) controls spatial and temporal patterns of expression in postgastrular embryos, but it requires portions of the proximal module for function. All three of these components of the regulatory domain are required for normal embryonic expression of *CyIIIa* transgenes, and they obviously must intercommunicate in some way; early in development distal–proximal intercommunication is required and later on distal–middle–proximal intercommunication is required. Formation of stable multimeric complexes between SpGCF1 molecules bound at the distal, middle, and proximal target sites could provide the specific mechanism for the physical intercommunication of transcription factors bound within individual regulatory modules with one another and with the basal transcription apparatus that will be located immediately adjacent to the proximal regulatory module. Thus, early in development we imagine a distal–proximal DNA loop (compare Fig. 3*D*) that would bring the activation function of distal SpGCF1 modules into direct proximity of the basal apparatus and the proximal transcription factors. Later in development multimerization of SpGCF1 bound at the A–B sites with SpGCF1 bound at the C and D and E sites would be predicted (compare Fig. 3 *G* and *H*). Similar arguments could apply to the *Endo16* gene regulatory system mentioned in the

Introduction, where SpGCF1 sites are again associated with cis-regulatory modules distributed over more than 2 kb of DNA. Thus, in general terms our speculation is that SpGCF1 carries out the essential function of a regulatory module intercommunicator, as well as serving as a transcriptional activator in its own right. The multimerization and DNA looping behavior of this protein that we observe *in vitro* would reflect this ubiquitously required *in vivo* function, and we propose that this function explains the widespread prevalence and local distribution of SpGCF1 sites in the cis-regulatory domains of sea urchin genes.

## APPENDIX

The accumulation of probe in the specific complexes shown in Fig. 2*A* as a function of protein concentration was tested against several models that might explain the formation and composition of these complexes. We compared the functions generated from these models to the data obtained from scanning the gels by PhosphorImager by means of a nonlinear least-squares procedure. The concentrations of recombinant SpGCF1 protein and of oligonucleotide probe were known *a priori*, as were the fractions of total protein ($\phi$) for each of the SpGCF1 molecular weight variants: where $i$ or $j$ denotes the 37-, 40-, 43-, 50-, and 55-kDa variants, and $P$ is protein, $P_i = \phi_i P$; $\Sigma\phi_i = 1$. We also knew (3, 12) that bands 1–5 contain only the 37-, 40-, 43-, 50-, and 55-kDa variants, respectively—i.e., if these complexes contain multimers of SpGCF1 protein they are homomultimers and not heteromultimers. Three models, and the degree to which they conformed to the data, expressed as $\chi^2$ values ($\sigma$ = 10%, $N = 13 \times 8$; $11 \times 8$ for the two data sets), are shown in the following Table. Here, where $n$ is the number of SpGCF1 (P) molecules, and $m$ is the number of DNA (D) probe molecules per complex, we have used the convention $n.m$ to describe the composition of the complexes:

| Model | Band | $n.m$; composition | $\chi^2$ (two data sets) |
|---|---|---|---|
| "Monomer" | 1–5 | 1.1; $P_iD$ | 18.1, 15.0 |
| | Region 5-A | 2.1; $P_j/P_iD$ | |
| | A | 4.2; $(P_j/P_iD)_2$ | |
| | B | 8.4; $((P_j/P_iD)_2)_2$ | |
| "Dimer" | 1–5 | 2.1; $(P_i)_2D$ | 5.3, 3.8 |
| | Region 5-A | 3.1; $P_j/(P_i)_2D$ | |
| | A | 6.2; $(P_j/(P_i)_2D)_2$ | |
| | B | 12.4; $((P_j/(P_i)_2D)_2)_2$ | |
| "Tetramer" | 1–5 | 4.1; $(P_i)_4D$ | 9.6, 10.2 |
| | Region 5-A | 5.1; $P_j/(P_i)_4D$ | |
| | A | 10.2; $(P_j/(P_i)_4D)_2$ | |
| | B | 12.4; $((P_j/(P_i)_4D)_2)_2$ | |

In constructing these models we assumed that all the mass variants of SpGCF1 may not interact with DNA with exactly the same equilibrium constant ($k_N$); $k_{N_i} = \rho_i k_N$ [$\rho_i$ values were known from earlier experiments (12)]. The equilibrium constant for SpGCF1/SpGCF1 interaction is $k_E$; for the interaction $(P_j/(P_i)_2D)/(P_j(P_i)_2D)$ it is $k_D$; for the interaction $(P_j/(P_i)_2D)_2/(P_j/(P_i)_2D)_2$ it is $k_Q$. Where $S_i = \rho_i\phi_i^2$, and $S = \Sigma_i S_i$, the equations used to describe the complexes formed in the dimer model are: bands 1–5, $PD_{2.1} = S_i k_N P^2 D$; Region 5-A, $PD_{3.1} = k_E S k_N P^3 D$; band A, $PD_{6.2} = k_D k_E^2 S^2 k_N^2 P^6 D^2$; and band B, $PD_{12.4} = k_Q k_D^2 k_E^4 S^4 k_N^4 P^{12} D^4$. It is worth noting also that in this model bands 1–5 approximately display an expected log–linear relation between mass and mobility in the gel, but this is no longer so for the higher order complexes.

We are pleased to acknowledge the useful insight of Prof. Robert Tjian, who in a telephone conversation early in this project alerted us to relevant studies on Sp1 carried out in his laboratory (21, 25). We are grateful to Profs. Carl Parker and Scott Fraser for their critical reviews of this manuscript and to Ms. Lora Cavallo for help with the EM preparations. This work was supported by National Institutes of Health Grants HD-05753 to E.H.D. and GM-31839 and GM-42342 to J.D.G. R.W.Z. was supported by an Office of Naval Research Augmentation Awards for Science and Engineering Research Training program (N00014-93-1-1400) and C.V.K. was supported by a training grant from the Austrian government.

1. Calzone, F. J., Thézé, N., Thiebaud, P., Hill, R. L., Britten, R. J. & Davidson, E. H. (1988) *Genes Dev.* **2,** 1074–1088.
2. Thézé, N., Calzone, F. J., Thiebaud, P., Hill, R. L., Britten, R. J. & Davidson, E. H. (1990) *Mol. Reprod. Dev.* **25,** 110–122.
3. Coffman, J. A., Moore, J. G., Calzone, F. J., Hood, L. E., Britten, R. J. & Davidson, E. H. (1992) *Mol. Mar. Biol. Biotechnol.* **1,** 136–146.
4. Char, B. R., Bell, J. R., Dovala, J., Coffman, J. A., Harrington, M. G., Becerra, J. C., Davidson, E. H., Calzone, F. J. & Maxson, R. (1993) *Dev. Biol.* **158,** 350–363.
5. Wang, D. G.-W., Kirchhamer, C. V., Britten, R. J. & Davidson, E. H. (1995) *Development (Cambridge, U.K.)*, in press.
6. Hough-Evans, B. R., Franks, R. R., Cameron, R. A., Britten, R. J. & Davidson, E. H. (1987) *Dev. Biol.* **121,** 576–579.
7. Cameron, R. A., Britten, R. J. & Davidson, E. H. (1989) *Mol. Reprod. Dev.* **1,** 149–155.
8. Livant, D., Cutting, A., Britten, R. J. & Davidson, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7607–7611.
9. Flytzanis, C. N., Britten, R. J. & Davidson, E. H. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 151–155.
10. Hough-Evans, B. R., Franks, R. R., Zeller, R. W., Britten, R. J. & Davidson, E. H. (1990) *Development (Cambridge, U.K.)* **110,** 41–50.
11. Franks, R. R., Anderson, R., Moore, J. G., Hough-Evans, B. R., Britten, R. J. & Davidson, E. H. (1990) *Development (Cambridge, U.K.)* **110,** 31–40.
12. Zeller, R. W. (1995) Ph.D. thesis (California Institute of Technology, Pasadena, CA).
13. Pei, D. & Shih, C. (1991) *Mol. Cell. Biol.* **11,** 1480–1487.
14. Mermod, N., O'Neill, E. A., Kelly, T. J. & Tjian, R. (1989) *Cell* **58,** 741–753.
15. Friedman, A. D. & McKnight, S. L. (1990) *Genes Dev.* **4,** 1416–1426.
16. Nocente-McGrath, C., Brenner, C. A. & Ernst, S. G. (1989) *Dev. Biol.* **136,** 264–272.
17. Ransick, A., Ernst, S., Britten, R. J. & Davidson, E. H. (1993) *Mech. Dev.* **42,** 117–124.
18. Akasaka, K., Frudakis, T. N., Killian, C. E., George, N. C., Yamasu, K., Khaner, O. & Wilt, F. H. (1994) *J. Biol. Chem.* **269,** 20592–20598.
19. Griffith, J. & Christiansen, G. (1978) *Annu. Rev. Biophys. Bioeng.* **7,** 19–35.
20. Griffith, J., Hochschild, A. & Ptashne, M. (1986) *Nature (London)* **322,** 750–752.
21. Su, W., Jackson, S., Tjian, R. & Echols, H. (1991) *Genes Dev.* **5,** 820–826.
22. Mastrangelo, I. A., Courey, A. J., Wall, J. S., Jackson, S. P. & Hough, P. V. C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5670–5674.
23. Dynan, W. S. & Tjian, R. (1983) *Cell* **35,** 79–87.
24. Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. (1986) *Science* **234,** 47–52.
25. Pascal, E. & Tjian, R. (1991) *Genes Dev.* **5,** 1646–1656.