MOLECULAR AND CELLULAR BIOLOGY, Oct. 1984, p. 2207–2213
0270-7306/84/102207-07$02.00/0

Vol. 4, No. 10

Dictyostelium Transposable Element DIRS-1 Preferentially Inserts into DIRS-1 Sequences

JOE CAPPELLO,* STEPHEN M. COHEN, AND HARVEY F. LODISH

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

Received 26 April 1984/Accepted 25 July 1984

Sequence analysis of genomic clones containing the intact *Dictyostelium* transposable element DIRS-1 reveals that in five of six cases DIRS-1 has inserted into other DIRS-1 sequences. The nucleotide sequences just beyond the endpoints of the terminal repeats of five different genomic clones can be aligned with different regions of the internal nucleotide sequence of DIRS-1. In the three genomic clones which contain flanking sequences on both sides of the element, both flanking sequences are homologous with DIRS-1. In one of these clones, both extended flanking sequences represent the full 4.1-kilobase *Eco*RI fragment of DIRS-1, which has been interrupted by the insertion of an intact DIRS-1 element. There is no duplication or deletion (except possibly 1 base) of the DIRS-1 sequence upon insertion of a second DIRS-1 transposon. DIRS-1-into-DIRS-1 insertions can occur in either a colinear or inverted orientation with respect to the target sequence; the target sequence need not be an intact DIRS-1 element. We also describe a cDNA clone which could be derived by transcription of a sequence that resulted from a DIRS-1-into-DIRS-1 insertion and discuss its significance concerning the function of the heat-shock promoters found in the terminal repeats of DIRS-1 and in other DIRS-1-related sequences.

DIRS-1 is a 4.7-kilobase (kb) repetitive and apparently transposable *Dictyostelium discoideum* DNA segment. Its expression is induced by cellular stresses such as heat shock, high cell density, and the initiation of development (3, 13). DIRS-1 consists of a 4.1-kb unique sequence flanked by ca. 330-base-pair (bp) inverted repeats. Partial DNA sequencing of DIRS-1 indicates that it encodes at least three long open reading frames and that it is transcribed into RNA which is polyadenylated and associated with polysomes (13, 13a; C. Zuker, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass., 1983). Compared with other eucaryotic transposons and transposon-like elements, DIRS-1 possesses a number of unusual features (13a). The left and right terminal repeats of DIRS-1 are of different lengths; all right repeats are longer than left repeats by a 28-bp extension of the canonical 330-nucleotide sequence (13a). In contrast, terminal repeats flanking other known transposons are nearly identical. There is no duplication of any nucleotides surrounding the ends of the repeats of three different cloned DIRS-1 elements (13a). In contrast, insertion site duplications are a common consequence of the insertion of most transposable elements. Two other transposons, Tc1 of *Caenorhabditis elegans* and Tn*554* of *Staphylococcus aureus*, do insert without apparent duplication of host sequences (8, 9).

The DIRS-1 terminal repeats contain heat-shock promoters which appear to be involved in its transcription. A highly conserved region of the terminal repeat contains a 14-nucleotide sequence that is nearly identical to the consensus heat-shock promoter sequence of *Drosophila melanogaster* (13; 13a). When introduced into *Saccharomyces cerevisiae* the DIRS-1 terminal repeat stimulates heat-shock-inducible transcription of adjacent sequences (2). The major DIRS-1 transcripts induced during heat shock and development in *Dictyostelium* species initiate within the terminal repeats ca. 210 bp from the heat-shock promoters (S. M. Cohen, J. Cappello, and H. F. Lodish, Mol. Cell. Biol., in press). One of the major transcripts is ca. 4.5 kb in length and represents the full-length transcription of the DIRS-1 element (13; Cohen et al., in press).

At first glance, the 40 intact copies of DIRS-1 appear to be dispersed throughout the *Dictyostelium* genome, typical of most known transposons (3; Zuker, Ph.D. thesis). Southern hybridization analysis of *Eco*RI-digested genomic DNA generates an intense band of 4.1 kb, which is produced from the 40 intact genomic copies of DIRS-1. The large number of single-copy *Eco*RI fragments of different sizes are from the remaining 200 partial DIRS-1-related fragments. We presumed that the heterogeneous DIRS-1-related fragments, generated by *Eco*RI digestion, were deletions of intact DIRS-1 elements. However, the studies reported here provide another explanation for the generation of some of these fragments. The nucleotide sequences immediately flanking five cloned intact DIRS-1 elements are homologous to internal sequences of DIRS-1. In one clone, SB41, the extended flanking regions can be joined to produce an intact 4.1-kb *Eco*RI fragment of DIRS-1. We propose that these genomic arrangements of DIRS-1 sequences originated from DIRS-1 insertions into preexisting DIRS-1-related sequences. All of the DIRS-1 genomic clones that we have examined contain such insertions. The apparent preference for DIRS-1-related target sequences might reflect an important characteristic of the mechanism of transposition of DIRS-1.

MATERIALS AND METHODS

Restriction and hybridization mapping. Restriction maps of genomic clones were generated as described previously (3; Zuker, Ph.D. thesis). Restriction enzymes were obtained from Bethesda Research Laboratories, Inc., and New England Biolabs. The hybridization map of clone SB41 was generated by filter hybridization experiments as detailed in the text. The electrophoresis of restriction fragments of SB41, transfer to nitrocellulose filters, and hybridization to

* Corresponding author.

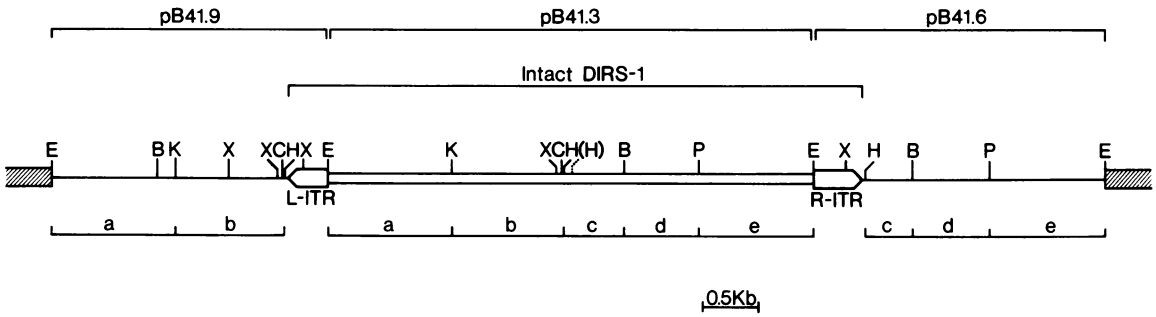

FIG. 1. Restriction and hybridization map of genomic clone SB41. The hatched boxes at each end of the map denote the lambda vector sequences. The intact DIRS-1 element of SB41 is boxed and indicated above the map. The L- and R-ITR sequences are indicated. The three *Eco*RI subclones of SB41, pB41.9 (left), pB41.3 (middle), and pB41.6 (right), are indicated. The hybridization map of SB41, as illustrated below the restriction map, was generated by digesting pB41.9 and pB41.6 with a combination of *Eco*RI and one of the following restriction enzymes: *Kpn*I, *Xba*I, *Cla*I, *Bgl*II, *Hin*dIII, or *Pvu*II. These digestions were size fractionated on a 1% agarose gel and transferred to nitrocellulose filters. The filters were hybridized separately with nick-translated fragments of pB41.3 corresponding to segments a, b, c, d, and e. Fragments of the flanking regions which hybridize with each of these probes are noted. Restriction enzyme sites are abbreviated as follows: E, *Eco*RI; K, *Kpn*I; X, *Xba*I; C, *Cla*I; H, *Hin*dIII; B, *Bgl*II; and P, *Pvu*II. The H denotes the position of a second *Hin*dIII site present in all DIRS-1 elements that we have characterized, except for the SB41 DIRS-1 element. This is due to a single-base change (data not shown).

nick-translated probes were performed as described previously (14).

DNA sequence analysis. DNA sequencing was carried out according to the chain termination procedure of Sanger et al. (11). M13mp8 and M13mp9 were used as sequencing vectors, and JM103 was the host strain. The M13 recombinants were generated as described by Messing et al. (7). Gel electrophoresis was done as described by Sanger and Coulson (10), and chemical sequencing was done as described by Maxam and Gilbert (6). Nucleotide sequence comparisons were performed by using the SEQP sequence analysis program (L. Domier, University of Minnesota).

RESULTS

Clone SB41 is an insertion of DIRS-1 into a second DIRS-1 element. Clone SB41 is a 9-kb fragment of *Dictyostelium* genomic DNA which contains an intact 4.7-kb DIRS-1 element (Fig. 1). Digestion of SB41 with *Eco*RI generates three fragments of 2.3 (pB41.9), 4.1 (pB41.3), and 2.5 kb (pB41.6). The 4.1-kb *Eco*RI fragment, designated pB41.3, is the unique internal portion of the intact DIRS-1 element. Immediately flanking the pB41.3 fragment are the inverted terminal repeats (ITRs) L-ITR and R-ITR, associated with all complete DIRS-1 elements. The ITRs of SB41 are 332 bp (left) and 360 bp (right) (13a) and are located in the flanking *Eco*RI fragments of SB41. Therefore, region pB41.9, mapping to the left of pB41.3, contains the L-ITR and 2.0 kb of flanking sequences; region pB41.6, on the right, contains the R-ITR and 2.1 kb of flanking sequence. The sum of these sequences, 4.1 kb, coincides with the size of the *Eco*RI fragment of pB41.3, the unique portion of the intact DIRS-1 element.

The restriction maps of pB41.9 and pB41.6 are very similar to the left and right parts, respectively, of the central 4.1-kb DIRS-1 fragment, pB41.3 (Fig. 1). Several distinctive restriction sites are characteristic of all DIRS-1 elements, including pB41.3. These are the *Kpn*I site, the clustered *Xba*I, *Cla*I, and *Hin*dIII sites, the *Bgl*II site, and the *Pvu*II site. Digestion of the pB41.3 *Eco*RI fragment of DIRS-1 with the enzymes *Kpn*I, *Hin*dIII, *Bgl*II, and *Pvu*II generates five fragments designated a through e, from left to right (Fig. 1). The restriction map of SB41 shows that fragments similar in size to fragments a and b are present in the left flanking *Eco*RI subclone, pB41.9, and that fragments similar in size to fragments c, d, and e are present in the right *Eco*RI subclone, pB41.6. The order of these fragments is identical to the left and right parts of the pB41.3 segment. The presence of restriction site polymorphisms, such as those between pB41.9 and the left part of pB41.3, have been found in a number of DIRS-1 genomic and cDNA clones.

To show that the flanking sequences in the pB41.9 and the pB41.6 *Eco*RI fragments represent the left and right portions of a preexisting DIRS-1 element, these subclones were digested with the enzymes noted above, separated by gel electrophoresis, and hybridized separately with fragments a through e of pB41.3. The hybridization data (not shown) confirm the results of the restriction mapping and indicate that only pB41.3 fragments a and b are present in the pB41.9 *Eco*RI fragment and that pB41.3 fragments c, d, and e are present in the pB41.6 fragment (see Fig. 1). These data indicate that the *Eco*RI fragments flanking the intact DIRS-1 element of SB41 are themselves two segments of DIRS-1 which comprise both halves of an intact element. Thus, SB41 appears to be an insertion of DIRS-1 into a second DIRS-1 element at some point near the border of fragments b and c.

If this is the case, the nucleotide sequence immediately flanking the inverted repeats of the pB41.3 DIRS-1 should be homologous with the internal sequence of DIRS-1 near the *Hin*dIII site at the junction of fragments b and c. We have determined 3.2 kb of the 4.7-kb total nucleotide sequence of the SB41 (pB41.3) DIRS-1 element. The DIRS-1 internal sequence shown in Fig. 2 is numbered from left to right, beginning arbitrarily at the *Hin*dIII site of the SB41 (pB41.3) element. In Fig. 2A, the nucleotide sequences immediately flanking the left and right terminal repeats in pB41.9 and pB41.6 are compared with the internal sequence of the pB41.3 DIRS-1. The sequence SB41 LEFT extends from within the left terminal repeat of the intact DIRS-1 element (the boxed residues at the right) and proceeds outward into the left flanking region. Figure 2A shows that the available nucleotide sequence just outside the terminal repeat is identical to internal residues 13 to 56 of DIRS-1. The right flanking sequence extends from within the right terminal repeat in the pB41.6 segment into the right flanking region. The available sequence immediately adjacent to the right

terminal repeat is homologous to residues 58 to 113 of DIRS-1. If joined together, the left and right flanking nucleotide sequences account for every nucleotide present in the intact DIRS-1 sequence except possibly one (the T residue at position 57).

We conclude that SB41 has the structure that would result from an insertion of a DIRS-1 element between bases 56 and 58 of a second DIRS-1 element. Furthermore, there is no duplication of nucleotides at the target site. However, this presumed insertion might have produced a deletion of one nucleotide.

DIRS-1-related sequences preferred target sites for DIRS-1 insertions. The nucleotide sequences flanking the terminal repeats of six different cloned DIRS-1 elements have no apparent similarities to each other (13a). However, our results with clone SB41 prompted us to examine whether the flanking sequences of the other DIRS-1 genomic clones might be DIRS-1 internal sequences. The published nucle-

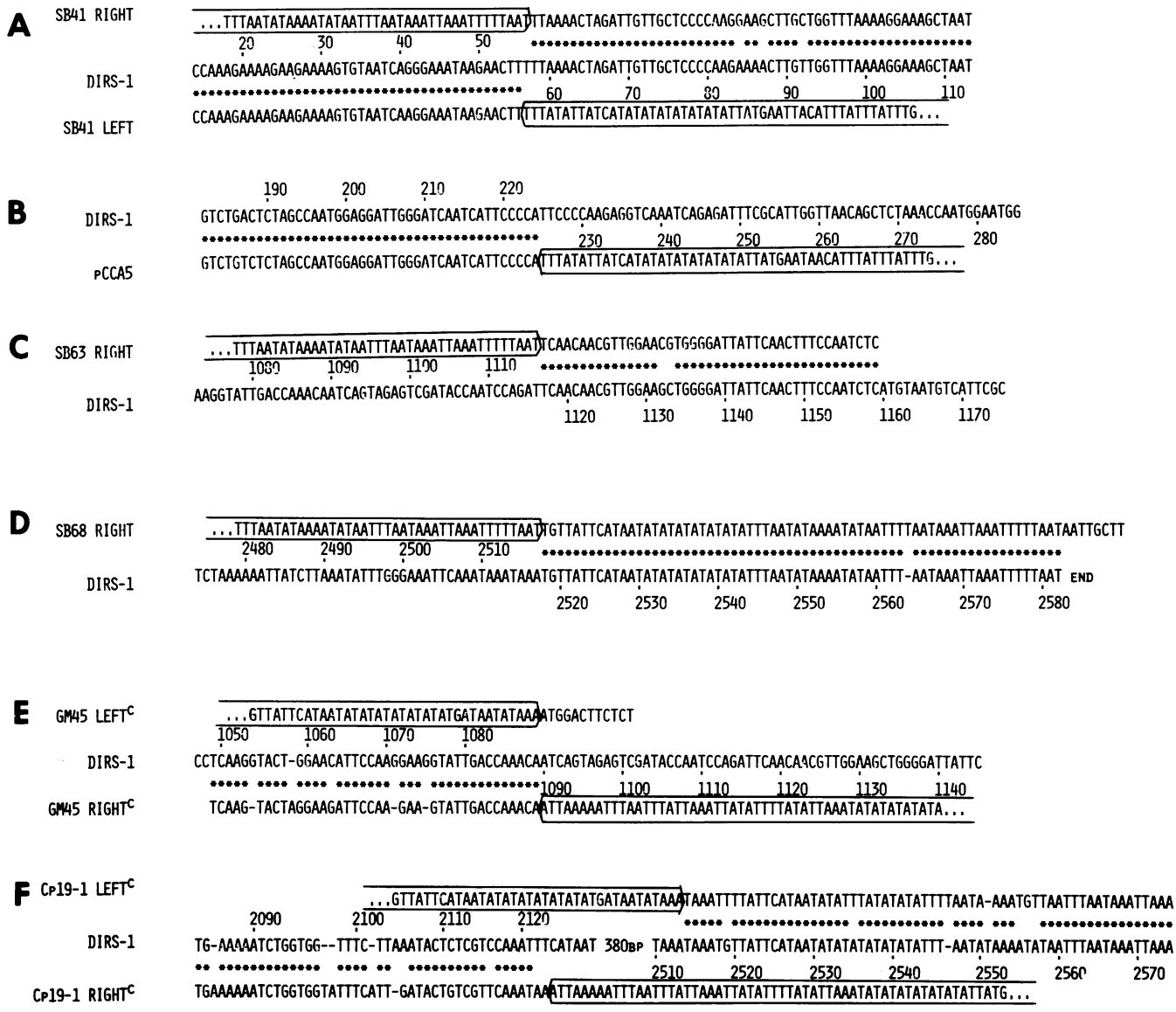

FIG. 2. Homology of DIRS-1 flanking sequences with the internal DIRS-1 nucleotide sequence. The nucleotide sequences surrounding the ends of the terminal repeats of six different genomic DIRS-1 clones, and also the cDNA clone pCCA5, were compared with our partial nucleotide sequence of DIRS-1. Except for pCCA5, all of these clones have been described previously (3). Our DIRS-1 sequence comprises 3.2 kb of a total of 4.7 kb and is continuous from the internal *Hin*dIII site to the end of the right terminal repeat of the intact DIRS-1 element of SB41. The numbering of the DIRS-1 sequence begins arbitrarily at the *Hin*dIII site. The sequences used in this comparison originate from positions indicated on the maps of the genomic clones displayed in Fig. 3. The sequences of each clone are displayed above and below the corresponding nucleotide sequence of DIRS-1 with which they are homologous. Homology is indicated by asterisks. The consensus nucleotide sequences of the terminal repeats are boxed. The sequences are written 5′ to 3′ and aligned in order to maximize homology. RIGHT and LEFT identify the positions of the flanking sequences with respect to the intact element of each clone. RIGHT^c and LEFT^c represent the complements of these sequences. The SB41 LEFT sequence is that of subclone pB41.9; the SB41 RIGHT sequence is that of pB41.6.

otide sequences flanking the intact DIRS-1 elements of genomic clones GM45, SB63, SB68, Cp19-1, and Cp19-5 were compared by computer with the available 3.2 kb of the 4.7-kb DIRS-1 element that we sequenced. Sequences that are homologous with an internal DIRS-1 sequence flank at least one side of every DIRS-1 clone except Cp19-5 (data not shown). The locations of these sequences and their relationship to the intact DIRS-1 element of each clone are diagrammed in Fig. 3.

The available sequences flanking the right side of the terminal repeats of DIRS-1 elements in clones SB63 and SB68 align almost perfectly with internal DIRS-1 sequences, at positions 1117 to 1159 and 2518 to 2582, respectively (Fig. 2C and D). Additionally, the complement of the sequences immediately adjacent to the right terminal repeats of the DIRS-1 element of GM45 and Cp19-1 are homologous with internal DIRS-1 sequences, at positions 1049 to 1089 and 2084 to 2121, respectively (Fig. 2E and F). The complement of the left flanking sequence of Cp19-1 aligns with bases 2514 and 2574. Therefore, internal DIRS-1 sequences inverted in orientation with respect to the central, intact, DIRS-1 element flank the DIRS-1 elements of clones GM45 and Cp19-1 (Fig. 4).

Thus, in those clones where the sequences flanking the intact DIRS-1 elements are known, seven of nine (all but GM45 LEFT and Cp19-5 LEFT; data not shown) are DIRS-1 related. The sequence comparisons shown in Fig. 2 establish that at least the first 40 bases immediately adjacent to the terminal repeats of the intact elements are 80 to 100% homologous with internal DIRS-1 sequences. Restriction and hybridization data on the flanking regions of these clones indicate that the homologies extend beyond these immediate 100 nucleotides (Zuker, Ph.D. thesis; Fig. 3). In at least two of the three cases where flanking regions are available for both sides of the intact DIRS-1 element, that is, clones SB41, GM45, and Cp19-1, both the left and right flanking sequences are homologous to DIRS-1. This is shown by either hybridization or nucleotide sequence homology, or both. In the case of SB41, both the hybridization data and the nucleotide sequence homology confirm that the regions flanking the DIRS-1 element add up to the intact 4.1-kb *Eco*RI fragment of DIRS-1 (see above).

Hybridization data on clone GM45 indicates that both flanking regions are homologous to internal DIRS-1 sequences as well (3). However, we failed to find any nucleotide sequence homology between the left flanking sequence of GM45 and the internal DIRS-1 sequence. This discrepancy might be explained by the short length of the sequence which is available, only 12 nucleotides. The nucleotide sequence homology found for the left flanking sequence of Cp19-1 is located ca. 70 bp from the end of the DIRS-1 right terminal repeat (positions 2514 to 2582; Fig. 2). The extent of this homology presumably marks the end of the DIRS-1-related sequence in the left flank of Cp19-1 and

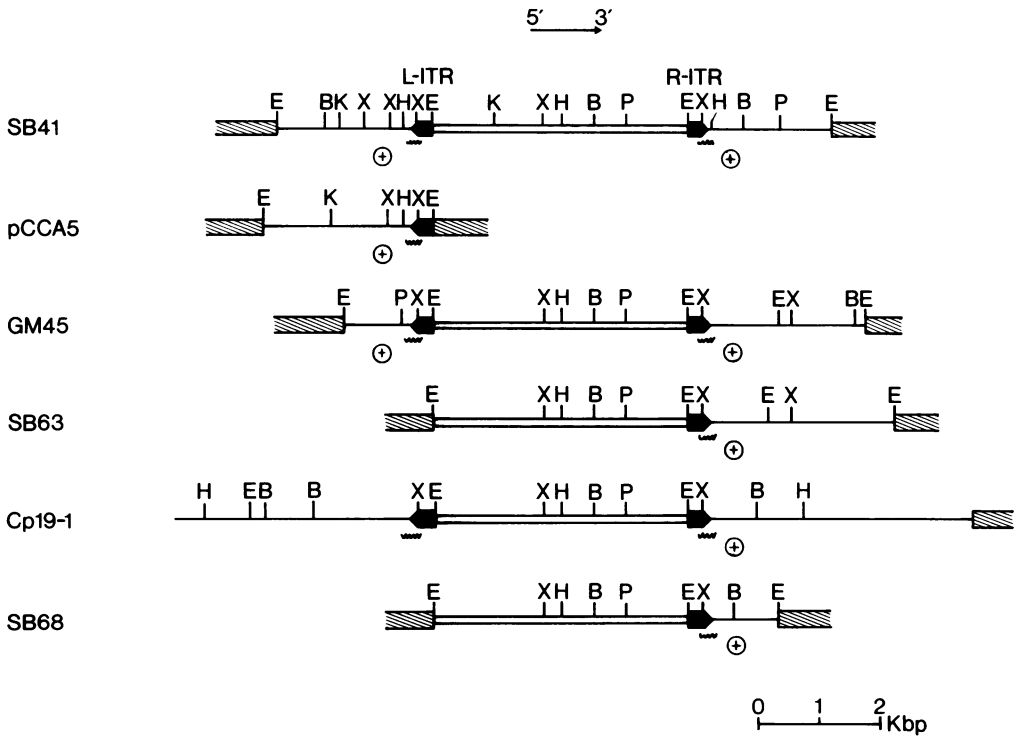

FIG. 3. Structure of DIRS-1 clones. The diagram displays the locations of the DIRS-1 flanking sequences used in the comparisons with the DIRS-1 internal sequence depicted in Fig. 2. The boxed regions of each clone are the intact DIRS-1 elements. L-ITR and R-ITR refer to the positions of the ITRs at both ends of DIRS-1. The wavy lines below the restriction maps denote the positions of the sequences which appear in Fig. 2. These extended from within the terminal repeats into the flanking regions of each DIRS-1 element. Flanking-region segments which have previously been shown to be homologous to DIRS-1 by direct hybridization are indicated by a ⊕ below the maps (3; Zuker, Ph.D. thesis): *Eco*RI digests of the cloned DNAs were hybridized to the pB41.3 fragment of SB41 (except for GM45). In clone Cp19-1 the hybridizing region is located in the *Eco*RI-*Bgl*II fragment at the right of the intact DIRS-1 element. No hybridization to the internal DIRS-1 fragment was detected with the left flanking fragments of Cp19-1 (data not shown). Data for GM45 was generated by hybridization to the 4.1-kb *Eco*RI fragment of DIRS-1 present in GM45 (3).

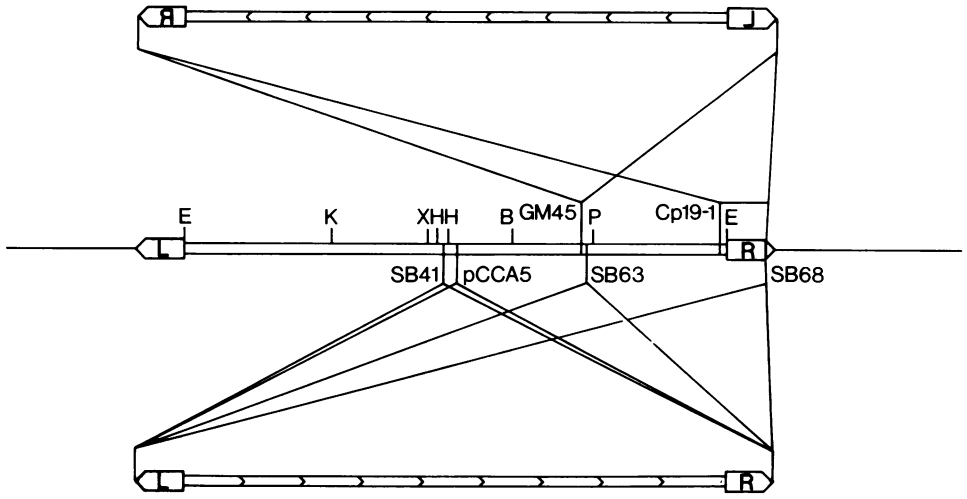

FIG. 4. Model for the generation of the sequences present in cloned DIRS-1 containing genomic segments. Intact DIRS-1 elements are flanked by internal DIRS-1 sequences. We interpret this as resulting from insertion of an intact DIRS-1 element into a second DIRS-1 or DIRS-1-related sequence. This is illustrated for five cloned genomic DIRS-1 sequences and for the cDNA clone pCCA5. The position along the DIRS-1 element denoted by the homology present in the flanking sequences is the putative point of insertion of the incoming element. In this model, the preexisting target element (the center element of the diagram) becomes the DIRS-1-related flanking sequences on both sides of the incoming element (those elements depicted above and below). As a generalization, the preexisting target sequence is depicted as an intact element. However, in at least two cases (GM45 and Cp19-1), the DIRS-1-related sequences flanking the intact element are not those expected from insertion at a single point in an intact DIRS-1 element (see the text). Only clone SB41 is known to contain both parts of the intact 4.1-kb *Eco*RI fragment of DIRS-1 flanking the central element. The insertion events associated with clones GM45 and Cp19-1 are depicted as inversions, that is, the incoming element is inverted with respect to the DIRS-1-related sequence at the target site. Those events depicted for SB41, pCCA5, SB63, and SB68 are depicted as colinear insertions, even though both flanking sequences are known only for clone SB41.

therefore explains why no hybridization with DIRS-1 was detected in this region (Fig. 3).

These data support the notion that the DIRS-1-related sequences flanking the intact DIRS-1 elements present in the above clones are a consequence of DIRS-1-into-DIRS-1 insertions. The location and the orientation of the proposed insertions, with respect to the preexisting DIRS-1-related sequence, are depicted in Fig. 4. The locations of the nucleotide homologies shown in Fig. 2 are given along the physical map of DIRS-1. The orientation of the incoming DIRS-1 element with respect to the surrounding DIRS-1-related flanking sequences is determined by the polarity of the DIRS-1 homology flanking the intact DIRS-1 element. The intact element in clones SB41, SB63, and SB68 are in the same orientation as the flanking DIRS-1-related sequences, whereas those of GM45 and Cp19-1 are inverted. Thus, in clones GM45 and Cp19-1 the incoming DIRS-1 element is inverted with respect to the DIRS-1-related sequence at the target site.

Transcription of DIRS-1 into DIRS-1 insertions. We isolated and characterized four cDNA clones complementary to DIRS-1 that were selected from a library generated by the reverse transcription of polyadenylate-containing RNA from 8-h developing *Dictyostelium* cells. Double-stranded cDNA was synthesized and ligated with *Eco*RI-digested lambda gt11 DNA (Cohen et al., in press). Three of the cDNA clones are 4.1 kb in length and, as they are colinear with DIRS-1, appear to represent end-to-end transcripts of the 4.1-kb unique portion of DIRS-1 (Cohen et al., in press).

One clone, pCCA5, is 2.3 kb in length and is homologous with the left part of DIRS-1. We sequenced 551 bases of pCCA5 leftward from the right *Eco*RI site (Fig. 3). For 338 bases this sequence is 95% homologous with the left terminal repeat of the SB41 DIRS-1 element. The sequence is identified as a left terminal repeat for the following reasons: (i) it does not contain the 28-base extension found at the end of all right repeats and (ii) it has the oligonucleotides TTGG and CT at positions 7 and 320 of the terminal repeat sequence, which are only present in left repeats (13a; Cohen et al., in press). Flanking this left repeat the sequence is homologous to residues 1 to 224 of the internal DIRS-1 sequence (Fig. 2B shows this homology from bases 184 to 224). Thus, the left part of pCCA5 is homologous with the left portion of DIRS-1. Importantly, 224 bases downstream of the *Hin*dIII site the sequence deviates from that of the internal DIRS-1 and is that of a consensus 332-bp left terminal repeat. The terminal-repeat sequence found in pCCA5 begins with the exact consensus endpoint nucleotides that are found at the end of all left repeats (13a).

These observations suggest that the RNA cloned in pCCA5 was transcribed from a genomic template that was generated by a DIRS-1-into-DIRS-1 insertion. This hypothetical genomic arrangement is diagrammed in Fig. 5. Since DIRS-1 is transcribed from left to right beginning from within the left terminal repeat (Cohen et al., in press), a DIRS-1-into-DIRS-1 insertion could contain two functional left promoters. Transcription of an RNA from the outer left promoter might continue through the interrupting element until it reaches a termination site. pCCA5 would have been formed from reverse transcription of this RNA, followed by *Eco*RI digestion of the double-stranded cDNA.

DISCUSSION

We examined the structure of several genomic clones containing DIRS-1 elements. Five of six cloned DIRS-1 elements are flanked by DIRS-1 internal sequences. The nucleotide sequences immediately flanking the terminal re-

peats of each cloned DIRS-1 element (except GM45 LEFT) are 80 to 100% homologous with an internal sequence of DIRS-1 (Fig. 2). In most of the clones the homology of the flanking regions to DIRS-1 can be shown by hybridization studies to be longer than the immediate flanking nucleotides presented in Fig. 2 (3; Zuker, Ph.D. thesis; Fig. 3). In clone SB41 the DIRS-1-related flanking sequences are clearly those that would result from a DIRS-1-into-DIRS-1 insertion (Fig. 1). The *Eco*RI fragments flanking the intact DIRS-1 element in SB41 equal the total 4.1-kb internal *Eco*RI fragment of DIRS-1. Similarly, our published hybridization data on clone GM45 demonstrates that the *Eco*RI fragments flanking the intact DIRS-1 element would comprise a DIRS-1 element from which ca. 1.5 kb of the left flanking remnant has been deleted (3). The cDNA pCCA5, insofar as it reflects the structure of its genomic template, can also be interpreted as an example of a DIRS-1 element that has been interrupted by the insertion of a second element (Fig. 2 and 5).

We failed to find a DIRS-1 homology with the left flanking sequence of one genomic clone, Cp19-5. However, since the right flanking sequence of this clone is not available, and since the DIRS-1 partial sequence used for these comparisons only comprises 3.2 kb of a total of 4.7 kb, this clone may yet be shown to contain a DIRS-1-related flanking sequence.

The locations of the flanking-region homologies, which define the proposed sites of insertion in each of our clones, are distributed throughout DIRS-1 (Fig. 2 and 4). This indicates that within the DIRS-1 sequence there is no absolute preference for insertion sites. Also, because the extent of homology in each of the flanking regions is variable, ranging from a full, uninterrupted complement of DIRS-1 sequences in SB41 to probably no more than a few hundred bases on either side of the intact element of Cp19-1, it is likely that the preexisting DIRS-1-related sequence need not be intact to function as a target of insertion. For example, the sequences flanking the left and right repeats of the DIRS-1 element in Cp19-1 are clearly internal DIRS-1 sequences (Fig. 2F). However, the left and right flanking sequences leave a gap of 393 bases in the internal DIRS-1 target sequence. A possible explanation for the origin of Cp19-1 is that a 393-bp deletion occurred in a solitary DIRS-1 sequence, followed by or cotemporaneous with an insertion of an intact DIRS-1 sequence at this site.

The orientation of the incoming element can be either colinear or inverted with respect to the preexisting DIRS-1-related sequence (Fig. 4). The DIRS-1-related flanking sequences of clones SB41, pCCA5, SB63, and SB68 are in the same orientation as the intact DIRS-1 element. In the case of the cDNA clone pCCA5, the left-terminal-repeat sequence is in the same orientation as the sequence homologous to the internal DIRS-1 sequence (Fig. 5). The flanking DIRS-1-related sequences of clones GM45 and Cp19-1 are inverted with respect to the intact element.

At least for the DIRS-1-into-DIRS-1 insertion in SB41, there is no duplication of the target nucleotide sequence at either side of the inserted DIRS-1 element (Fig. 2). Flanking sequences on both sides of the DIRS-1 element can be rejoined and aligned almost perfectly with the internal sequence of DIRS-1. The alignment accounts for every nucleotide present in the DIRS-1 sequence before insertion except one. This possible deletion of a T residue could be explained merely by an ambiguity in the definition of the final residue of the terminal-repeat sequences, which are also T's. The transposable elements Tc1 from *C. elegans* and Tn*554* from *S. aureus* also transpose without an apparent duplication of host sequences at the insertion site (8, 9).

Several hypotheses explaining the preference of DIRS-1 for self-insertion can be imagined. The simplest involves the recognition of internal DIRS-1 sequences as insertional "hot spots." Since the exact location of insertion sites within a DIRS-1 sequence can vary, a level of recognition more complex than the actual nucleotide sequence might be involved. The secondary structure of an integrated element, possibly in conjunction with specifically bound proteins, might serve as a recognition signal. However, we cannot rule out the possibility that a specific nucleotide sequence homologous to DIRS-1 is initially recognized by the incoming element. This initial homology could define a region into which the incoming element might insert that could span several kilobases of the preexisting DIRS-1 sequence.

Some transposons insert at locations that are homologous to the ends of the transposable element. A hot spot for insertion of Tn*3* into a plasmid is near a site that exhibits significant homology with the ends of Tn*3* (12). IS*1* and Tn*9* insertions also occur more frequently near sequences that contain a 7-base match with the ends of these transposons (5). A study of cloned copia elements revealed a weak

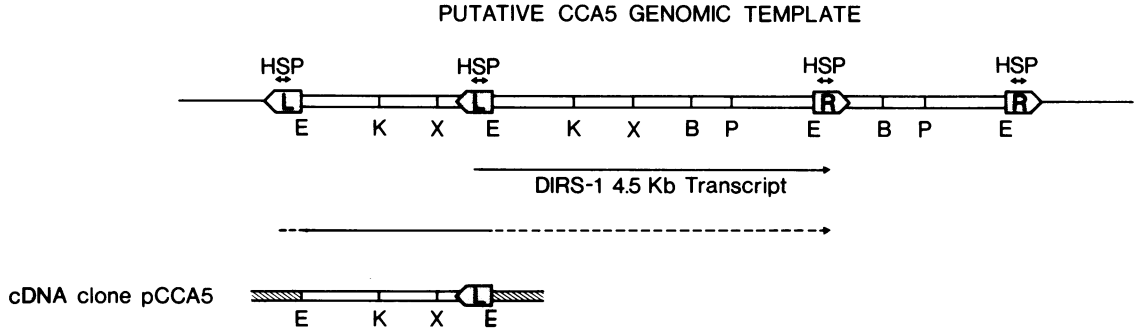

FIG. 5. Putative genomic template of cDNA clone pCCA5. The putative structure of a DIRS-1-into-DIRS-1 insertion is diagrammed. The DIRS-1 sequences are boxed, and the position and orientation of the ITRs are indicated. The positions of characteristic DIRS-1 restriction sites are shown. HSP and the double-headed arrow refer to the 18-base palindrome which contains the heat-shock promoter sequence found in all terminal repeats. The long arrows below the map indicate the proposed transcribed regions of the DIRS-1-into-DIRS-1 insertion. The 4.5-kb DIRS-1 RNA is transcribed from intact DIRS-1 elements and initiates within the left terminal repeat (Cohen et al., in press). The partially dotted arrow indicates the putative RNA which would be transcribed from the upstream promoter contained in the left terminal repeat of the interrupted DIRS-1 element. The solid region of this arrow designates the portion of this RNA cloned in pCCA5.

homology of the extended flanking regions with the copia terminal repeat. However, nucleotide sequence analysis of 150 bases flanking each element showed no homology with the copia element (4).

LITERATURE CITED

1. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.
2. **Cappello, J., C. Zuker, and H. F. Lodish.** 1984. Repetitive *Dictyostelium* heat-shock promotor functions in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:591–598.
3. **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.
4. **Dunsmuir, P., W. J. Brorein, Jr., M. A. Simon, and G. M. Rubin.** 1980. Insertion of the *Drosophila* transposable element *copia* generates a 5 base pair duplication. Cell **21**:575–579.
5. **Galas, D. J., M. P. Calos, and J. H. Miller.** 1980. Sequence analysis of Tn9 insertions in the Lac Z gene. J. Mol. Biol. **144**:19–41.
6. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
7. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
8. **Murphy, E., and S. Lofdahl.** 1984. Transposition of Tn554 does not generate a target duplication. Nature (London) **307**:292–294.
9. **Rosenzweig, B., L. W. Liao, and D. Hirsh.** 1983. Target sequences for the *C. elegans* transposable element Tc1. Nucleic Acids Res. **11**:7137–7140.
10. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**:107–110.
11. **Sanger, F., S. Niclen, and A. R. Coulson.** 1977. DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
12. **Tu, C. P., and S. N. Cohen.** 1980. Translocation specificity of the Tn*3* element: characterization of sites of multiple insertions. Cell **19**:151–160.
13. **Zuker, C., J. Cappello, R. L. Chisholm, and H. F. Lodish.** 1983. A repetitive *Dictyostelium* gene family that is induced during differentiation and by heat shock. Cell **34**:997–1005.

13a. **Zuker, C., J. Capello, H. F. Lodish, P. George, and S. Chung.** 1984. The *Dictyostelium* transposable element DIRS-1 has 350-basepair inverted terminal repeats that contain a heat shock promotor. Proc. Natl. Acad. Sci. U.S.A. **81**:2660–2664.

14. **Zuker, C., and H. F. Lodish.** 1981. Repetitive DNA sequences cotranscribed with developmentally regulated *Dictyostelium discoideum* RNAs. Proc. Natl. Acad. Sci. U.S.A. **78**:5386–5390.