MOLECULAR AND CELLULAR BIOLOGY, Nov. 1984, p. 2332–2340
0270-7306/84/112332-09$02.00/0

Vol. 4, No. 11

Transcription of *Dictyostelium discoideum* Transposable Element DIRS-1

STEPHEN M. COHEN, JOE CAPPELLO, AND HARVEY F. LODISH*
Department of Biology, Massachusetts Institute of Technology

Received 25 June 1984/Accepted 17 August 1984

DIRS-1 is a *Dictyostelium discoideum* transposable element that contains heat shock promoter sequences in the inverted terminal repeats. We showed that transcription of a 4.5-kilobase polyadenylated RNA initiates at a discrete site within the left-terminal repeat of DIRS-1, downstream from heat shock promoter and TATA box sequences. This RNA represents a full-length transcript of DIRS-1. We describe a cDNA clone that contains the 4.1 kilobases of internal sequence of DIRS-1, a cDNA clone that spans the junction between the internal sequences and the right-terminal repeat, and a cDNA clone that appears to have been transcribed from a rearranged genomic copy of DIRS-1. A second DIRS-1 RNA, named E1, is transcribed on the opposite strand of DIRS-1 from the 4.5-kilobase RNA and is under control of the heat shock promoter in the right-terminal repeat. E1 transcription initiates at multiple positions both within and downstream from the right-terminal repeat. The same transcriptional initiation sites are used during normal development and during heat shock, suggesting that in all cases DIRS-1 transcription is regulated by the heat shock promoters contained within the two terminal repeats.

Dictyostelium intermediate repeat sequence 1 (DIRS-1) is a moderately repetitive and apparently transposable element with several unusual features (4, 7, 15). DIRS-1 consists of 4.1 kilobases (kb) of unique internal sequence flanked by inverted terminal repeats of unequal length, 332 and 360 base pairs (bp) (6, 23). There are about 40 copies of the intact DIRS-1 element interspersed throughout the *Dictyostelium* genome, as well as an additional 200 copies of related sequences (7). Different *Dictyostelium* strains have different sequences flanking some of their DIRS-1 elements, indicating that the element had transposed after these strains were separated (7, 15).

Transcription of DIRS-1-related sequences results in the production of a heterogeneous population of polyadenylated cytoplasmic RNAs that are differentially expressed during development (7, 24; C. Zuker, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge). DIRS-1 RNAs are present at very low levels in growing cells, begin to accumulate within the first hour after initiation of development, and reach their maximal level by 15 h. Transcription of DIRS-1-related RNAs is induced in vegetative cells subjected to stresses such as heat shock and high cell density (22).

Examination of the DNA sequence of the DIRS-1 terminal repeats has led to the identification of a putative heat shock promoter sequence (22, 23) homologous to the *Drosophila* consensus heat shock promoter (13, 14). Stress-induced transcription of DIRS-1-related RNAs is presumed to be directed by these sequences. Cappello et al. (5) showed that a cloned partial copy of DIRS-1 (pB41.6) directs transcription of a heat-shock-inducible RNA in yeasts and that heat-inducible transcription of this RNA is dependent on the presence of the DIRS-1 terminal repeat. The isolated terminal repeat also directs heat-inducible transcription of the flanking vector sequences (5).

In *Dictyostelium* strains, the direction of transcription of the majority of the DIRS-1 RNAs is from the left-terminal (332-bp) repeat, rightward into the element (see Fig. 2A for a representation of the structure of DIRS-1) (15; Zuker, Ph.D. thesis). Among the heterogeneous population of RNAs transcribed from the left repeat is a major species of ca. 4.5 kb. We show here that this RNA represents a nearly full-length transcript of an intact DIRS-1 element. Transcription of the 4.5-kb RNA initiates at a discrete site within the left-terminal repeat. Transcription of another RNA (E1) occurs from the opposite strand of DIRS-1, initiating near the right-terminal repeat (15). In both cases, the same transcriptional initiation sites are used by cells during normal development and during heat shock. These results suggest that transcription of DIRS-1 RNAs is regulated by the heat shock promoter sequences contained in the inverted terminal repeats.

MATERIALS AND METHODS

***Dictyostelium* culture conditions and genomic clones.** *Dictyostelium discoideum* AX-3 was grown axenically and plated for development on filters as described previously (1). Growing cells were harvested when the concentration reached 2×10^6 cells per ml.

The DIRS-1 genomic clones used here (pB41.6 and pB41.3) are subclones of the λ genomic clone SB41 and have been described in detail elsewhere (7, 22, 24). pB41.3 contains the entire 4.1 kb of DIRS-1 unique internal sequence flanked by *Eco*RI sites. pB41.6 contains a 2.5-kb *Eco*RI fragment that includes the right-terminal repeat of DIRS-1 and 2.2 kb of flanking sequence, which in SB41 is another copy of the right half of DIRS-1 (4, 24).

RNA preparation, transfer, and hybridization. RNA was isolated from suspension cultures and filter-developed cells by the guanidine thiocyanate extraction procedure of Chirgwin et al. (6). Cells were harvested and washed as described previously (3), dissolved in 50% guanidine thiocyanate–0.5% *N*-lauryl sarcosine–0.1% 2-mercaptoethanol–25 mM EDTA, and centrifuged through a cesium chloride cushion. The RNA pellet was dissolved in water and stored at −70°C. Polyadenylated [poly(A)$^+$] RNA was purified by oligodeoxythymidylate [oligo(dT)]-cellulose chromatography as described previously (3). RNA samples used for gel transfer

* Corresponding author.

hybridization analysis were adjusted to 50% formamide–18 mM Na_2HPO_4–2 mM NaH_2PO_4–6% formaldehyde, heat denatured for 10 min at 65°C, and electrophoresed on 1% formaldehyde agarose gels (Zuker, Ph.D. thesis). RNA gels were transferred to Gene Screen, and the filters were hybridized to nick-translated DNA probes as described previously (22, 24).

cDNA library construction and screening. Double-stranded cDNA was synthesized essentially as described by Efstradiatis et al. (8) from poly(A)$^+$ RNA prepared from 8-h filter-developed *D. discoideum* cells. This RNA was a gift of R. Chisholm. Before cDNA synthesis, 10 μg of RNA was denatured in 10 mM methylmercury hydroxide for 10 min at room temperature followed by the addition of 2-mercaptoethanol to a concentration of 100 mM. cDNA synthesis was carried out for 2 h at 42°C in 100 mM Tris hydrochloride (pH 8.3)–10 mM $MgCl_2$–140 mM KCl–1 mM dATP, dTTP, dCTP and dGTP, containing 25 μg of oligo (dT), 400 U of RNAsin (Promega Biotec), 200 μCi of [^{3}H]dGTP, and 150 U of AMV reverse transcriptase (gift of the laboratory of D. Baltimore) in a final volume of 250 μl. The reaction mixture was extracted with phenol, chloroform, and ether and heat denatured at 95°C for 3 min. Second-strand synthesis was carried out by addition of 67.5 U of *Escherichia coli* DNA polymerase I large fragment, 50 μCi of [^{32}P]dATP, and 1/10 volume of a solution containing 0.5 M Tris hydrochloride (pH 7.2), 0.1 M $MgSO_4$, and 1 mM dithiothreitol (DTT). The reaction was incubated for 16 h at 15°C followed by extraction with phenol, chloroform, and ether and precipitation with ethanol. The double-stranded cDNA was treated with 500 U of S1 nuclease for 1 h at 37°C to remove any hairpin loops and to generate blunt ends; extracted with phenol, chloroform, and ether; and precipitated with ethanol. cDNA (2.5 μg) was then ligated to 5 μg of a 5′-phosphorylated *Eco*RI synthetic oligonucleotide linker (GGAATTCC) for 16 h at 4°C in 50 μl of 70 mM Tris hydrochloride (pH 7.6)–5 mM $MgCl_2$–5 mM DTT–1 mM ATP containing 10 U of DNA ligase. The ligation mix was digested with *Eco*RI (20 U/μg of linker) for 4 h at 37°C, phenol and chloroform extracted, and adjusted to 1 mM EDTA and 0.3 M NaCl. Excess linkers were removed by size fractionation on a Sephadex G-100 column (10 by 1 cm) in 10 mM Tris hydrochloride (pH 8.1)–1 mM EDTA–0.3 M NaCl.

The cDNA was then cloned into the λgt11 expression vector described by Young and Davis (21). λgt11 DNA was cut with *Eco*RI and treated with calf intestinal alkaline phosphatase to minimize religation in the absence of an insert. The cDNA was ligated into the vector DNA by incubation for 16 h at 4°C and packaged into lambda phage in vitro (18). The resulting library was composed of ca. 120,000 recombinant phage, as assessed by the loss of β-galactosidase activity which results from insertion of DNA into the *Eco*RI site of wild-type λgt11. After amplification, the library was screened according to the procedure of Benton and Davis (2). Phage that contained cDNA complementary to DIRS-1 were isolated through two cycles of plaque purification and hybridization. cDNA inserts from these lambda phage clones were excised with *Eco*RI and recloned into pBR322.

DNA sequence analysis. The ends of the cDNA clones were sequenced by the chain termination method of Sanger et al. (17). M13mp8 and M13mp9 were used as the sequencing vectors, and JM103 was used as the host strain for transformation. M13 recombinants were prepared by the procedures of Messing et al. (11). Sequencing reactions were carried out as described previously (23). The samples were analyzed on 6% acrylamide (19:1 acrylamide-bisacrylamide) gels containing 7 M urea (16).

Primer extension. Primers for reverse transcription of DIRS-1 RNAs were prepared by cloning DNA segments near the 5′ ends of the DIRS-1 coding regions into M13 bacteriophage. The orientation of the DNA inserts was selected so that the DNA strand synthesized from the M13 phage template was complementary to the RNA. For the 4.5-kb DIRS-1 transcript, a 219-bp *Sau*3a-to-*Eco*RI fragment from the left end of DIRS-1 internal sequence was cloned into *Bam*HI- and *Eco*RI-digested M13mp8 (see Fig. 4a). For the E1 RNA, a 604-bp *Sau*3a-to-*Eco*RI fragment from the right end of DIRS-1 internal sequence (see Fig. 6a) was cloned into *Bam*HI- and *Eco*RI-digested M13mp8. The structures of the M13 clones were verified by DNA sequencing. For synthesis of the primer, 20 μl of single-stranded M13 DNA solution (ca. 0.1 to 0.5 μg of DNA) was annealed to 6 ng of a 17-base oligonucleotide M13 primer (P-L Biochemicals) in 70 mM Tris hydrochloride (pH 7.6)–10 mM $MgCl_2$. The reaction mixture was adjusted to contain 10 mM DTT; 0.5 mM dGTP, dCTP, and dTTP; 20 μCi of [^{32}P]dATP (400 Ci/mmol); and 5 U of *E. coli* DNA polymerase I large fragment were added. After 15 min of incubation at room temperature, dATP was added to a concentration of 0.5 mM, and synthesis was continued for 15 min. The reaction was terminated by heating to 70°C for 10 min.

The asymmetrically labeled double-stranded primer was excised from the M13 DNA by digestion with 10 U of each of the appropriate restriction enzymes for 30 min at 37°C. For the 4.5-kb and related RNAs, the 112-bp primer fragment was excised with *Sau*3a and *Eco*RV and purified on a 6% acrylamide gel. For E1 RNA, the 604-bp *Sau*3a-*Eco*RI fragment was excised and purified on a 1.5% agarose gel, recovered, and digested with *Rsa*I and *Dde*I. The 95-bp

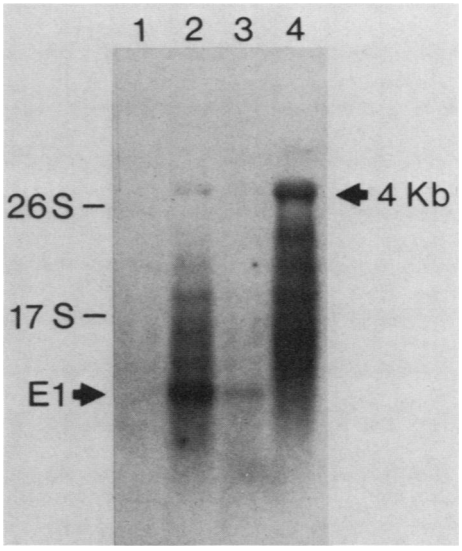

FIG. 1. Developmental regulation of DIRS-1 RNAs. RNA filter hybridization of *D. discoideum* RNAs with labeled DIRS-1 DNA. Four micrograms of poly(A)$^+$ RNA isolated from vegetative cells (lane 1) heat shock-treated vegetative cells (lane 2), and 1-h (lane 3) and 15-h (lane 4) filter-developed cells were size fractionated on a 1% agarose–formaldehyde gel and transferred to Gene Screen. The filter was hybridized for 48 h at 65°C with 10^6 cpm of nick-translated DNA from clone pB41-6 per ml. The positions of the ribosomal RNAs are indicated as 26S and 17S. The positions of the ca. 4.5-kb and E1 RNAs are indicated.

*Rsa*I-to-*Dde*I primer fragment was purified on a 6% acrylamide gel. The primers were eluted from macerated gel slices by diffusion for 16 h at 37°C in 0.5 M NaCl–10 mM Tris hydrochloride (pH 7.5)–10 mM EDTA and precipitated with ethanol in the presence of 10 μg of carrier tRNA.

The primers were hybridized to 20 μg of poly(A)$^+$ RNA for 16 to 24 h at 37°C in 20 μl of 70% formamide–400 mM NaCl–40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); pH 6.4]–1 mM EDTA (19), precipitated with ethanol, and redissolved in 50 μl of 50 mM Tris hydrochloride (pH 8.1)–50 mM NaCl–10 mM $MgCl_2$–5 mM DTT–1 mM each of dATP, dTTP, dCTP, and dGTP; 30 U of AMV reverse transcriptase was added, and the reaction mix was incubated for 2 h at 42°C (J. Pinkham and T. Platt, manuscript submitted). The reactions were terminated by precipitation with ethanol, and the products were size fractionated on 6% acrylamide–7 M urea sequencing gels (16).

RESULTS

Developmental regulation of DIRS-1-related RNAs. Transcription of DIRS-1 is regulated during development of *D. discoideum* (7, 22). Figure 1 shows that poly(A)$^+$ DIRS-1-related RNAs were present at very low levels in vegetative cells, below the limit of detection of the RNA filter hybridization experiment (but see lane 2 of the primer extension experiments in Fig. 4B and 6B). DIRS-1 RNAs began to accumulate within 1 h of development (lane 3) and increased significantly by 15 h of development (lane 4). There was a predominant RNA species of 4.5 kb as well as a heteroge-

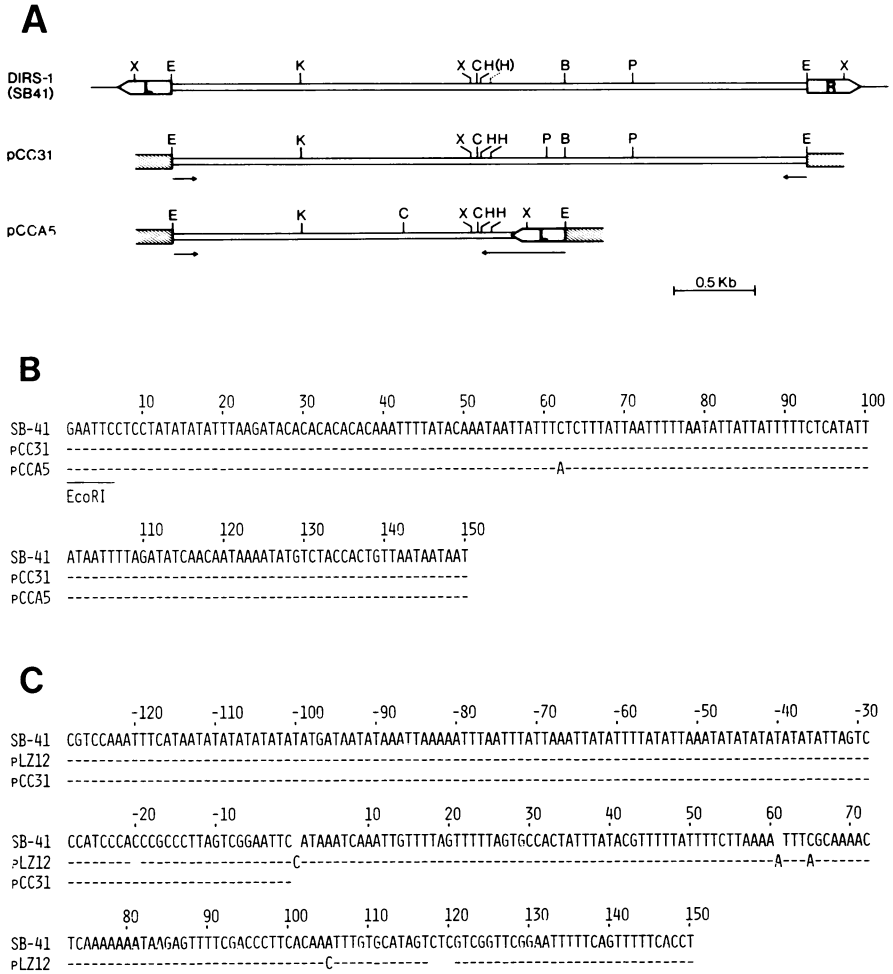

FIG. 2. Characterization of two cDNA clones complementary to DIRS-1. (a) Restriction maps of DIRS-1 and cDNA clones pCC31 and pCCA5. The restriction map of DIRS-1 is from genomic clone SB41. The *Hin*dIII restriction site indicated by a dashed line is absent from SB41 because of a single nucleotide substitution (see Fig. 3) but is present in all other genomic copies of DIRS-1. L and R denote left and right inverted terminal repeats. Hatched boxes represent pBR322 sequences. The portions of the cDNA clones that have been sequenced are indicated by arrows under the restriction maps. Abbreviations: B, *Bgl*II; C, *Cla*I; H, *Hin*dIII; K, *Kpn*I; R, *Eco*RI; P, *Pvu*II; X, *Xba*I. (B) Comparison of the DNA sequence of the 5′ ends [left ends in (A)] of cDNA clones pCC31, pCCA5, and the unique internal region of DIRS-1 (SB41) adjacent to the left-terminal repeat. The 5′ ends of the cDNA clones were isolated as *Eco*RI-*Alu*I fragments and were cloned into *Eco*RI- and *Hin*cII-digested M13mp9. The DNA sequence is written 5′ to 3′ [left to right in (A)] and numbered from the *Eco*RI site. Dashes indicate homology with the DIRS-1 sequence (22, 23; Zuker, Ph.D. thesis). (C) Comparison of the DNA sequence of the right ends of cDNA clones pLZ12 and pCC31 with the region of DIRS-1 spanning the junction of unique internal sequence and the right-terminal repeat. The sequence of pLZ12 was kindly provided by J. Morgenstern, R. Chisholm, and H. F. Lodish. The 3′ *Eco*RI-*Hin*dIII fragment of pCC31 was isolated and cloned into *Eco*RI- and *Hin*dIII-digested M13mp9. The DNA sequence is written 5′ to 3′ [left to right in (A)] and numbered in both directions from the *Eco*RI site. Dashes indicate homology with the DIRS-1 sequence (22, 23; Zuker, Ph.D. thesis).

neous population of smaller RNAs. Based solely on its size, the 4.5-kb RNA might represent a full-length transcript of the DIRS-1 element. This suggestion is supported by the observation that each of 10 subfragments of the unique internal 4.1-kb of DIRS-1 sequence hybridizes to the 4.5-kb RNA (7). One 900-bp RNA, which we call E1, was present at 1 h of development but disappeared as development proceeded (22) (lanes 3 and 4, Fig. 1). In heat-shock-treated vegetative cells most, if not all, of the developmentally induced DIRS-1 RNAs were transcribed; however, E1 accumulated to a disproportionately greater extent (lane 2, Fig. 1).

Isolation of cDNA clones. To determine whether the 4.5-kb RNA represented a transcript of an intact DIRS-1 element, we isolated cDNA clones complementary to DIRS-1 RNAs. We screened ca. 320,000 recombinant λgt11 phage and identified seven cDNAs that hybridized to DIRS-1. Three of these contained 4.1-kb inserts, one contained a 2.3-kb insert, and three contained inserts of less than 200 bp. We chose the 2.3-kb cDNA and one of the 4.1-kb clones for examination. The cDNA inserts, flanked by *Eco*RI restriction sites, were recloned into pBR322. Restriction maps of these clones are presented in Fig. 2A. pCC31 is a 4.1-kb cDNA clone delimited by *Eco*RI restriction sites. Restriction mapping (Fig. 2A) and DNA sequence analysis of the ends of pCC31 (Fig. 2B and C) indicated that it represented a complete transcript of the internal 4.1-kb *Eco*RI segment of a DIRS-1 element. There were a small number of restriction site polymorphisms between pCC31 and the DIRS-1 element contained in SB41 (Fig. 2A), indicating that this cDNA was probably derived from a different genomic copy of DIRS-1.

Since the cDNA molecules were truncated by *Eco*RI digestion during construction of the cDNA library, we could not determine from pCC31 whether transcription of the 4.5-kb RNA extended beyond the *Eco*RI sites on either side. A previously identified cDNA clone complementary to DIRS-1, pLZ12 (24), contains sequences that are homologous to 230 bp at the 3′ end of pCC31 and extend at least 152 bp into a right-terminal repeat sequence (Fig. 2C). We do not know whether pLZ12 was derived from a transcript of an intact DIRS-1 element; however, the structure of pLZ12 suggests that transcription of DIRS-1 can proceed through at least half of the right-terminal repeat. Transcription termination near the end of the right-terminal repeat, adjacent to flanking sequences, was suggested by the presence of three polyadenylation signals (AATAAA) in the distal 70 nucleotides of the repeat sequence (TTTATT at positions −278, −282 and −334 in Fig. 7) (23).

The restriction map of most of cDNA clone pCCA5 was similar to that of the left half of DIRS-1 (Fig. 2A). The left (5′) end of pCCA5 had the same nucleotide sequence as pCC31 and DIRS-1 (Fig. 2B). The similarity in the restriction maps extended to the internal *Hin*dIII restriction sites. pCCA5 then contained the *Xba*I and an *Eco*RI site spaced 240 bp apart, a pattern reminiscent of the DIRS-1 terminal-repeat structure. We sequenced 551 nucleotides at the 3′ end of pCCA5 extending from the 5′ *Hin*dIII site to the *Eco*RI site (Fig. 3). The sequence of the first 223 nucleotides from the 5′ *Hin*dIII site was virtually identical to the internal sequence of the equivalent region of genomic DIRS-1 clone SB41. There were a small number of single-base differences, one of which accounts for the absence of the second *Hin*dIII restriction site in SB41 (Fig. 2A). Beginning at position −328 (224 nucleotides from the 5′ *Hin*dIII restriction site) the cDNA sequence was nearly identical to that of the SB41 left-terminal repeat. Transcription of the RNA from which pCCA5 was derived was presumed to have initiated upstream from the 5′ *Eco*RI site of the cDNA, probably in a DIRS-1 left-terminal repeat. Approximately halfway through the normal DIRS-1 internal sequence, the genomic template

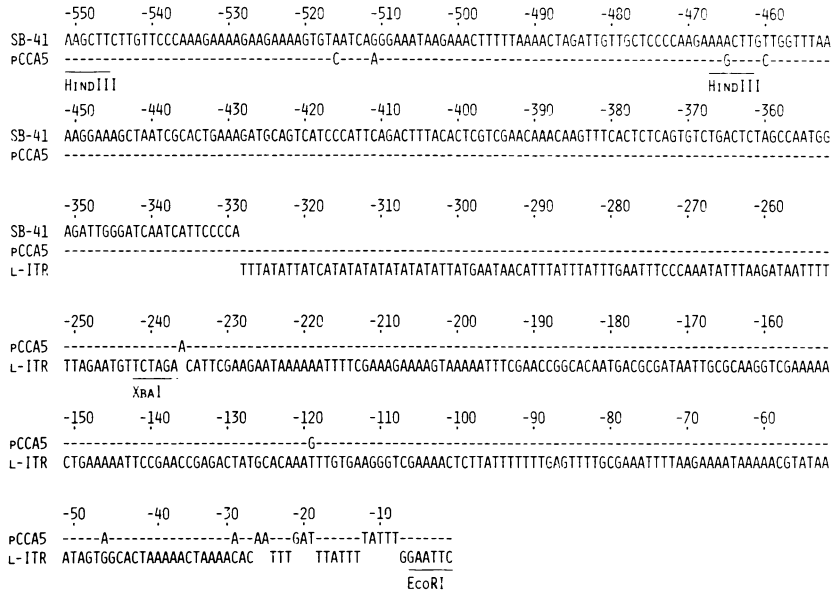

FIG. 3. DNA sequence of the 3′ end of cDNA clone pCCA5. A total of 551 nucleotides from the 3′ end of pCCA5 were sequenced. The *Hin*dIII-*Eco*RI fragment was cloned into *Hin*dIII- and *Eco*RI-digested M13mp8 and M13mp9. The *Hpa*II site at position −186 was used to clone the *Hpa*II-*Hin*dIII fragment into *Acc*I- and *Hin*dIII-digested M13mp9 and to clone the *Hpa*II-*Eco*RI fragment into *Acc*I- and *Eco*RI-digested M13mp8. The sequence of pCCA5 is numbered backward from the *Eco*RI site at the right end of the cDNA clone. The sequence of pCCA5 is homologous to two noncontiguous regions of DIRS-1 (see Fig. 2A). The sequence labeled SB41 is the unique internal DIRS-1 sequence extending rightward from the *Hin*dIII site (22). The sequence labeled L-ITR is the left-terminal repeat of SB41 (23).

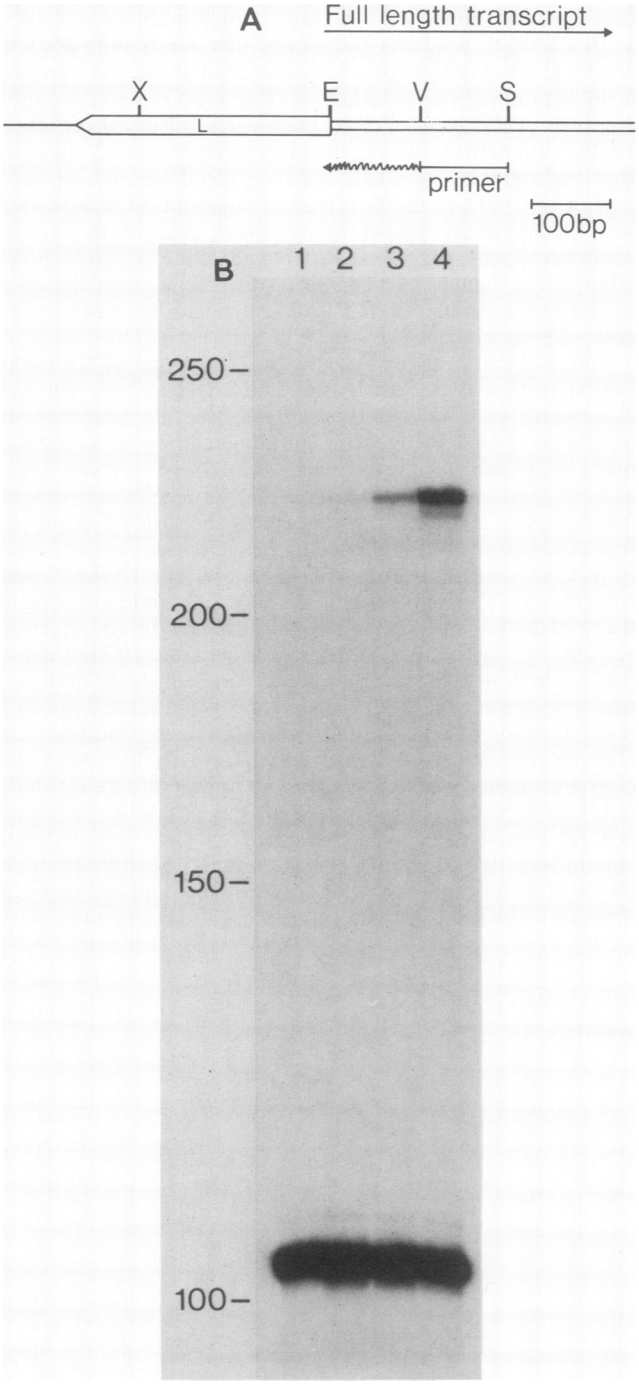

FIG. 4. Initiation of transcription of the 4.5-kb RNA. The initiation point of the 4.5-kb RNA was determined by primer extension. The structure of the primer and its relationship to the left end of DIRS-1 are shown in (A). The left terminal repeat is indicated by the box labelled L, and the open box represents internal DIRS-1 sequences. The direction of transcription of the 4.5-kb RNA is indicated. Restriction sites: E, *Eco*RI; S, *Sau*3a; V, *Eco*RV; X, *Xba*I. The labeled *Sau*3a-*Eco*RV fragment was hybridized to RNA and extended by reverse transcription, indicated by the wavy line. (B) The cDNA extension products were size fractionated on a 6% acrylamide–7 M urea gel panel B. The primer was hybridized to tRNA (lane 1) and poly(A)$^+$ RNA from vegetative cells (lane 2), heat shock-treated vegetative cells (lane 3), and 15-h filter-developed cells (lane 4). A predominant band of 230 bp was synthesized in lanes 2 through 4. The 112-bp primer is visible at the bottom of the gel. Size standards were provided by a DNA sequencing ladder (not shown).

of pCCA5 is interrupted by a second left-terminal repeat. the genomic template of pCCA5 could have arisen through a DIRS-1-into-DIRS-1 transposition event similar to the one observed in genomic clone SB41 (4). pCCA5 reflects transcription of a genomic copy of a rearranged DIRS-1 element that would appear as a small *Eco*RI fragment on a genomic Southern blot.

Transcription initiation from the left-terminal repeat. We mapped the precise point at which transcription of the 4.5 kb RNA is initiated by using primer extension (19). For the 4.5-kb RNA and related transcripts, a 112-bp *Sau*3a-to-*Eco*RV fragment located 111 bp from the 5′ *Eco*RI restriction site was used as a primer for reverse transcriptase (Fig. 4A). The uniformly labeled double-stranded primer was hybridized to 40 μg of poly(A)$^+$ RNA, extended by reverse transcription, and size fractionated on a 6% acrylamide–7 M urea sequencing gel (Fig. 4B). One predominant cDNA extension product of 230 bp was primed by the 112-bp *Sau*3a-*Eco*RV fragment by using RNA from vegetative cells, heat shock-treated vegetative cells, and 15-h filter-developed cells. Quantitatively minor products were also generated. The major initiation site was located 9 bp upstream from the center of the *Eco*RI restriction site that defines the end of the left-terminal repeat (Fig. 5).

Sequences homologous to the *Drosophila* consensus heat-shock promoter (13, 14) were located at positions −205 and −216 upstream from the major initiation site. A sequence identical to the *D. discoideum* consensus TATA box TATAAA(T/A)A (9) was located at position −27. Transcription initiated at the end of a short oligo(dT) sequence, as has been observed for a number of *Dictyostelium* genes (9). Since DIRS-1 transcription initiates at the same position during both normal development and heat shock, the same promoter apparently directs transcription under both sets of conditions. We also mapped the transcriptional initiation of the 4.5-kb and related RNAs by S1 nuclease protection experiments, with similar results (data not shown).

Transcription initiation from the right-terminal repeat. Transcription of the E1 RNA occurs on the opposite strand of DIRS-1 from that of the 4.5-kb RNA (15) and appears to be directed by the heat-shock promoters in the right-terminal repeat. We mapped the transcription initiation of the E1 RNA by primer extension. The primer was a 98-bp *Dde*I-*Rsa*I fragment located 332 bp from the right-terminal repeat (Fig. 6A). The asymmetrically labeled primer was hybridized to 40 μg of poly(A)$^+$ RNA from vegetative, heat-shock-treated, and 1-h filter-developed cells; extended by reverse transcription; and size fractionated on a sequencing gel (Fig. 6B). Unlike the 4.5-kb RNA, initiation of the E1 RNA is heterogeneous. A large number of cDNA extension products were generated, ranging in size from ca. 120 to 450 bp. One predominant extension product of 420 bp (indicated by the asterisk in Fig. 6B) was located 7 bp downstream from the center of the *Eco*RI restriction site that defines the end of the right-terminal repeat. The positions of the other major initiation sites are indicated in Fig. 7. E1 transcription initiated at the same positions in vegetative, filter-developed, and heat shock-treated cells (Fig. 6B). The profile was not affected by increasing the temperature of the reverse transcription reaction from 42 to 46°C, suggesting that the multiple initiation sites are not due to premature termination of the reverse transcription reaction due to RNA secondary

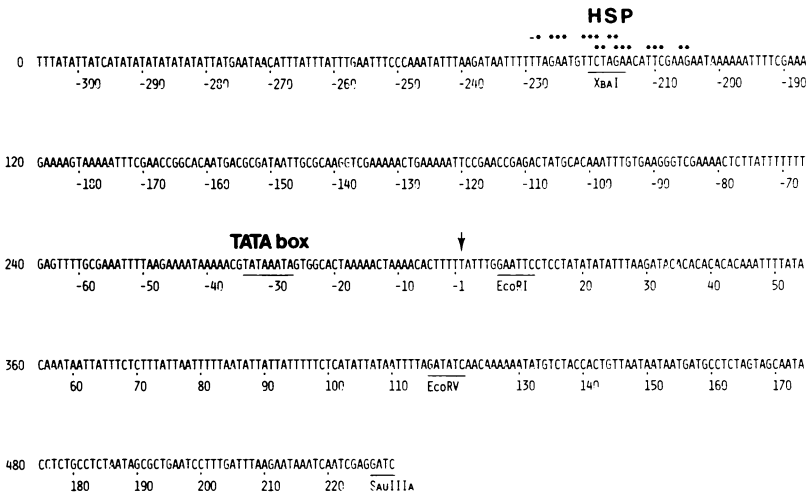

FIG. 5. Nucleotide sequence of the left-terminal repeat and internal region of DIRS-1. The nucleotide sequence of the left-terminal repeat (23) and a portion of the 5′ internal region of genomic clone SB-41 is presented. The sequence is written from the distal end of the left-terminal repeat rightward to the *Sau*3a restriction site. The size of the cDNA extension product (Fig. 4) places the initiation point of transcription 230 bp upstream from the *Sau*3a site (arrow), within the terminal repeat. The major transcription initiation site is designated position 1 (arrow). The locations of the putative TATA box and heat shock promoter (HSP) sequences are indicated. The regions of homology between the *Drosophila* and *Dictyostelium* heat shock promoters are indicated by asterisks.

structure (data not shown). A similarly heterogeneous pattern was obtained in S1 nuclease protection experiments (data not shown).

DISCUSSION

DIRS-1 is a 4.7-kb transposable element in *D. discoideum*. Transcription of DIRS-1-related RNAs is regulated during development and results in the accumulation of a heterogeneously sized population of RNA species (7, 22). This heterogeneity reflects transcription of both the 40 intact and 200 partial copies of the DIRS-1 element. cDNA clone pCC31 represents a transcript of an intact element, whereas cDNA clone pCCA5 was derived from a rearranged copy of the element. We do not know whether any of the multiple RNA species reflect processing of the 4.5-kb transcript.

Transcription of the majority of the DIRS-1-related RNAs is directed inward from the left repeat (left to right in Fig. 2A) (15; Zuker, Ph.D. thesis). Transcription initiates predominantly at a site 9 bases upstream from the center of the *Eco*RI site that defines the end of the terminal repeat. This initiation site is located at the downstream end of a short oligo(dT) sequence. An oligo(dT) sequence preceding transcriptional initiation sites is a characteristic feature of *Dictyostelium* protein-coding genes (9) and may be required for precise initiation. A sequence identical to the *Dictyostelium* consensus TATA box (9) begins at position −27. The heat shock promoter sequences (13, 14) are located at positions −205 and −216.

Although we did not directly demonstrate that this heat-shock promotor is responsible for regulating transcription of the DIRS-1 RNAs in *D. discoideum*, this suggestion was strongly supported by the observations that these RNAs were heat shock inducible and that transcription initiated at the same site during heat shock and normal development. Sequences contained in the terminal repeat are required for heat shock-inducible transcription of DIRS-1 RNAs in yeasts (5; J. Cappello, unpublished data). Recent experiments have shown that specific deletion of the heat shock promoter consensus sequences in the terminal repeat abolishes heat-inducible transcription of DIRS-1 RNA in yeasts (J. Cappello, unpublished data).

Transcription of one 900-bp cytoplasmic poly(A)$^+$ DIRS-1-related RNA, E1, exhibits a different pattern of regulation than the majority of the DIRS-1 RNAs (15, 22; Zuker, Ph.D. thesis). E1 RNA accumulates early in development, reaches its maximal level by 1 h of development, and decreases in abundance thereafter (22). Transcription of E1 occurs from the opposite strand of DIRS-1 (15) and is induced to much greater extent during heat shock than are other DIRS-1-related RNAs (Fig. 1) (15, 22; Zuker, Ph.D. thesis).

Transcriptional initiation of E1 RNA is significantly more heterogeneous than that of the 4.5-kb RNA. There are several possible explanations for this heterogeneity. The prominent 420-bp and the minor 450-bp cDNA extension products indicate that at least some of the E1 RNAs initiate near or within the right-terminal repeat. These initiation sites may be directed by the TATA box at position −42. The sequence of the first 120 nucleotides downstream from the *Eco*RI restriction site is unusual in that it contains a 100-base stretch consisting almost entirely of adenine and thymine residues (22). Consequently, another perfect match for the consensus TATA box is located within this sequence at position +45. It is possible that this and other less-perfect TATA boxes might contribute to the multiple E1 initiations. In addition, there is a 19-bp perfect inverted repeat located at positions +22 to +40 and +88 to +107. This inverted repeat could form a base-paired structure that might cause premature termination of primer extension by reverse transcriptase. Three quantitatively minor cDNA extension products of 287, 310, and 314 bp terminate at positions +106, +110, and +133, respectively. These products could result from premature termination if the base-paired structure is stable under the hybridization conditions used. We could not find any unusual structural features that would explain the major cDNA extension products of 151, 161, 167, 178, 192, 202, and 257 bp. We suggest that these cDNAs reflect authentic heterogeneity in the initiation of E1 transcription. This heterogeneity does not result from structural divergence

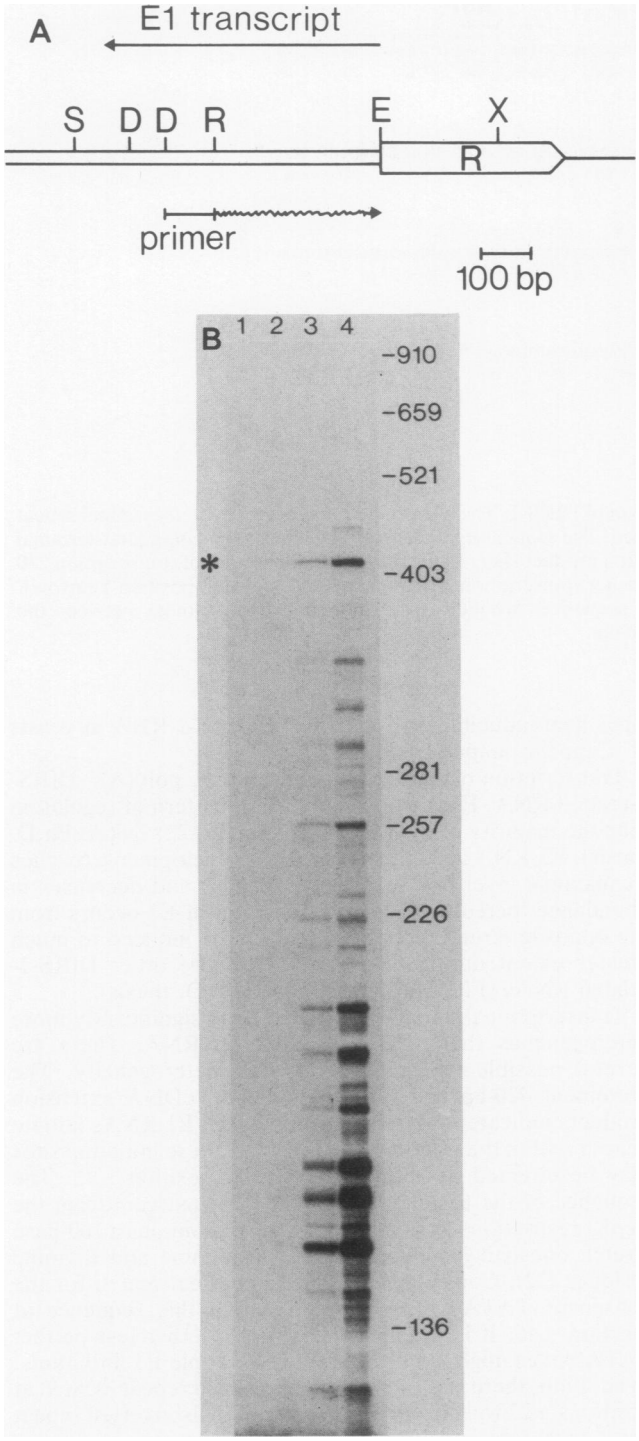

FIG. 6. Initiation of transcription of the E1 RNA. The initiation points of transcription of the E1 RNA were determined by primer extension. The structure of the primer and its relationship to the right end of DIRS-1 are indicated in panel A. The right terminal repeat is indicated by the box labeled R, and the open box represents DIRS-1 internal sequences. The direction of E1 transcription is indicated. Restriction sites: D, *Dde*I; E, *Eco*RI; R, *Rsa*I; S, *Sau*3a; X, *Xba*I. The labeled *Dde*I-*Rsa*I fragment was hybridized to RNA and extended by reverse transcription, indicated by the wavy line. (B) The cDNA extension products were size fractionated on a 6% acrylamide, 7M Urea gel. The primer was hybridized to tRNA (lane 1) and poly(A)$^+$ RNA from vegetative cells (lane 2), 1-h filter-developed cells (lane 3), and heat shock-treated vegetative cells (lane 4). Major cDNA extension products of 151, 161, 167, 178, 192, 202, 257, and 420 bp, as well as a number of quantitatively minor products, were produced. The 420-bp cDNA is indicated with an asterisk. Size standards were *Alu*I-digested pBR322.

among different genomic copies of DIRS-1, since the DNA sequences of four different DIRS-1 elements are nearly identical in the region encoding the E1 RNA (15, 22; J. Cappello, unpublished data).

Transcription of DIRS-1-related RNAs appears to be under the control of the heat shock promoters contained in the inverted terminal repeats of the element. Left- and right-terminal repeats have been sequenced from a number of genomic copies of DIRS-1 (23). All terminal repeats contain two tandem, overlapping copies of the *Drosophila* consensus heat shock promoter, CT-GAA--TTC-AG, defined by Pelham and Beinz (13, 14). These sequences are indicated by asterisks in Fig. 5 and 7. *Drosophila* heat shock promoters and synthetic oligonucleotide promoters that confer heat shock inducibility on heterologous genes typically share 8 to 10 bp of homology with the consensus sequence. The DIRS-1 heat shock promoters match the consensus sequence at 9 or 10 positions. The evolutionary conservation of the sequence of the heat shock promoter has also been suggested by numerous observations that heat shock genes and promoters can function in heterologous systems (5, 10, 13, 14; see reference 13 for additional references). Heat shock genes in heterologous cells are recognized by the cellular transcriptional machinery and are activated in response to the same stresses as the endogenous heat shock genes (5, 10, 13, 14). Parker and Topol have recently identified a protein that is specifically required for the transcription of the *Drosophila* hsp70 gene (12). This protein binds to the region of the hsp70 gene that contains the heat shock promoter sequences. Similar indications of protein binding to the heat shock promoters of the hsp70 and hsp83 genes of *Drosophila* have also been reported (20).

A consequence of the palindromic structure of the consensus heat shock promoter is the presence of an equivalent copy of the promoter in both strands of the DNA. These promoters should therefore be able to function bidirectionally. Yeasts transformed with plasmids carrying isolated DIRS-1 terminal repeats show heat shock-inducible transcription of vector sequences in both directions (J. Cappello, unpublished data). In addition, previous experiments have shown heat shock-inducible transcription directed outward from a DIRS-1 terminal repeat into flanking sequences in yeasts transformed with a DIRS-1-related clone, pB41.6 (5). The genomic fragment contained in pB41.6 was generated by a transpositional event in which a DIRS-1 element inserted into a preexisting genomic copy of DIRS-1, with the result that sequences flanking the intact DIRS-1 element are portions of DIRS-1 (4). We do not yet have any evidence for transcription outward into flanking sequences in *D. discoideum*.

Both terminal repeats of DIRS-1 appear to contain all of the upstream regulatory elements required to direct transcription of DIRS-1 RNAs. Transcription from the left repeat into the element produces the 4.5-kb RNA. A number of other heat shock-inducible RNAs are transcribed from the same strand as the 4.5-kb RNA. These smaller RNAs may reflect processing of the 4.5-kb RNA, transcription of rearranged or deleted genomic DIRS-1 templates, or transcrip-

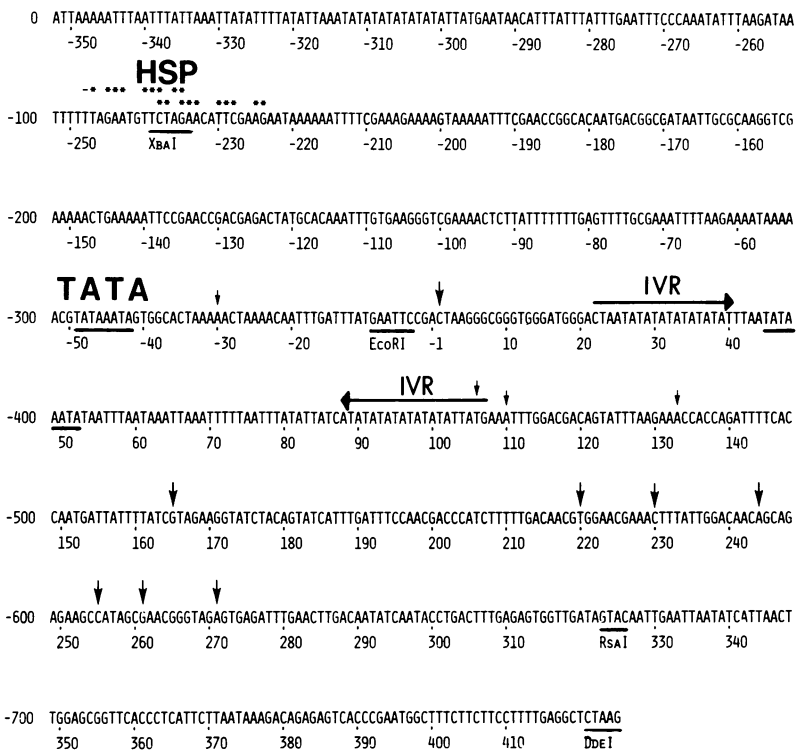

FIG. 7. Nucleotide sequence of the right-terminal repeat and internal region of DIRS-1. The nucleotide sequence of the right-terminal repeat and a portion of the 5′ internal region of genomic clone SB-41 is presented (22, 23). The sequence is written for the 5′-to-3′ strand from the distal end of the right-terminal repeat leftward to the *Dde*I restriction site of the primer (right to left in Fig. 6A). The size of the cDNA extension product indicated by the asterisk in Fig. 6B places this initiation point of transcription 420 bp upstream from the *Dde*I site, adjacent to the terminal repeat. This initiation site is arbitrarily designated position +1, and the sequence is numbered accordingly. Other major initiation sites are indicated by large arrows. Some minor initiation sites are indicated by small arrows. The locations of the putative TATA boxes are underlined. A 19-bp perfect inverted repeat is indicated by horizontal arrows (IVR) above the DNA sequence. The regions of homology between the *Drosophila* and *Dictyostelium* heat shock promoters are indicated by asterisks.

tional initiation from other heat shock promoters contained within the internal region of DIRS-1. One such internal putative heat shock promoter has been identified by DNA sequence (Zuker, Ph.D. thesis) but has not yet been shown to be functional. Transcription from the right repeat into internal sequences produces E1 RNA. In both cases transcription appears to be regulated by the heat shock promoters contained in the terminal repeats. Transcription of the 4.5-kb RNA and at least a portion of the transcription of E1 RNA initiate near the junction between the terminal repeat and internal DIRS-1 sequences. Both isolated DIRS-1 terminal repeats are capable of directing heat shock-inducible transcription in yeasts (Cappello, unpublished data). In spite of the nearly identical nucleotide sequences of the right- and left-terminal repeats (23) regulation of the 4.5-kb and E1 RNAs differs significantly. E1 RNA is only expressed during early development and is induced to a much greater extent by heat shock than are the other DIRS-1 RNAs.

ACKNOWLEDGMENTS

This work was supported by grants PCM 79-00839 from the National Science Foundation and GM-32751 from the National Institutes of Health. S.C. was supported by a Postdoctoral Fellowship from the Natural Sciences and Engineering Research Council of Canada. J.C. was supported by Postdoctoral Fellowship PF-2235 from the American Cancer society.

We thank Naomi Cohen and Karl Handlesman for technical assistance and Miriam Boucher for secretarial assistance.

LITERATURE CITED

1. **Alton, T. H., and H. F. Lodish.** 1977. Developmental changes in messenger RNA's and protein synthesis in *Dictyostelium discoideum*. Dev. Biol. **60**:180–206.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λ recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
3. **Blumberg, D. D., and H. F. Lodish.** 1980. Complexity of nuclear and polysomal RNAs in growing *Dictyostelium discoideum* cells. Dev. Biol. **78**:268–284.
4. **Cappello, J., S. M. Cohen, and H. F. Lodish.** 1984. *Dictyostelium* transposable element DIRS-1 preferentially inserts into DIRS-1 sequences. Mol. Cell. Biol. **4**:2207–2213.
5. **Cappello, J., C. Zuker, and H. F. Lodish.** 1983. Repetitive *Dictyostelium* heat-shock promoter functions in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:591–598.
6. **Chirgwin, J. M., A. E. Pryzbyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
7. **Chung, S., C. Zuker, and H. F. Lodish.** 1983. A repetitive and apparently transposable DNA sequence in *Dictyostelium discoideum* associated with developmentally regulated RNAs. Nucleic Acids Res. **11**:4835–4852.
8. **Efstradiatis, A., F. C. Kafatos, A. M. Maxam, and T. Maniatis.** 1976. Enzymatic in vitro synthesis of globin genes. Cell **7**:279–288.
9. **Kimmel, A. R., and R. A. Firtel.** 1983. Sequence organization in *Dictyostelium*: unique structure at the 5′ ends of protein coding genes. Nucleic Acids Res. **11**:541–552.

10. **Lis, J., N. Costlow, J. deBanzie, D. Knipple, D. O'Connor, and L. Sinclair.** 1982. Transcription and chromatin structure of *Drosophila* heat shock genes in yeast, p. 57–62. *In* M. Schlesinger, M. Ashburner, and A. Tissiers (ed.), Heat shock from bacteria to man. Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.
11. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun sequencing DNA. Nucleic Acids Res. **9**:309–321.
12. **Parker, C. S., and J. Topol.** 1984. A *Drosophila* RNA polymerase II transcription factor specific for the heat-shock gene binds to the regulatory site of an hsp70 gene. Cell **37**:273–283.
13. **Pelham, H. R. B.** 1982. A regulatory upstream promoter element in the *Drosophila* HSP 70 heat shock gene. Cell **30**:517–538.
14. **Pelham, H. R. B., and M. Bienz.** 1982. A synthetic heat-shock promoter confers heat inducibility on the herpes simplex thymidine kinase gene. EMBO J. **1**:1473–1477.
15. **Rosen, E., A. Sivertsen, and R. A. Firtel.** 1983. An unusual transposon encoding heat shock inducible and developmentally regulated transcripts in *Dictyostelium*. Cell **35**:243–251.
16. **Sanger, F., and A. R. Coulson.** 1978. The use of thin gels for DNA sequencing. FEBS Lett. **87**:107–110.
17. **Sanger, F., S. Niclen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
18. **Scalenghe, F., E. Turco, J. E. Edstrom, V. Pirrotta, and M. Melli.** 1981. Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. Chromosoma **82**:205–216.
19. **Treisman, R., N. J. Proudfoot, M. Shander, and T. Maniatis.** 1982. A single-base change at a splice site in a B^O-thalassemic gene causes abnormal RNA splicing. Cell **29**:903–911.
20. **Wu, C.** 1984. Two protein binding sites in chromatin implicated in the activation of heat shock genes. Nature (London) **309**:229–234.
21. **Young, R. A., and R. W. Davis.** 1983. Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. U.S.A. **80**:1194–1198.
22. **Zuker, C., J. Cappello, R. Chisholm, and H. F. Lodish.** 1983. A repetitive *Dictyostelium* gene family that is induced during differentiation and by heat shock. Cell **34**:997–1005.
23. **Zuker, C., J. Cappello, H. F. Lodish, P. George, and S. Chung.** 1984. A *Dictyostelium* transposable element DIRS-1 has 350 bp inverted repeats which contain a heat shock promoter. Proc. Natl. Acad. Sci. U.S.A. **81**:2660–2664.
24. **Zuker, C., and H. F. Lodish.** 1981. Repetitive DNA sequences cotranscribed with developmentally regulated *Dictyostelium discoideum* mRNAs. Proc. Natl. Acad. Sci. U.S.A. **78**:5386–5390.