MOLECULAR AND CELLULAR BIOLOGY, June 1985, p. 1465–1472
0270-7306/85/061465-08$02.00/0

Vol. 5, No. 6

Structure of the Promoter of the *Dictyostelium discoideum* Prespore EB4 Gene

ERIC BARKLIS, BRIAN PONTIUS,† KEN BARFIELD, HARVEY F. LODISH*

Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 17 December 1984/Accepted 4 February 1985

EB4 is one of several cloned cDNAs that is expressed as mRNA only after the aggregation stage of *Dictyostelium discoideum* differentiation and exclusively in prespore and spore cells (E. Barklis and H. F. Lodish, Cell 32:1139–1148, 1983). We have isolated the unique genome fragment corresponding to the 5′ portion of the EB4 message and the EB4 promoter. The EB4 transcript has an unusually long, G+C-rich, 5′ noncoding region, but initiates at several start sites within a region of DNA that is 96% A+T. The sequence GTGGTGG, along with slight variations, occurs several times in the promoter. We have used the EB4 promoter to drive the transcription of an EB4/β-galactosidase fusion transcript in yeast cells. Although the cap sites of the fused transcript in yeast cells are located in the region where multiple EB4 transcripts are initiated in *Dictyostelium*, the unregulated expression of the fusion transcript in yeast does not mimic the normal regulated pattern of EB4 mRNA expression in *D. discoideum*.

Dictyostelium discoideum is an excellent system for the study of differentiation of two or more cell types from a single type of cell. During the *Dictyostelium* developmental program, individual amoeboid cells are induced to aggregate into mounds of approximately 10^5 cells by pulsatile signalling of cyclic AMP (1). Aggregates then undergo a series of morphological changes which result in the formation of fruiting bodies, composed primarily of spore and stalk cells.

Previously, we (1, 6; E. Barklis, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, 1984) and others have identified and analyzed *Dictyostelium* cDNA clones complementary to mRNAs enriched in cells programmed to become either spores (prespore cells) or stalk cells (prestalk cells). Several studies suggest that there is a coordinate control of expression of genes specific for the two principal cell types. In particular, prespore RNAs begin to accumulate just after aggregation. Levels of prespore messages are drastically reduced by disaggregation of multicellular aggregates, but can be maintained in disaggregated cells by high levels of cAMP (1, 6, 22; Barklis, Ph.D. thesis). Prestalk messages can be subdivided into two groups on the basis of their temporal patterns of accumulation. Prestalk clones of class I are complementary to messages which are present at low levels in amoeboid and early developing cells, but accumulate to 10 to 20-fold-higher levels during later stages of development (1, 6). Prestalk mRNAs of class II accumulate only between 8 and 15 h of development and decrease thereafter (1, 6).

In an effort to further clarify mechanisms of prespore and prestalk gene regulation, we have undertaken the characterization of *Dictyostelium* genomic clones encoding cell type-specific messages. Here we report the characterization of the prespore EB4 promoter sequence, describe the function of the EB4 promoter in yeast cells, and compare it with other well-characterized *Dictyostelium* promoters. In the following paper (2), the structure of the prestalk D11 gene is described. These papers provide a foundation for further study of the regulation of cell type-specific messages in *D. discoideum*.

* Corresponding author.
† Present address: Department of Biochemistry, Stanford University Medical Center, Stanford, CA 94305.

MATERIALS AND METHODS

***Dictyostelium* strains, growth, differentiation, and DNA and RNA isolation.** *Dictyostelium* strain AX3 (17) was used for almost all experiments, although strains NC4 (24) and V12 (25) were used as described in the text. *Dictyostelium* growth and differention conditions, as well as the isolation of *Dictyostelium* polyadenylated [poly(A^+)]RNA, have been described previously (1). Isolation of *Dictyostelium* nuclear DNA was as described before (7).

***Dictyostelium* genomic library, enzymes, and cloning procedures.** A genomic library of *Dictyostelium* vegetative nuclear DNA was constructed and generously supplied by Rex Chisholm. The library was constructed by inserting partial *Sau*3A digests of *Dictyostelium* genomic DNA in place of the *Bam*HI "stuffer" fragment of the vector lambda Charon 28 (26). Restriction endonuclease, DNA ligase, DNA polymerase, Klenow fragment (15), DNase I, calf intestine alkaline phosphatase, and T4 polynucleotide kinase were from Bethesda Research Laboratories, New England Biolabs, or Boehringer Mannheim Biochemicals. Radiolabeled nucleotides were from Amersham Corp. The M13 cloning and sequencing reagents were from P-L Biochemicals. Unless otherwise noted, all cloning methods used were essentially as described by Maniatis et al. (20) or by Maxam and Gilbert (21).

Plaque screening and DNA and RNA blot hybridization. Generally, probes for hybridizations were made by nick translation (16). Occasionally single-stranded probes were synthesized from recombinant M13 templates, using large-fragment DNA polymerase I. In these cases, 10 ng of M13 pentadecamer primer (TCCAGTCACGACGT; New England Biolabs) was annealed to 1 μg of an M13 single-stranded (SS) DNA template in the presence of 2.5 μg of 10× *Pol*I

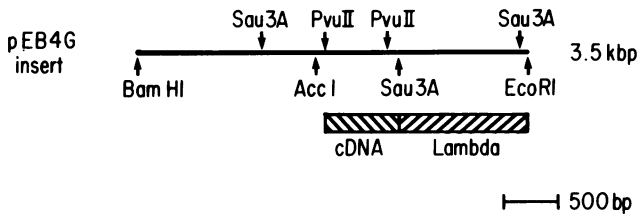


FIG. 1. Structure of the EB4 genomic fragment. Restriction map of the pEB4G 3.5-kbp insert. Striped bars indicate regions of lambda DNA and the region which hybridizes to the EB4 cDNA probe. Restriction sites are as shown.

buffer (10× buffer = 70 mM Tris [pH 7.5], 70 mM $MgCl_2$, 500 mM NaCl) in a final volume of 18.5 μl. For labeling, the primer-annealed template was then incubated for 1 h at room temperature in a final reaction volume of 25 μl containing 50 μCi of [α-^{32}P]dATP, 20 μM deoxynucleoside triphosphates, 2 mM dithiothreitol, 1× *Pol*I buffer, and 5 U of large-fragment polymerase I.

Lambda plaques were transferred to nitrocellulose and screened by the method of Benton and Davis (3). DNA fragments were separated by electrophoresis on agarose gels, transferred onto nitrocellulose, and hybridized to ^{32}P-labeled DNA probes by the method of Southern (28). RNA was fractionated by electrophoresis, transferred to nitrocellulose or diazobenzoxymethyl paper, and hybridized to probes as described previously (18). In all cases, nitrocellulose from Schleicher & Schuell was used.

DNA sequencing and S1 mapping. DNA sequencing was carried out by the chain terminator procedure of Sanger et al. (27), using M13mp8 and M13mp9 as sequencing vector (23). Two strategies were used for S1 mapping of RNA transcripts. For our initial mapping of transcripts (20), approximately 1μg (10 μl) of a recombinant M13 ssDNA was hybridized to 10 μg of poly(A$^+$) RNA or 80 to 100 μg of total RNA in 20 to 25 μl of 500 mM NaCl–20 mM Tris (pH 7.5)–2 mM EDTA. Reactions were covered with mineral oil, heated to 95°C for 4 to 5 min, and incubated at 65°C for 18 to 24 h. Reaction mixtures were removed from mineral oil and incubated at 37°C for 1 h in 200 μl of S1 buffer (30 mM sodium acetate [pH 4.6], 50 mM NaCl, 1 mM $ZnSO_4$, 5% glycerol) containing 300 to 400 U of S1 nuclease. Hybrids were ethanol precipitated after addition of 20 μg of yeast tRNA, resuspended, gel fractionated, and blotted onto nitrocellulose or diazobenzoxymethyl paper. RNA-DNA hybrids were visualized by hybridization to radiolabeled probes and autoradiography.

Precise S1 mapping of transcripts was performed following the procedure of Berk and Sharp (4), using slight modifications (8, 29). Essentially, uniformly labeled ssDNA (8) was hybridized to RNA and incubated with S1 nuclease as described above. Hybrids were precipitated with 5 μg of carrier yeast tRNA, resuspended in a minimal volume of dye mix (85% deionized formamide, 10 mM disodium EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol FF), denatured at 100°C for 2 to 3 min, run on a DNA sequencing gel (21), and visualized by autoradiography.

Yeast transformation and RNA preparation procedures. *Saccharomyces cerevisiae* strain BWG2-9a-1 (α *his519 ade ura3×42 gal40*) (5) was used for all transformations, using vectors detailed in Results. Transformations were performed by the method of Ito et al. (12), and *ura*$^+$ transformants were selected on yeast minimal medium plates containing 40 μg of adenine and histidine per ml (5). Total yeast RNA for RNA blot hybridizations and S1 mapping studies was isolated from log-phase cells as described previously (5).

RESULTS

Isolation of the EB4G genomic fragment. The regulation of the 2,080-base mRNA complementary to the EB4 cDNA clone is typical of the coordinately controlled class of prespore genes and has been described previously (1). Using the EB4 cDNA clone as a probe, we screened a lambda Charon 28 genomic library of *Dictyostelium* vegetative nuclear DNA by the method of Benton and Davis (3). From an initial screen of 40,000 recombinant lambda plaques, one, designated λEB4G, was positive on successive rescreenings.

By restriction analysis, we determined that a 3.5-kilobase pair (kbp) *Eco*RI-*Bam*HI fragment contained all *Dictyostelium* DNA sequences of the 46.1-kbp λEB4G DNA. Hybridization of λEB4G restriction fragments to EB4 cDNA indicated that sequences within this fragment were related to EB4. The 3.5-kbp *Eco*RI-*Bam*HI fragment was inserted into the large *Eco*RI-*Bam*HI fragment of pBR322 to facilitate further analysis (Fig. 1). The resultant pBR322 derivative was designated pEB4G. The 3.5-kilobase (kb) pEB4G insert contains approximately 1.1 kb of lambda Charon 28 DNA

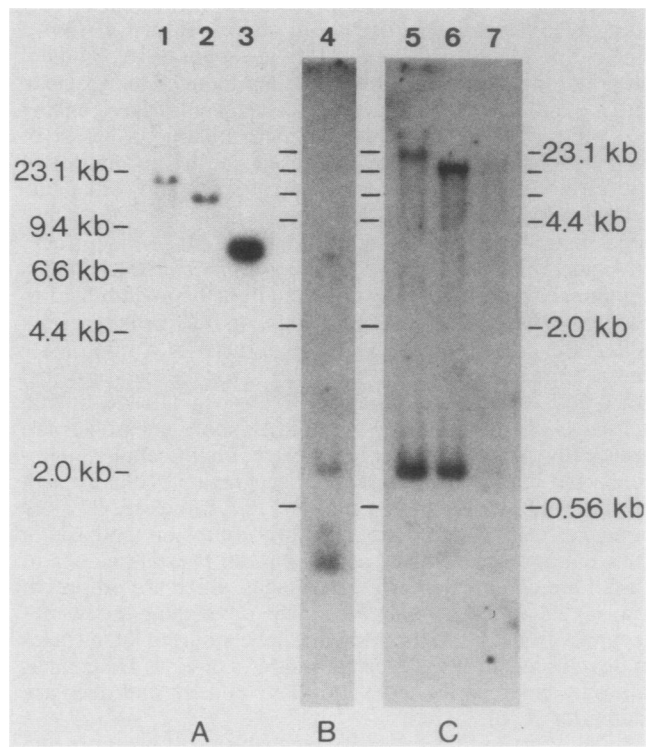


FIG. 2. Analysis of the EB4 region in genomic DNA. *Dictyostelium* total nuclear DNA (approximately 2 μg per lane) was digested with restriction endonuclases, fractionated by agarose gel electrophoresis, blotted onto nitrocellulose, and probed with ^{32}P-labeled DNA probes. (A) and (B) were probed with the EB4 cDNA clone. (C) was probed with pEB4G. Lanes 1 to 5 contain DNA from *D. discoideum* strain AX3; lane 6 contains DNA from strain V12; and lane 7 contains DNA from strain NC4. Lane 1 DNA was digested with *Hin*dIII; lane 2, *Bgl*II; lane 3, *Eco*RI; and lanes 4 to 7, *Pvu*II. Size markers for (A) are on the left; those for (B) and (C) are on the right.

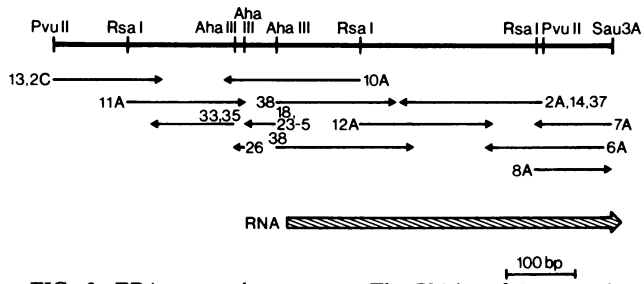


FIG. 3. EB4 sequencing strategy. The 794 bp of *Dictyostelium* DNA bordering the lambda sequences in pEB4G (see Fig. 1) was sequenced by the chain termination method. Numbers refer to recombinant M13 clones; i.e., 2a refers to clone mEB4G/2a. Arrows indicate the direction of the sequencing reaction and the amount of sequence obtained from a given clone. (Arrow sizes do not necessarily correlate with *Dictyostelium* insert sizes in M13 clones.) The direction of RNA transcription is as shown. Restriction endonuclease sites and the size scale are as indicated.

(from the *Eco*RI site to the *Sau*3A site) and 2.4 kb of *Dictyostelium* DNA (Fig. 1). By hybridization of the EB4 cDNA probe to DNA blots of restriction digests of the pEB4G insert, we showed that the region complementary to the EB4 clone is contained in the 700-base pair (bp) *Pvu*II fragment and the 100-bp *Pvu*II-*Sau*3A fragment (Fig. 1).

Hybridizations of λEB4G and pEB4G probes to blots of *Dictyostelium* RNA fractionated on formaldehyde gels (Northern blots) are identical to Northern blot hybridizations with the EB4 cDNA probe (data not shown). Each probe hybridizes strongly to a 2,030-base mRNA which is present during the postaggregation stages of *Dictyostelium* development but not during the earlier stages, evidence that pEB4G is the genomic fragment that encodes the EB4 mRNA.

To substantiate this point we performed DNA blot hybridizations (Southern blots) of restriction digests of *Dictyostelium* nuclear DNA to the EB4 cDNA and pEB4G probes. Figure 2A shows that the EB4 cDNA probe hybridizes to unique *Hin*dIII, *Bgl*II, and *Eco*RI genomic fragments, suggesting that the EB4 transcript is encoded by a unique *Dictyostelium* gene. Figure 2B indicates that the EB4 cDNA probe hybridizes to two genomic *Pvu*II fragments, and one of these corresponds to the pEB4G 700-bp *Pvu*II fragment. Figure 2C shows the hybridization of pEB4G to *Pvu*II digests of DNA from three different strains of *D. discoideum*, and in each lane a 700-bp *Pvu*II frgment hybridizes to the pEB4G probe. (Hybridization to DNA in lane 7 is faint because approximately one-fifth of the DNA in other lanes was used.) These results agree with the restriction mapping data of pEB4G and indicate that pEB4G contains a unique *Dictyostelium* genomic fragment which encodes a portion of the 2,030-base EB4 message.

The direction of transcription of the EB4 message was determined by two methods: hybridization of single-stranded probes to Northern blots of *Dictyostelium* 22-h poly(A$^+$) RNA, and S1 mapping of the 22-h transcripts with recombinant M13 ssDNA clones (see Materials and Methods). Each of these procedures indicated that the direction of EB4 message transcription is from *Dictyostelium* sequences toward lambda sequences in pEB4G (left to right in Fig. 1 and 3). Furthermore, the initial S1 mapping studies

```
CAGCTGCTCC AAGTGGTGGT GATAGTGGTG GTATTTGGAC AGGTGGTGAT   50
AAACATAAAA AAGATTGTGG TGGCACAGGA TGTTGTAAAG AAGGTCAATA  100
TTGTACCAAA TTGATGGAAA AGAACAATGT GTATATTATC CAGCATGACA  150
AAGGTCCACA ATGGTGGTTC ATTCTGTAGA GAAGGTGAAA TTTGTGTTTT  200
AGAACATGGT GTTTTGGGTT GTCTTGAAAA TCCAGTTGAA GATAGTAAAG  250
GTAAATATTT AAATATTTTT ATTTTAAATC TATAATAATT ATTTTTTATA  300
AAATTAATTT AATAACATTT TAAATTTTTT TTTTTTTTTA TTAATATTTC  350
TATTACTAGC AAATTGTGGT TTAGTCCATT GTAATTACAA TGAATATTGT  400
ATTATGGTTA ATGGTTATCT TCAATGTTTA TTCTCAAATG GTACTACACC  450
ATTTGCATGT GGTCTCCAAA TTGTGTTCCA CCACAATTAT GTGTAAAGGT  500
TGAAAATTGT CAACAATGTG TTTCTCCACC AAGACCAGAC AATAACTGTG  550
GTGACAAATA TTGTGATGAT GAACATCAAT GTGTTAGAAA AGGACTCGGT  600
TATGAATGTA TTCAAAACGT CTTACTTGTG AGACCAAAAA ATGCGAAGCC  650
TCTCAAGTTG TATCATGGTA AATGGTGATG CTCAATGTAT CGTACCACCA  700
GCTGCCGCCG CTGCTGAAAT TAGAATAAAC AATTTCAGAT TCAACCGTAA  750
CTCTGTGTTC AACGTAATGC CAAACCAGCT CAACAACAAA GATC        794
```

FIG. 4. DNA sequence of the EB4 promoter. The 794 bp from the internal *Pvu*II site of the pEB4G insert (nucleotide 1) to the *Sau*3A site at the junction of pEB4G *Dictyostelium* and lambda sequences (nucleotide 794) was sequenced by the chain terminator method, using the strategy described in the legend to Fig. 2. Only the RNA sense strand is shown. Putative TATA boxes are underlined; oligo(dT) stretches and the initiator methionine codon are boxed. The 16-bp palindrome (nucleotides 377 to 392) is underscored by a dashed line, as are occurrences (5′ to 3′ or 3′ to 5′ in either the sense or antisense strand) of the repeated TGGTGG motif. Black triangles indicate transcription initiation sites in *Dictyostelium* as determined in Fig. 4; open triangles indicate transcription initiation sites of the fused EB4/β-galactosidase message (see Fig. 5) in yeasts.

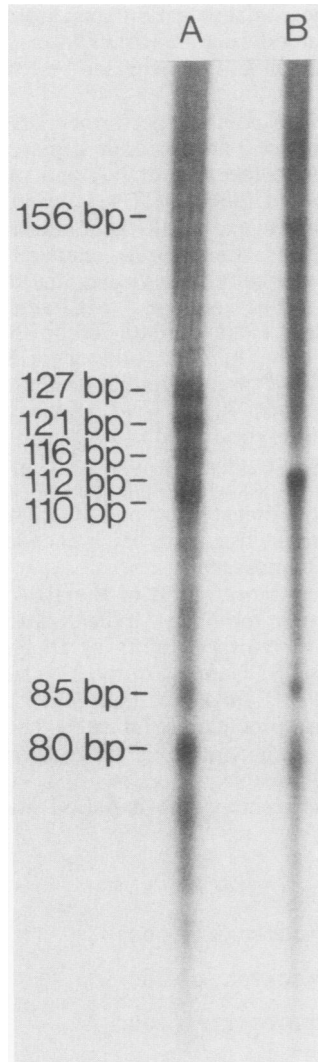


FIG. 5. S1 mapping of EB4 transcription initiation sites. A uniformly ^{32}P-labeled ssDNA probe, extending from the pEB4G *Rsa*I site at position 442 (Fig. 3) to the *Rsa*I site at position 104, was generated from the mEB4G/10c template by DNA synthesis, using large-fragment polymerase I. This probe was incubated with either 10 μg of cytoplasmic poly(A$^+$) RNA from 22-h developing *Dictyostelium* cells or 80 μg of total RNA from a yeast ΔUAS/EB4-1 transformant (see Fig. 6). Incubations were for 20 h at 65°C in 25 μl of 500 mM NaCl–20 mM Tris (pH 7.5)–2 mM EDTA and were followed by digestion with S1 nuclease (400 U per incubation) at 37°C for 1 h. S1-protected material was ethanol precipitated after addition of 5 μg of carrier yeast tRNA, resuspended, denatured, and applied to a 6% polycarylamide DNA sequencing gel. (A) *Dictyostelium* cytoplasmic poly(A$^+$) 22-h RNA. (B) Total RNA from the yeast ΔUAS/EB4-1 transformant. Incubation of probe with yeast tRNA produced no protected bands (not shown). Sizes were determined from DNA sequencing reactions run in parallel on the gel.

indicated the EB4 message is complementary to approximately 350 bases of the pEB4G 700-bp *Pvu*II piece. These results suggested that the 5′ end of the EB4 message is midway between the two *Pvu*II sites in the pEB4G insert.

Structure of the EB4 promoter. The 794 bp of pEBa4G DNA which hybridizes to the EB4 cDNA probe, including the 701-bp *Pvu*II fragment and the 93-bp *Pvu*II-*Sau*3A fragment, was sequenced by the dideoxy chain terminator procedure of Sanger et al. (27). The sequencing strategy is shown in Fig. 3; the sequence is shown in Fig. 4. Only the strand corresponding to the RNA strand, from the internal *Pvu*II site (position 0) to the *Sau*3A site (794) is shown.

To precisely assign the site of EB4 transcription initiation, transcripts were mapped by using a modification (4) of the Berk-Sharp S1 mapping procedure (see Materials and Methods). A uniformly ^{32}P-labeled ssDNA probe, extending from the pEB4G *Rsa*I site (GT/AC) at position 442 (Fig. 4) to the *Rsa*I site at position 104, was generated from the mEB4G/10c template as described by Dierks et al. (8). This labeled fragment was hybridized to *Dictyostelium* 22-h poly(A$^+$) RNA, and unprotected regions were subsequently digested with S1 nuclease. The remaining hybrids were denatured and fractionated on a 6% polyacrylamide sequencing gel (21). Protected DNA bands were visualized by autoradiography. *Dictyostelium* RNA protects a series of bands with the sizes of 110, 111, 112, 116, 121, 122, 128, and 129 bases (Fig. 5). Minor protected bands also appear in the region of 79 to 80 and 84 to 85 bases. These fragments indicate a heterogeneous set of EB4 transcription initiation sites at positions 314, 315, 321, 322, 327, 331, 332, 333, 358 to 359, and 363 to 364 (Fig. 4).

The overall percentage of A+T nucleotides in the EB4 promoter fragment is 67%. However, in the 100 bases from positions 251 to 350, the percent A+T is 96%, characteristic of regions adjacent to transcription initiation sites in *Dictyostelium* (14). Within this A+T-rich region there are several candidate TATA boxes, similar to the *Dictyostelium* consensus TATA sequence TATAAA(T/A)A (13, 14). Also, there are three oligodeoxythymidylate [oligo(dT)] stretches in this region (positions 266 to 275, 289 to 297, 325 to 339) which occur between TATA boxes and cap sites of *Dictyostelium* actin, discoidin, M3L, M3R, M4, and D2 genes (14; see also Fig. 7). The transcription initiations at positions 314 to 315, 321 to 322 and 327 are 28 to 40 bases downstream from the potential TATA box, TATAATA, at positions 381 to 287; the starts at positions 331 to 333 are 27 to 19 bases downstream from the potential TATA box, TATAAAAT, at positions 297 to 304; and the starts at positions 358 and 363 do not follow any obvious TATA sequence.

Because EB4 transcription initiates between positions 314 and 364, only the segment representing the 5′ 420 to 480 bases of the 2,030-base EB4 message are represented on the pEB4G genomic fragment. Consequently, if a protein is initiated within the pEB4G fragment, an open translational reading frame would be expected to read through position 794 of our sequence. The only such reading frame begins at a methionine codon at position 641. This ATG is immediately preceded by several A nucleotides, characteristic of translation starts in *Dictyostelium* (13, 14). A termination codon, UGA, appears in frame, four codons upstream from the methionine codon. The only other long open reading frame within the sequenced region begins at position 424 and the ATG initiator is preceded by ony a single A nucleotide. This open reading frame codes for 83 amino acids and is terminated by a UGA codon at position 676. If EB4 translation begins at position 641, the EB4 message possesses a remarkably long 5′ untranslated region of approximately 300 nucleotides. The G+C content of this untranslated region is 32%, significantly higher than the 10% G+C content of most *Dictyostelium* 5′ untranslated regions (14).

Function of the EB4 promoter in yeast cells. To assess the functional ability of the EB4 promoter, the 1.4-kb *Sau*3A fragment from pEB4G was fused to the β-galactosidase gene of the yeast transformation vector ΔUAS (9, 10; see Fig. 6).

The ΔUAS vector is a derivative of pLG669-Z (10, 11), generously provided by L. Guarente, and has been described previously (9, 10). The ΔUAS vector has the pBR322 ori and β-lactamase gene, and the yeast 2 μm circle and *ura3* gene (Fig. 6). These sequences permit replication and selection in both *Escherichia coli* and yeasts. Also, ΔUAS contains the structural gene for β-galactosidase fused to the yeast *cyc1* promoter (10, 11). The yeast *cyc1* promoter has been inactivated by deletion of the upstream activator sequence (UAS) and produces no detectable β-galactosidase RNA or protein (9, 10). In the ΔUAS/EB4-1 construct, the pEB4G 1.4-kb *Sau*3A fragment was inserted into the *Bam*HI site at the beginning of the β-galactosidase structural gene to form a fused gene (Fig. 6). In this construct, the β-galactosidase message is expected to be translated in the wrong reading frame, producing no active β-galactosidase enzyme.

The ΔUAS and ΔUAS/EB4-1 plasmids were used to transform *S. cerevisiae* strain BWG2-9A-1 (α *his4-519 ade ura3-52 gal40*) (5) by the methods of Ito et al. (12), using the selectable *ura3* marker. Total RNA was isolated from the host strain, three ΔUAS transformants and eight ΔUAS/EB4-1 transformants as described before (5; see also Materials and Methods). RNA from the host and each transformant was fractionated by formaldehyde gel electrophoresis, transferred onto nitrocellulose, and hybridized to labeled DNA specific for either EB4 (pEB4G) or β-galactosidase (M13mp8).

The pEB4G probe detects no transcripts in the yeast host strain BWG-9a-1 (Fig. 7A, lane a). In ΔUAS transformants (lanes b to d), two low-molecular-weight messages, of approximately 1,800 and 2,800 bases, are detected by the pEB4G probe. These messages probably derive from the pBR322 sequences in pEB4G (9, 10). In all ΔUAS/EB4-1 transformants (lanes e to l), a large RNA, greater than 4,000 bases, is detected. The size of this RNA corresponds to the expected size of an EB4/β-galactosidase fusion transcript. Figure 7B shows the hybridization patterns of these 12 RNAs to M13mp8, which was used as a probe for transcripts complementary to β-galactosidase DNA. No RNA from the BWG-9a-1 host or ΔUAS transformants is complementary to this probe (lanes a to d). In contrast, RNA species complementary to the M13mp8 probe is present in all of the ΔUAS/EB4-1 transformants (lanes e to l); the high-molecular-weight band in each of these lanes corresponds exactly to

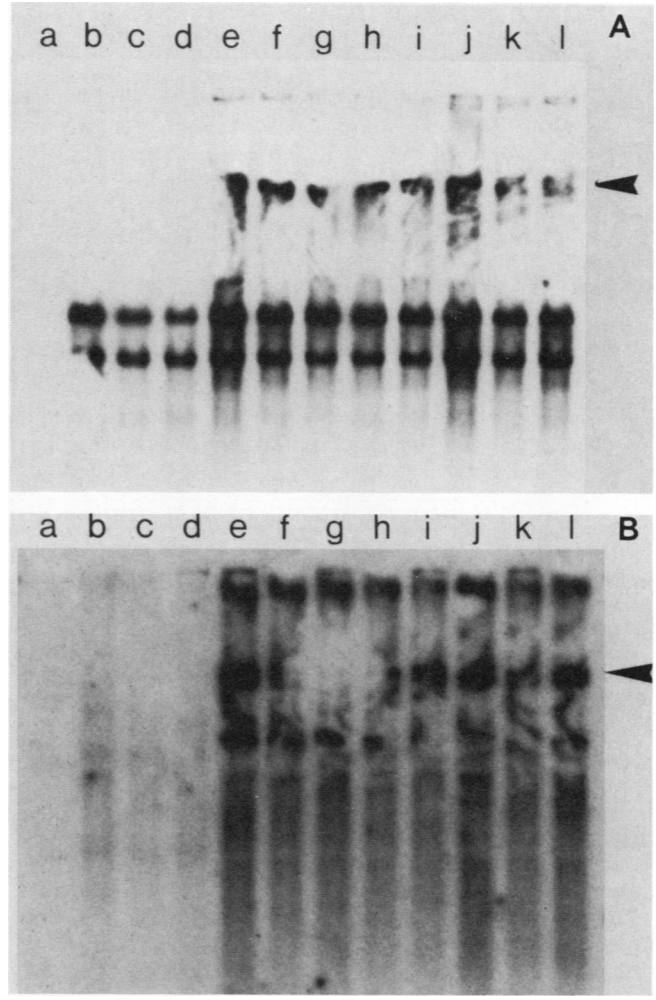
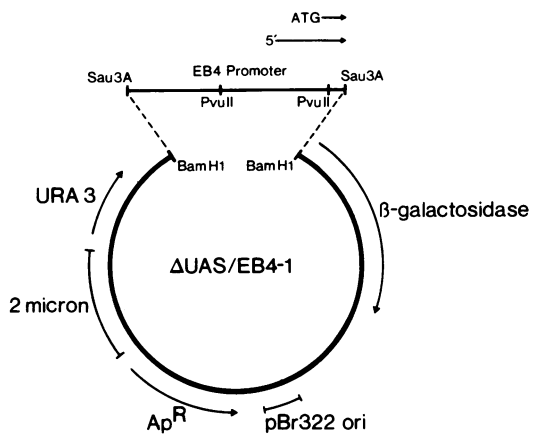


FIG. 7. Expression of the EB4/β-galactosidase-fused mRNA in ΔUAS/EB4-1 yeast transformants. Total RNA (15 μg per lane) from the untransformed yeast host strain BWG2-9a-1 (a), from yeast ΔUAS transformants (b to d), and from yeast ΔUAS/EB4-1 transformants (e to l) was fractionated on a formaldehyde gel, blotted onto nitrocellulose, and hybridized to ^{32}P-labeled pEB4G DNA (A) or to ^{32}P-labeled m13mp8 RF DNA (B), which was used as a probe to detect RNA sequences complementary to the β-galactosidase gene. The arrowhead indicates the RNA band which corresponds to the expected size of the fused EB4/β-galactosidase message and is present in only ΔUAS/EB4-1 transformants. The RNA fractionation, blotting, and hybridization of (A) and (B) were performed at the same time from the same gel.



β-galactosidase ATG ACC GGA TCC GGA . . . . . ▸

EB4/β-gal ATG . . . AAA GAT CCG GAG . . . . . ▸

FIG. 6. Structure of the ΔUAS/EB4-1 plasmid. The ΔUAS/EB4-1 plasmid was constructed by insertion of the 1.5-kbp *Sau*3A fragment from the pEB4G insert into the unique *Bam*HI site of the ΔUAS yeast transformation vector. The vector contains sequences of the yeast 2 μm circle, permitting replication in yeast cells; the pBR322 origin of replication, permitting replication in *E. coli*; the pBR322 β-lactamase gene; and the yeast *ura3* gene. The EB4 promoter was inserted into the vector so as to yield transcription of a fused EB4/β-galactosidase message under the direction of the EB4 promoter. The approximate initiation site of transcription (from Fig. 5) and translation ATG are shown. Assuming that translation occurs from the ATG at positions 641 to 643 (see Fig. 4), translation of the β-galactosidase-coding region will be out of phase, as illustrated at the bottom of the diagram.

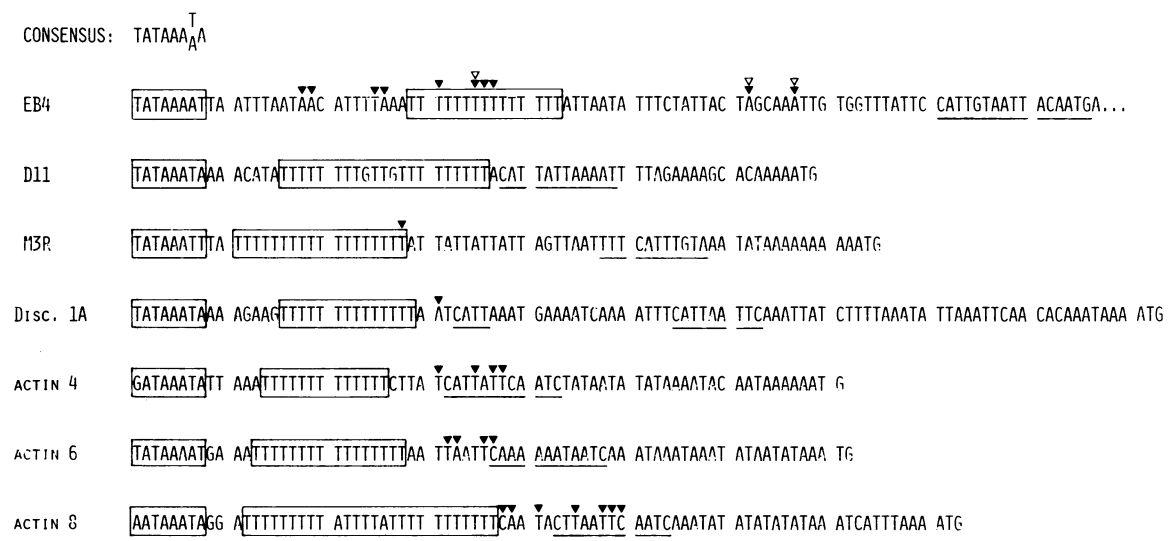


FIG. 8. Comparison of *Dictyostelium* promoters. The DNA sequence of the sense strand of seven *Dictyostelium* promoters is shown. Sequences are aligned by TATA sequences. TATA boxes and oligo(dT) stretches are boxed. CATT sequences, which appear at the 5′ untranslated region of many *Dictyostelium* mRNAs, are underlined. Black arrows indicate precise transcription initiation sites, when determined. Open arrows indicate initiation sites of the fused EB4/β-galactosidase RNA in yeast ΔUAS/EB4-1 transformants. The sequence of the EB4 promoter region is depicted from nucleotides 296 to 393 (Fig. 3). The putative EB4 initiator codon is locatd 250 bases downstream from the 3′ end shown here. The promoter regions of the bottom six promoters are depicted from the TATA box to the initiator methionine codon. The *Dictyostelium* consensus TATA sequence and the actin, discoidin 1A, and M3R promoter sequences are from Kimmel and Firtel (13). The sequence of the D11 promoter region is from Barklis et al. (2).

the size of the high-molecular-weight band detected by the pEB4G probe. From this we conclude that an EB4/β-galactosidase fusion transcript is produced in all ΔUAS/EB4-1 transformants.

To determine the transcription initiation site of the fusion message in yeast cells, RNA from one of the yeast transformants (lane j, Fig. 7) was mapped by the Berk-Sharp S1 mapping procedure. Yeast total RNA was mapped as described above for mapping of *Dictyostelium* poly(A^+) RNA; results are shown in Fig. 5B. As shown, fusion transcripts in yeasts appear to initiate very close to initiation sites of the EB4 message in *Dictyostelium*. In particular, a major protected fragment appears at 112 bases, whereas minor bands appear at 79 to 80, 84 to 85, and 156 bases. Although the 156-base fragment does not correspond to any protected fragments seen in *Dictyostelium*, this initiation site (Fig. 4, position 287) is 12 bases downstream from a possible TATA box (TAAATAT; positions 252 to 258). Initiation sites of the EB4/β-galactosidase fusion message are shown in Fig. 4.

DISCUSSION

We have cloned and sequenced *Dictyostelium* genomic segment corresponding to the EB4 prespore mRNA. Hybridization of restriction digests of *Dictyostelium* nuclear DNA to the EB4 cDNA clone and to the pEB4G genomic clone indicate that the EB4 message is encoded by a single-copy gene. The pEB4G genomic clone contains the 5′ region of the 2,030-base EB4 message and includes at least a portion of the EB4 promoter. This is the first prespore-specific *Dictyostelium* promoter region to be sequenced. Because message abundance levels of developmentally regulated mRNAs are controlled, at least partially, at the level of transcription, the EB4 promoter should contain sequences that respond to stage-specific transcription regulators.

Sequence analysis (Fig. 4) and S1 mapping of the EB4 transcript (Fig. 5) indicate that the EB4 message initiates within an A+T-rich region similar to the 5′ flanking regions of *Dictyostelium* actin, discoidin, M3L, M3R, M4, and D11 genes which encode messages which are synthesized in growing cells or during early development (14; see Fig. 8). The initiation sites of the EB4 mRNA are somewhat heterogeneous compared with the cap sites of other *Dictyostelium* genes (Fig. 8). However, with the exception of the cap sites mapped at positions 358 and 363, the initiation sites we have mapped are characteristic of *Dictyostelium* initiation sites; they follow oligo(dT) stretches and are located 25 to 40 bases downstream of TATA-like sequences. Nevertheless, we cannot exclude the possibility that some of the heterogeneity in the mapped EB4 cap sites is due to S1 overdigestion of transiently melted A+T-rich regions of RNA/DNA hybrids.

An unusual feature of the EB4 trancript is the long G+C-rich 5′ untranslated region. Assuming that translation begins at the methionine codon at position 641, this 5′ noncoding region is 300 bases long and is 32% G+C. In contrast, other *Dictyostelium*-transcribed but noncoding regions tend to be less than 100 nucleotides and have a G+C content of 10% or less (13). The EB4 message differs from other sequenced *Dictyostelium* messages in that it is preferentially degraded in disaggregated *Dictyostelium* cells (1, 19). It is possible that the EB4 5′ noncoding region mediates this sensitivity.

Within the EB4 noncoding region is a CATT sequence, which appears slightly downstream of the cap site in the 5′ untranslated region of almost all characterized *Dictyostelium* messages (Fig. 8). The developmentally regulated *Dictyostelium* M4 band 4-3 mRNA contains the CATT sequence within the sequence TGATTCATTTG, which is complementary to a sequence in the *Dictyostelium* 17S rRNA (13). In the coregulated M3R and M3L mRNAs, this CATT sequence is nested within a conserved dodecamer nucleotide stretch, TTCATTTGTA. In the EB4 transcript, the CATT

sequence appears at the beginning of a 16-base perfect palindrome, CATTGTAATTACAATG.

Another feature of the EB4 promoter is the occurrence of a repeated GTGGTGG motif in the region surrounding the transcription initiation sites. The sequence GTGGTGGT is present at positions 13 to 20 and 25 to 32, well upstream of where RNA synthesis begins at positions 312 to 362 (see Fig. 4). The sequence GTGGTGGCACAGG is present at positions 67 to 70 and occurs in a 3'-to-5' orientation with a 1-base deletion at positions 37 to 48. Minor variations of this motif occur at positions 212 to 219 and in the opposite strand at positions 478 to 485, 524 to 531, and 694 to 700. Shorter variations, such as TGGTG and TGTGT, occur in either strand an additional 10 to 12 times. The EB4 GTGGTGG motif is similar to the potential enhancer core sequence, GTGGTTTG, which is present, in a number of variations, in several well-characterized viral and mammalian enhancers (15, 19). Kimmel and Firtel (14) have observed that the $(TTG)_n$ repeated sequence occurs in the anti-sense strand of the 5' flanking region of the M4 band 4-3 gene, is repeated many times throughout the *Dictyostelium* genome, and is complementary to the $(AAC)_n$ repeat sequence found in the 5' noncoding regions of one class of aggregation-stage, *Dictyostelium*-dependent messages. A series of cDNA clones bearing the $(GGT)_n$ repeat have also been isolated (14). At the present time, the function of these sequences is not understood.

Experiments with the ΔUAS/EB4-1 yeast transformation vector indicate that the 1.4-kb pEB4G *Sau*3A fragment is capable of initiating transcription of the fused EB4/β-galactosidase RNA in yeast cells (Fig. 7). S1 mapping experiments locate the cap sites of the fused transcript in yeasts to the region where EB4 transcripts are initiated in *Dictyostelium*. Expression of the EB4/β-galactosidase RNA in yeasts is not merely a consequence of high plasmid copy number. Since the levels of the 1.8- and 2.8-kb RNAs present in ΔUAS and ΔUAS/EB4-1 transformants are the same (Fig. 7A), we estimate that copy numbers of these plasmids in host cells are similar. In contrast, transcripts detected by the M13mp8 probe (Fig. 7B) are made in ΔUAS/EB4-1 transformants but not transformants by the ΔUAS clone. The specific elements of the 1.4-kb *Sau*3A fragment responsible for message transcription in yeasts are under study.

The functional activity of the EB4 promoter in yeast cells is somewhat surprising, since the EB4 mRNA is present and transcribed in differentiating but not vegetatively growing *Dictyostelium* cells. Attempts to increase expression of the EB4/β-galactosidase RNA in ΔUAS/EB4-1 transformants by addition of exogenous cyclic AMP (1 mM), or by sporulation after formation of diploids with the yeast strain BJW-T13A (*his*$^+$ *ade*$^+$ *ura*$^-$), have been unsuccessful. Thus, regulation of the EB4/β-galactosidase RNA in yeast cells does not mimic the pattern of EB4 RNA regulation in *Dictyostelium*.

We hope that it will be possible to apply our knowledge of the EB4 promoter towards the study of *Dictyostelium* prespore gene regulation by transformation of *Dictyostelium* cells with fusion genes similar to the EB4/β-galactosidase gene.

LITERATURE CITED

1. **Barklis, E., and H. F. Lodish.** 1983. Regulation of *Dictyostelium discoideum* mRNAs specific for prespore or prestalk cells. Cell **32**:1139–1148.
2. **Barklis, E., B. Pontius, and H. F. Lodish.** 1985. Structure of the *Dictyostelium discoideum* prestalk D11 gene and protein. Mol. Cell. Biol. **5**:1473–1479.
3. **Benton, N., and R. W. Davis.** 1977. sequence λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**: 180.
4. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–732.
5. **Cappello, J., C. Zuker, and H. F. Lodish.** 1984. Repetitive *Dictyostelium* heat-shock promoter functions in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:591–598.
6. **Chisholm, R., E. Barklis, and H. F. Lodish.** 1984. Mechanism of sequential induction of cell-type specific mRNAs in *Dictyostelium* differentiation. Nature (London) **310**:67–69.
7. **Chung, S., C. Zuker, and H. F. Lodish.** 1983. A repetitive and apparently transposable DNA sequence in *Dictyostelium discoideum* associated with developmentally regulated mRNAs. Nucleic Acids Res. **11**:4835–4852.
8. **Dierks, P., A. van Ooyen, M. Cochron, D. Dubkin, J. Reiser, and C. Weissman.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell **32**:695–705.
9. **Guarente, L., B. Lelonde, P. Gifford, and E. Alani.** 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. Cell **36**:503–511.
10. **Guarente, L., and T. Mason.** 1983. Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. Cell **32**:1279–1286.
11. **Guarente, L., and M. Ptashne.** 1981. Fusion of *Escherichia coli lacZ* to the cytochrome c gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A. **78**:2199–2203.
12. **Ito, A., Y. Kukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.
13. **Kimmel, A., and R. Firtel.** 1982. The organization and expression of the *Dic*tyostelium genome, p. 233–324. *In* W. F. Loomis (ed.), The development of *Dictyostelium doscoideum*. Academic Press, Inc., New York.
14. **Kimmel, A., and A. Firtel.** 1983. Sequence organization in *Dictyostelium*: unique structure at the 5' ends of protein coding genes. Nucleic Acids Res. **11**:541–552.
15. **Klenow, H., and I. Henningsen.** 1970. Selective elimination of the exonuclase activity of the deoxyribonucleic acid polymerase from *Escherichia coli B* by limited proteolysis. Proc. Natl. Acad. Sci. U.S.A. **65**:168.
16. **Landfear, S., P. Lefebvre, S. Chung, and H. F. Lodish.** 1982. Transcriptional control of gene expression during development of *Dictyostelium discoideum*. Mol. Cell. Biol. **2**:1417–1426.
17. **Loomis, W.** 1970. Temporal control of differentiation in the slime mold, *Dictyostelium discoideum*. Exp. Cell. Res. **60**:285–289.
18. **Mangiarotti, G., S. Chung, C. Zuker, and H. F. Lodish.** 1981. Selection and analysis of cloned developmentally regulated *Dictyostelium discoideum* genes by hybridization competition. Nucleic Acids Res. **9**:947–963.
19. **Mangiarotti, G., P. Lefebvre, and H. F. Lodish.** 1981. Differences in the stability of developmentally regulated mRNAs in aggregated and disaggregated *Dictyostelium discoideum* cells. Dev. Biol. **89**:82–91.
20. **Maniatis, T., E. Fritsch, and J. Sambrook.** 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Maxam, A., and J. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–500.
22. **Mehdy, M., D. Ratner, and R. Firtel.** 1983. Induction and modulation of cell-type-specific gene expression in *Dictyostelium*. Cell **32**:763–771.
23. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
24. **Raper, K. B.** 1935. *Dictyostelium discoideum*, a new species of slime mold from decaying forest leaves. J. Agric. Res. **50**:135–147.

25. **Raper, K. B.** 1951. Isolation, cultivation and conservation of simple slime molds. Q. Rev. Biol. **26**:169–190.
26. **Rimm, R., D. Horness, J. Huchera, and F. Blattner.** 1980. Construction of coliphage lambda charon vectors with BamHI cloning sites. Gene **12**:301.
27. **Sanger, F., S. Nichen, and A. Colsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
28. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503.
29. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β globin mRNA precursor and mature 10S β globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.