MOLECULAR AND CELLULAR BIOLOGY, Jan. 1992, p. 229–239
0270-7306/92/010229-11$02.00/0

Vol. 12, No. 1

Structure of DRE, a Retrotransposable Element Which Integrates with Position Specificity Upstream of *Dictyostelium discoideum* tRNA Genes

ROLF MARSCHALEK, JÖRG HOFMANN, GERALD SCHUMANN, ROGER GÖSSERINGER, AND THEODOR DINGERMANN*

Institut für Biochemie der Medizinischen Fakultät, Universität Erlangen-Nürnberg, Fahrstrasse 17, D-8520 Erlangen, Germany

Received 2 August 1991/Accepted 11 October 1991

Different *Dictyostelium discoideum* strains contain between 2 and 200 copies of a retrotransposable element termed DRE (*Dictyostelium* repetitive element). From the analysis of more than 50 elements, it can be concluded that DRE elements always occur 50 ± 3 nucleotides upstream of tRNA genes. All analyzed clones contain DRE in a constant orientation relative to the tRNA gene, implying orientation specificity as well as position specificity. DRE contains two open reading frames which are flanked by nonidentical terminal repeats. Long terminal repeats (LTRs) are composed of three distinct modules, called A, B, and C. The tRNA gene-proximal LTR is characterized by one or multiple A modules followed by a single B module (A_nB). With respect to the distal LTR, two different subforms of DRE have been isolated. The majority of isolated clones contains a distal LTR composed of a B module followed by a C module (BC), whereas the distal LTR of the other subform contains a consecutive array of a B module, a C module, a slightly altered A module, another B module, and another C module (BC · ABC). Full-length as well as smaller transcripts from DRE elements have been detected, but in comparison with the high copy number in *D. discoideum* strains derived from the wild-type strain NC4, transcription is rather poor.

Genetic and molecular analyses of many eukaryotic cells uncovered different repetitive elements, all of which have in common a coding capacity for reverse transcriptase (RT). These retrotransposons are divided into families which contain or lack long terminal repeats (LTRs). LTR-lacking retrotransposons are collectively termed L1-like elements. These elements are found from mammals (e.g., the L1 element of *Mus domesticus* [20]) to invertebrate organisms (e.g., R2 elements in *Bombyx mori* [5] and I factor and F element in *Drosophila melanogaster* [14]). Most L1-like elements studied show a dispersed genomic organization, although some elements, e.g., Ingi and Slacs elements in *Trypanosoma brucei* (1, 28), show a preferred integration specificity. Rough estimation of the copy numbers revealed from 20 to up to 10^4 copies per cell. L1 elements show higher similarity to fungal group II introns than to retrotransposons of the viral group (34) and are frequently disrupted at their 5′ ends.

Examples of LTR-containing elements are found in many eukaryotic organisms. Members of the Ty1-copia family such as Ty1 (9) and Ty2 (47) in *Saccharomyces cerevisiae*, Tal in *Arabidopsis thaliana* (46), Tnt1 in tobacco (17), and copia in *D. melanogaster* (35) encode *gag* and *pol* proteins but not an *env*-like protein. Functional domains of *pol* gene products of this family are ordered as protease–integrase–RT–RNase H. Domains of *pol* gene products from gypsy group retrotransposons (e.g., Ty3 in *S. cerevisiae* [10, 18] and the *Drosophila* elements gypsy [23], 297 [40], 17.6 [22], and 412 [52]) are organized as protease–RT–RNase H–integrase.

In *Dictyostelium discoideum*, nearly all tRNA gene loci show a high frequency of restriction fragment length polymorphism (12, 13). To gain insight into this interesting phenomenon, we analyzed the genomic organization of previously isolated tRNA genes (32). In the course of this study, we repeatedly identified the retrotransposable element DRE as well as the previously characterized transposable Tdd-3 element (30, 38). Both elements were found associated with several unrelated tRNA genes in a characteristic position-specific manner.

Up to now, DRE has never been found other than in association with tRNA genes in *D. discoideum*. In all analyzed clones, tRNA genes are separated by 50 ± 3 nucleotides from the associated DRE element, and DRE always occurs in a constant orientation relative to tRNA genes. Since no sequence similarity is apparent at the DRE integration site, a mechanism other than sequence-specific integration must be responsible for this striking position-specific and orientation-specific integration of DRE elements in *D. discoideum*. Similar results with respect to position specificity but not with respect to orientation specificity were obtained for Ty3, a retrotransposon identified in *S. cerevisiae*. Ty3 elements or a remnant LTR, termed sigma, were also found in front of various tRNA genes, separated in this case by 16 to 19 nucleotides from the mature coding tDNA sequence (10, 41). Both Ty3 and DRE, therefore, are members of a family of retrotransposons which integrate by a position-specific rather than by sequence-specific or random mechanisms into the chromosome (8, 31). Sequence-specific recombination has been described for many retrotransposable elements in different organisms (1, 5, 15, 21, 22, 50), but many retrotransposons and most retroviruses are thought to integrate more or less at random (for reviews, see references 3, 42, and 45). In the cases of Ty3 and DRE, a great number and variety of tRNA genes mark the specific integration sites. This is a remarkable fact which puts these elements into a different category from other retrotransposons and

* Corresponding author.

retroviruses which can act as lethal mutagens. Studying these specialized retrotransposons may shed light on a new type of integration mechanism which apparently neither causes any harm to the host cell nor interferes with relevant genetic information, since in eukaryotes signals for efficient transcription of tRNA genes are located inside the structural gene.

MATERIALS AND METHODS

Strains and plasmid DNAs. All *D. discoideum* strains used in this study were obtained from D. Welker, Utah State University, Salt Lake City, and from Keith Williams, Macquarie University, Sydney, Australia. Wild-type strains (WS380B, WS472, V12, OHIO, NP20, and NC-4) or non-axenic laboratory strains (HU526, HU1628, HU1184, HUD311, and HU2274) were grown in association with *Klebsiella aerogenes* on SM agar plates at 21°C (49). Axenic strains AX-2 and AX-3 were grown in HL5 medium (48) at 21°C.

Recombinant clones carrying 5′ ends of different DRE elements and tRNA genes were isolated by screening a λZap library prepared from sheared and *Eco*RI-linker-ligated genomic DNA from *D. discoideum*, which was kindly provided by Herb Ennis (LaRoche Institute, Nutley, N.J.). Additionally, plasmid libraries were constructed by ligating *Hin*dIII- or *Bgl*II-restricted genomic DNA into pUC18 plasmids. Recombinant plasmids were transformed into the *Escherichia coli* SURE strain (Stratagene) and were screened for the presence of DRE by using LTR-specific probes.

To isolate a complete DRE element, a partial *D. discoideum Sau*3A plasmid library was screened. This screening was kindly done by Jeffrey Williams (Imperial Cancer Research Fund, London, United Kingdom).

Sequence analysis of genomic tDNA clones. The sequence of DRE was originally established from pB3. The plasmid was subcloned, and Bal31 deletions were performed to construct sets of overlapping DNA sequences. All subcloned DNAs were sequenced as double-stranded DNAs according to the dideoxy-chain termination method (43). On the basis of the deduced sequence, we designed a set of 36 different oligonucleotides that is sufficient to provide overlapping sequence data of newly isolated DRE elements. Oligonucleotides relevant to this report are Rep1 (5′-GAGATTGTTGT TGTTGTTTGTT-3′ [34 to 12, 233 to 211, 5702 to 5680, and 5901 to 5879]), Rep7 (5′-GATGGCTTTTTAAAATAG-3′ [116 to 99 and 5784 to 5767]), Rep9 (5′-TCTTCGACT TCACCAGCCC-3′ [1280 to 1298]), Rep15 (5′-CCTCTTT TAGAGACCCTG-3′ [5551 to 5568 and 6315 to 6332]), Rep17 (5′-TTAAATTGGCTACTAGTGC-3′ [4839 to 4821]), Rep18 (5′-CAATACTATTGTGTTGTGG-3′ [4116 to 4098]), Rep19 (5′-GATGATTTGTGACACCC-3′ [1936 to 1952], Rep22 (5′-TTTCTGGATTAACTTGGGTTC-3′ [335 to 316, 5243 to 5224, and 6003 to 5984]), and Rep37 (5′-GCATCTCAAGGA GAAGC-3′ [5606 to 5622 and 6375 to 6391]). Numbers correspond to sequence coordinates.

Southern and dot blot analysis. Southern and dot blot analyses were carried out as described elsewhere (31). Genomic analyses (mapping of the structure of the DRE element in the background of different *D. discoideum* strains and rough estimation of the DRE frequency in *D. discoideum* cells) were done by digesting 3.0 μg of genomic DNA with the restriction endonucleases *Bam*HI, *Bgl*II, *Cla*I, *Eco*RI, *Eco*R5, *Hin*dIII, *Sau*3A, and *Sca*I. Genomic DNA fragments were fractionated on agarose gels and, after Southern transfer to GeneScreen Plus membranes, hybridized with suitable radiolabeled oligonucleotides.

Northern (RNA) blot analysis. RNA was isolated from vegetatively growing cells and cells starved in phosphate buffer under various conditions. After separation on agarose gels containing formaldehyde and after transfer to a nylon membrane, RNAs were probed with nick-translated DNA fragments as well as labeled, strand-specific oligonucleotides.

Nomenclature. The entire DRE element (A_nB–BC · ABC form) is approximately 6.4 kb in length. Nucleotides were numbered from left to right, the left side being the tRNA gene-proximal end. Nucleotides 1 to 199 constitute the core unit of the tRNA gene-proximal A module. This core unit occurs frequently in multiple, tandemly organized copies or may even be missing (31). In case a particular nucleotide within consecutive A modules of an individual DRE element needs to be identified, the first 199 nucleotides will be indicated by a small number (1 to 199_n). The two known subforms of A modules (31) are termed A_a and A_b. The two subforms are of identical length but are distinguished by three point mutations at positions 7, 103, and 178 (see Fig. 2). Because of the missing nucleotide at position 7 in the A_b subform, the direct terminal repeats of the A module comprise here 71 nucleotides, as opposed to the A_a subform, in which these repeats comprise 72 nucleotides.

Nucleotide +1 of tRNA genes is conventionally the first nucleotide of the mature tRNA coding region.

Nucleotide sequence accession numbers. The sequence data reported have been assigned GenBank/EMBL accession numbers M24053, M24058, M24060, M24064, M24565, M24567, M24569-71, X00325, X01763, X03499, X57034, X59561, X59563, X59565-66, X59569, X59571-73, X59576-81, X59583-84, X59587, and X59783.

RESULTS

Isolation of clones containing DRE sequences. To isolate a complete DRE element, a plasmid library containing partially digested genomic *Sau*3A fragments from the axenic *D. discoideum* strain AX-3 was screened with Rep1, an oligonucleotide which binds to both LTR structures (this library was kindly provided and screened by Jeffrey Williams). One clone obtained (pB3) carries 6.0 kb of genomic *D. discoideum* DNA, characterized by a modular structure shown in Fig. 1, group 1; it starts with the 72-bp repeat of the A module, followed by an oligo$(dA)_{39}$ track and a B module, a 5.4-kb coding region preceding a second B module, a C module, and three A-module core units. To confirm this structure, genomic clones from different libraries were selected and analyzed by restriction mapping, oligonucleotide hybridization, and sequence analysis. Finally, all isolated clones were grouped into five different categories (Fig. 1), all representing fractions of a 6.4-kb compound consensus DRE element. The entire sequence of this element is depicted in Fig. 2.

Group 1 contains the unique clone pB3, which according to the consensus DRE element contains neither an intact 5′ end nor an intact 3′ end. If libraries were screened with an A-module-specific oligonucleotide (Rep1) which according to the pB3 structures should bind to both LTR structures, surprisingly the vast majority of clones isolated represented left ends of DRE elements. All of these plasmids (a total of 56) were classified as group 2 clones, and they all carry at least one tRNA gene (determined by dot blot hybridization with end-labeled, bulk tRNA from *D. discoideum*). On

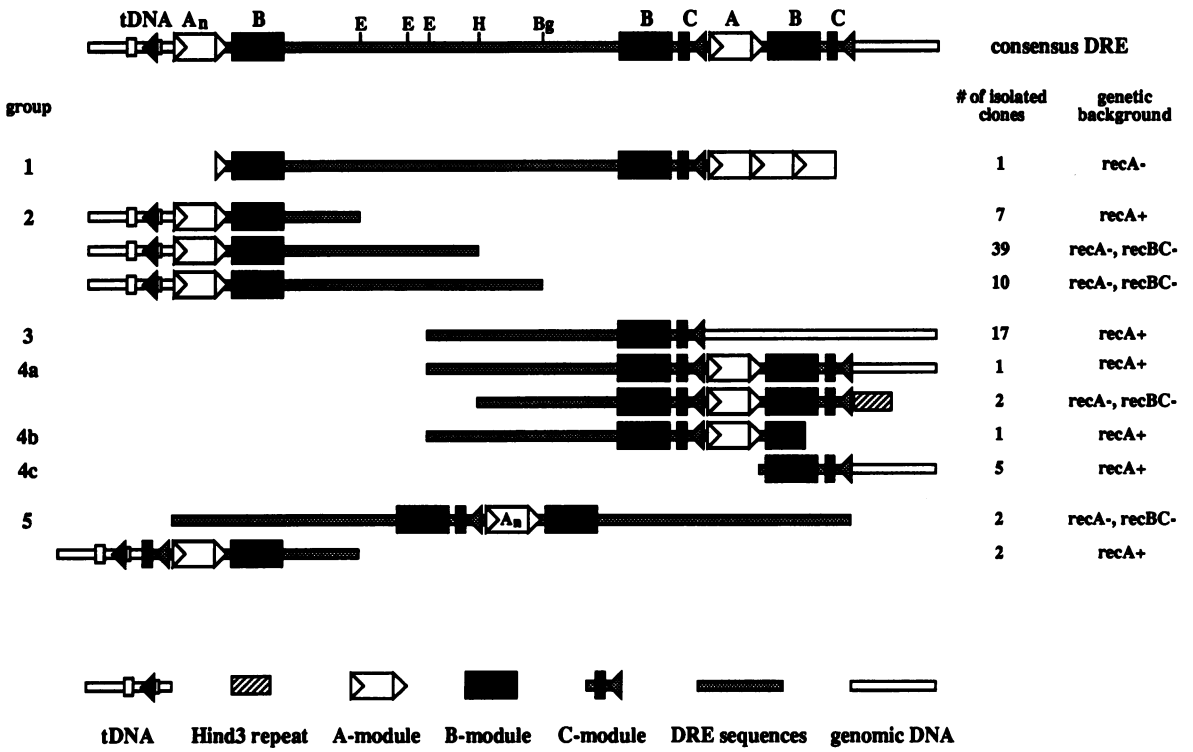

FIG. 1. Schematic organization of a 6.4-kb consensus DRE element whose structure has been derived from cloned *D. discoideum* DNA fragments containing various parts of individual DRE elements. These clones are subdivided into five major groups (see text). The single clone of group 1 (pB3) is the only clone which contains the entire coding region of DRE elements. This clone, however, contains an aberrant LTR organization. From these analyses and also from genomic mapping experiments, it can be concluded that DRE elements contain modularly composed LTRs which are differently organized at both ends. The tRNA gene-proximal LTR has an A_nB–structure, whereas the tRNA gene-distal LTR is characterized by a –BC or –BC · ABC structure. E, H, and Bg correspond to recognition sites for *Eco*RI, *Hin*dIII, and *Bgl*II, respectively. The column marked "genetic background" indicates the phenotype of the *E. coli* host in which plasmids have been amplified. Symbols for the various DNA parts are indicated.

characterized clones, the DRE element always resides 50 ± 3 nucleotides upstream of a tRNA gene (Fig. 3), and in all cases, transcription orientations of the tRNA gene and of the DRE element are divergent. These results confirmed the suggestion (31) that DRE elements occur associated with tRNA genes due to position- and orientation-specific integration.

Right-end LTR-containing clones had to be selected for with oligonucleotides, such as Rep15, that bind within the C-module region. Using this strategy, we isolated a total of 26 clones which were grouped into four categories (groups 3, 4a, 4b, and 4c; Fig. 1). The majority of these clones (group 3) contain, in addition to a part of the coding sequence, only a terminal –BC structure, thus representing DRE elements of only 5.7 kb.

Only three clones were isolated which, like pB3, carried a C module followed by an A module at the 3′ end. Further characterization of these clones revealed a terminal –BC · ABC structure (group 4a). These plasmids are very unstable in *E. coli*. If they are transformed back into *E. coli* (either *recA*$^+$ or *recA* mutant strains), a great fraction of the recovered plasmid DNA has lost the terminal ABC structure (see below). One clone (E.15–17) is linked to regular genomic DNA, while the modular terminal BC · ABC structure of the other two clones (H.37-1 and H.64-1) is linked to another repetitive *Dictyostelium* DNA sequence, the *Hin*dIII repeat (H3R). This element is 261 nucleotides long and has a *Hin*dIII restriction site on one end. It has been identified in either orientation in front of different tRNA genes (Fig. 3). Clones H.37-1 and H.64-1 can be distinguished from each other by a single point mutation within H3R. The one 4b clone and the five 4c clones were derived from a library containing sheared genomic DNA, which explains the truncations within the right-end LTR.

Clones grouped in class 5 of Fig. 1 contain linked LTRs which are characterized by a –BC · A_nB– structure. Together with pB3 (group 1), such clones might be explained by the presence of circular DNA intermediates of replicated transcripts (37) or by tandemly repeated DRE elements in the genome of *D. discoideum*.

Complete DRE elements are composed of nonsymmetric LTR structures encompassing an internal coding part. Detailed sequence information (Fig. 2) derived from all isolated DRE subclones confirmed the modular structure of the elements. According to this information, DRE contains nonsymmetric LTRs which flank an extensive internal coding part. The structure of a consensus DRE element can be described as A_nB–BC · ABC in the case of the 6.4-kb version or as A_nB–BC in the case of the 5.7-kb version, where A, B, and C represent the modular structures of the LTRs, n indicates the possible multiple occurrence of A-module core units within the left LTR, and the line represents the central coding region.

The A module contains a core unit of 199 nucleotides plus a direct repetition of nucleotides 1 to 72. The A-module core unit may be missing (in the cases of ValAAC-8 and

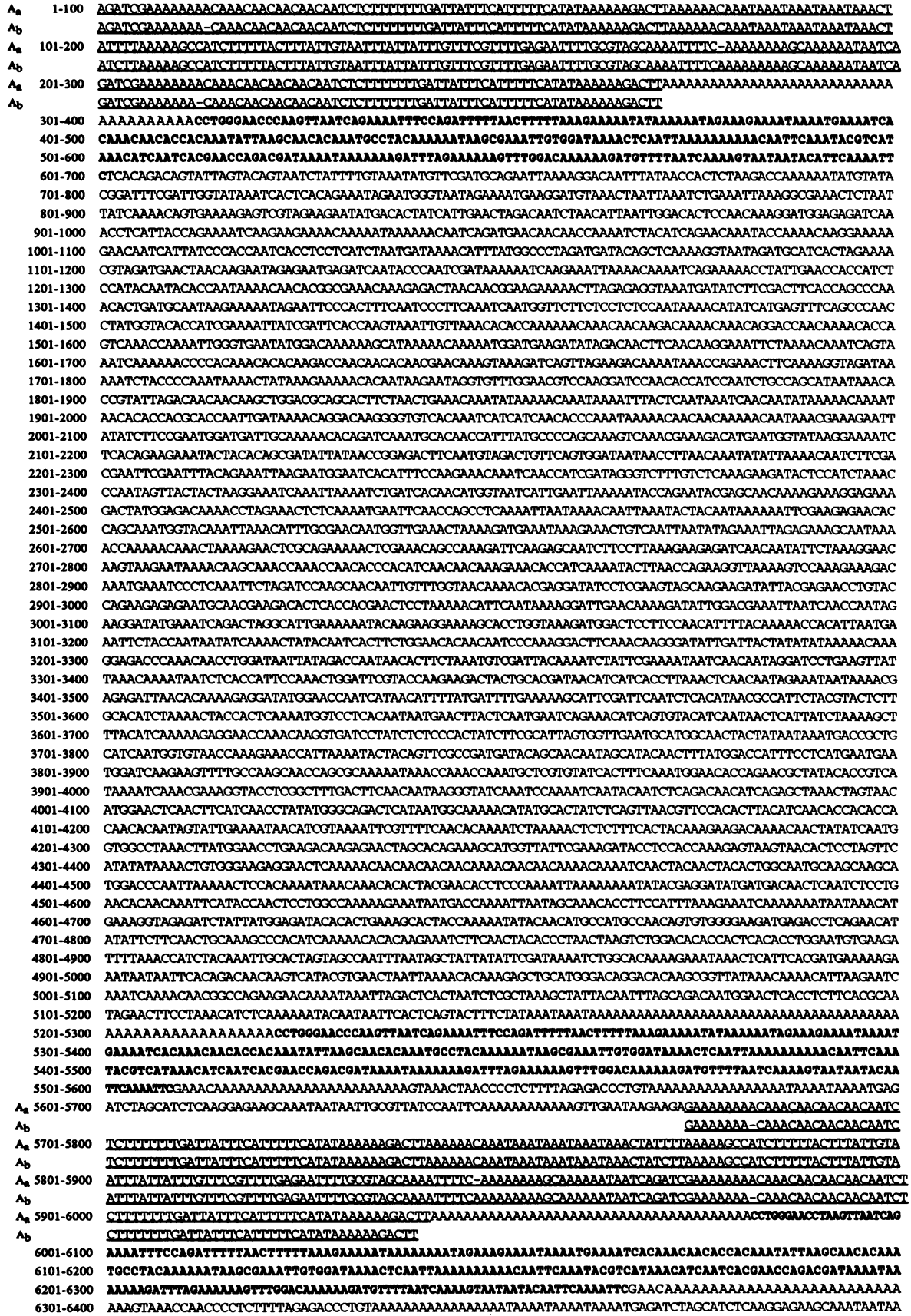

FIG. 2. Complete sequence of a 6.4-kb consensus DRE element. A modules (underlined nucleotides) are located from coordinates 1 to 271 and 5674 to 5939. Both subforms of A modules (A_a and A_b) are shown. B modules (boldface nucleotides) are located from coordinates 311 to 601, 5219 to 5510, and 5979 to 6270. C modules are defined by nucleotides 5511 to 5773 and by nucleotides 6271 to 6428. ORF1 starts with nucleotide 379 and ends with a UAA stop codon at position 1723. ORF2 spans from nucleotides 1701 to 5155.

LysUUU-3) or can occur as repeats within the 5′ LTR (Fig. 3) but has always been found once within an intact 3′ LTR. Characteristic for the A module of the 3′ LTR is the absence of the first five nucleotides (5′-AGATC-3′). Two subforms of the A-module core unit (A_a and A_b) have been identified. The two subforms are of identical length but are distinguished by three point mutations at positions 7, 103, and 178 (Fig. 2). Because of the missing nucleotide at position 7 in the A_b subform, the direct terminal repeat of the A module here comprises 71 nucleotides, as opposed to the A_a subform, in which this repeat comprises 72 nucleotides. Remarkably, certain A-module subtypes seem to be constantly associated with certain tRNA gene families, even in different *D. discoideum* strains (Fig. 3). Also in clone pB3, 5′ and 3′ A modules subforms are of the same type.

The B module comprises 290 nucleotides and provides the translation initiation site of the first open reading frame (ORF1). The first 34 nucleotides of the B module (nucleotides 311 to 344) show a strong homology to heat shock promoter elements identified in yeast and *Drosophila* cells (44). The B module is present once within the 5′ LTR and once or twice in the 3′ LTR of 5.7- and 6.4-kb DRE elements, respectively.

The C module has been defined as the sequence which terminates DRE elements. In the case of 6.4-kb DRE elements, the 14 nucleotides which generally separate the

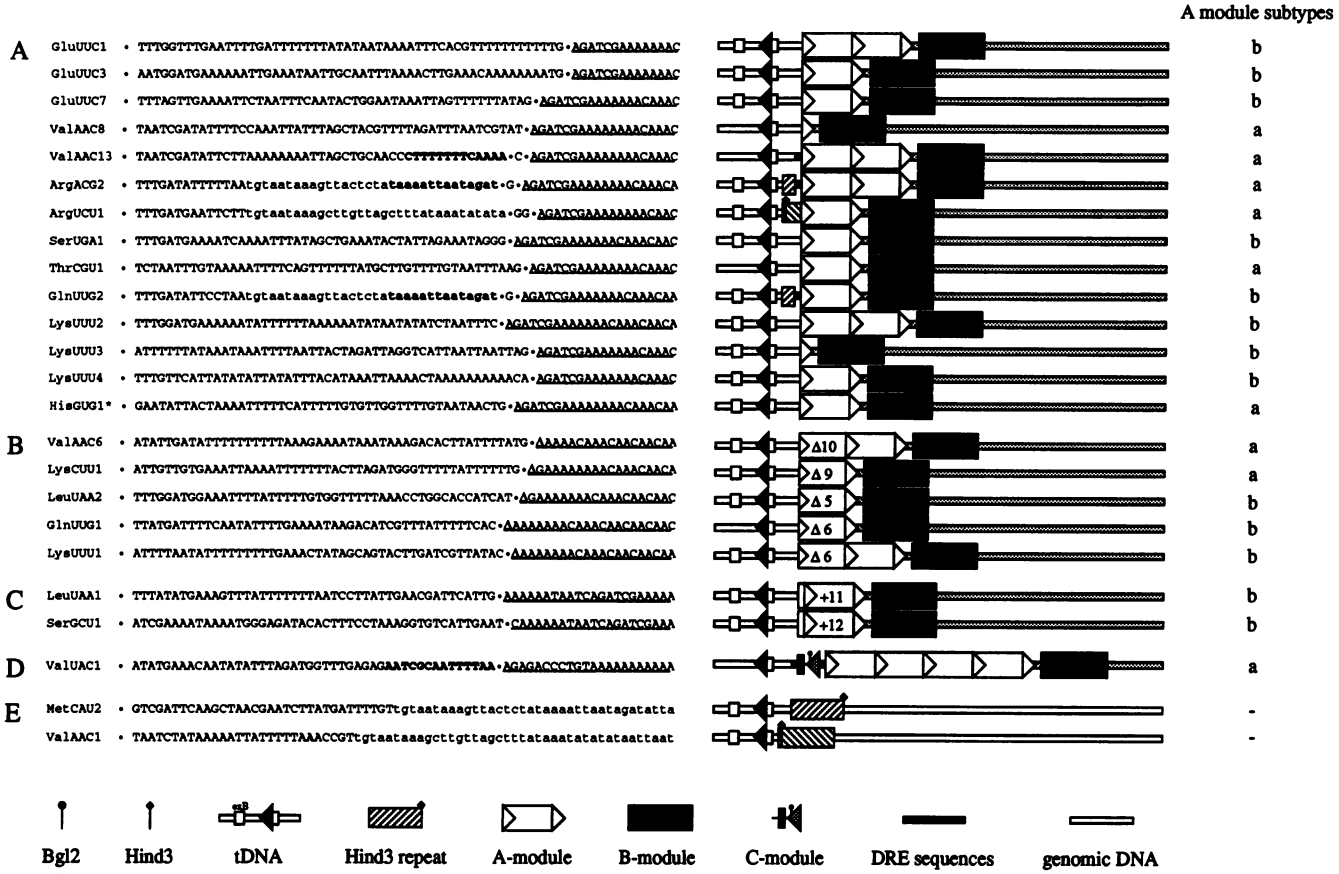

FIG. 3. Association of DRE elements with different *D. discoideum* tRNA genes. On the left are shown nucleotides which separate the tRNA gene (transcribed from right to left) and the DRE element (transcribed from left to right). Underlined nucleotides are part of DRE elements; nucleotides in lowercase are derived from another repetitive element, termed H3R. This element is 261 nucleotides long and has a *Hin*dIII restriction site (marked by a black rhombus in the schematic drawing at the right) on one end. It has been identified in either orientation relative to tRNA genes. From circumstantial evidence, it can be suggested that G nucleotides (and a C nucleotide in the case of ValAAC-13) which are flanked by dots are generated during the transposition event through an unknown mechanism (see text). Nucleotides in boldface have been identified as target site duplications. On the right side, the organization of regions around the tRNA genes (black arrowheads pointing to the left) is shown schematically. (A) All elements that inserted correctly in front of tRNA genes, indicated by the fact that the first A module of DRE starts with nucleotide 1 (AGATCGA$_{7/8}$. . .). (B to D) Misintegrated DRE forms which carry ΔDRE::tRNA gene associates (B) with small deletions (5 to 10 nucleotides) of the left-terminal A_n module, DRE(+)::tRNA gene associates (C) with a few additional A-module nucleotides (11 to 12 nucleotides) in front of an intact A_n module, and a putative solo C-module integration into a preexisting DRE::tRNA gene associate (D). (E) Clones that are associated only with H3R.

ORF1

```
1-50    KENKMKITNN TTNIKQHKCL QKISEIVDKT QLKKKQFKYV ININHEPDDK
51-100  IKKDLEKSLD KKDVLIKSNN TFKILTDSIS TVIYFVNMFD AELKGQFITT
101-151 LRPKNMYTDF DWYKSLTEIE WVIENEGCKL IKSEIKGETL IIKTVKRVVE
151-200 EYDTIIELDN LTLIGHSNKG WRDQTSLPEN QEENKNKKQS DEQQPKSTSE
201-250 QIPKQGKRTI IIPPITSSSN DKTFMALDDT AQKVIDASLE NVDELTRIEN
251-300 EINTQSIKNQ EIKTKSEKPI EPPSPYNTPI KQHGETKRLT TEEKLREVND
301-350 IFDFTSPNTD AIRKIEFPLS IPSNQWFFSS PIKHIMSFSP TYGTPSKIID
351-400 SPSKLLNTPK NKQQDKTNRT NKTPVKPKLG EYGQKSIKTK MDEDIDNFNK
401-448 EILKQISKSK NPTNTQDQQH NEQSKDQLED KINQKLQKVD KNLPQIKL*
```

ORF2

```
1-50    KSTPNKTIKK NTIRIGVWNV QGSNTIQSAS IINTVLDNNK LDAALLTETN
51-100  IKTNKIYSIN QQYKNKITHH APIDKTGQGV SQIIINTQIK TTTKTINERI
101-150 ISSEWMIAKT QIKCTTIYAP AKSNERHEWY KENLTEEILH SDIITGDFNV
151-200 DCSVDNNLNK YIKTIFDEFE FTEIKNGITF PRNKSTIDRV FVSKKILHLN
201-250 PIVTTKEIKL KSDHNMVIIE LKIPEYEQQK KGERLWRQNL ETLKMNSTSL
251-300 KINKTIKYYN KKFEENTSKW YKLNICEQWL KLKDEIKKLS INIEIRESNK
301-350 TKNKLKELAE KLETAKDSRA IFLKEEINNI LKEQVRIKQA NQTNTHINNK
351-400 ETPSKYLTRR LKVQRKTNEI PQILDPSNNC LVTKHEDILE VARRYYENLY
401-450 QKRECNEDTH HELLKTFNKR IEQKILDEIN QPIEGYEIRL GIEKIQEGKA
451-500 PGKDGLLPTF YKNHINEILP IISKLYNHFW NTTIPKDFKQ GILITIYKNK
501-550 GDPNNLDNYR PITLLNVDYK IYSKIINNRI LKLLNKIISP FQTGFVPRRL
551-600 LHDNIITLNS TIEIIKREIN TKEDMEPIIT FTDFEKAFDS ISHNAILRTL
601-650 AHLKLPLKMV LTIMNLLNES ETSVYINNSL SKSFTSKRGT KQGDPISPTI
651-700 FALVVECMAT TIINDRCING VTKETIKILQ FADDTATIAY NFMDHFLMNE
701-750 WIKKFCQATG AKINQTKCSC ITFKWNTRTL YTVIKSNERY LGFDFNNKGI
751-800 KSKINTISDN IRAKLVTWNS TSSTYMGRLI MAKTYALSQL TFHTYINTTP
801-850 QHNSIENNIV KFVFNTKSKN SLSLQRRQNN YINGGLNLWN LKTRELAQKA
851-900 WLFERYLHQR VSNTPSSYIK LWEEELKNNN NNKTTTKQNQ LQLEWQCKQA
901-950 WTQLKTPQNK QTEYEELPKL KKIYEDMMTT QSPEHNKFIP TPGQKEIMTK
951-1000 INSKHLPFKE IKKIINMKGR DLLWRYTLKA LPKIYNMPCQ QCGEDETSEH
1001-1050 IFFNCKAEIK NTQEIFNYTL TKSGETTETW NVKILNELQI ALVANLIAII
1051-1100 FDKIWHKRNK LIHDEKEIII HRQQVIRELI KTQRAAWDRT QAVINKTLRI
1101-1150 KSKQRPEEQN KLDSLISLKL LQFSRQWNSP LHAIELPKHL KKYNNSLSTF
1151-1152 YK*
```

FIG. 4. Amino acid sequence of the two ORFs encoded by the internal part of DRE. ORF1 resides between nucleotide positions 379 and 1725, and ORF2 lies between nucleotide positions 1701 and 5159. Starting with the first ATG codon, ORF1 codes for 443 amino acids. Lysine amino acids are underlined; proline clusters are in boldface and underlined. They are typically found in some retroviral *gag* proteins. ORF2 encodes 1,152 amino acids and overlaps ORF1 by 22 nucleotides. Domains which share similarities with LINE RT (amino acid positions 446 to 745) and with cystine/histidine motifs (amino acid positions 894 to 1037) are indicated.

internal C module from the following A module are idiosyncratic for an individual DRE element. The terminal C modules of 6.4-kb DRE elements characteristically differ from C modules of 5.7-kb DRE elements and from internal C modules of 6.4-kb DRE elements by five nucleotides: the sequence motif TAAAA occurs twice within the C module of –BC structures (coordinates 5587 to 5596) but three times within the terminal C module of –BC · ABC structures (coordinates 6351 to 6365). On this basis, C modules of group 4c clones were identified as terminal C modules of 6.4-kb DRE elements.

Internal structure and ORF analysis. The internal structure codes for two long ORFs (ORF1 [nucleotide positions 379 to 1725] and ORF2 [nucleotide positions 1701 to 5159]) (Fig. 4). Starting from a putative ATG initiation codon at amino acid position 5 in ORF1, a protein sequence of 443 amino acids can be deduced which terminates with a typical *Dictyostelium* UAA translational stop signal. This putative peptide has a molecular mass of 51.8 kDa and is a slightly basic (approximate charge of +8 at pH 7.0). Comparison of the amino acid sequence of ORF1 with known protein sequences uncovered no apparent similarities with the exception of some proline clusters and many basic amino acids (especially K) which are also found in some retroviral *gag* proteins.

ORF2 encodes 1,152 amino acids and also terminates with UAA. This stop codon resides within the sequence 5′-TATAAATAAATAAAA$_{55}$-3′, which may also function as a polyadenylation signal. The deduced polypeptide has a molecular mass of approximately 135 kDa. There is no methionine start codon within the first 105 amino acids of ORF2. Since ORF1 and ORF2 overlap by only 22 nucleotides, it might be suggested that both are translated into a single polyprotein involving either splicing, ribosomal hopping, or a −1 frameshift event. Colinear translation of two separated overlapping reading frames involving nonsense or frameshift suppression has already been documented for several other retrotransposons and retroviruses as well (2, 7, 9, 19, 24, 25, 33).

ORF2 contains a domain (amino acid positions 446 to 745) which shares significant similarity with RT, and three cystine motifs can be defined (amino acid positions 894 to 1037) which may constitute metal binding domains functionally involved in RNA or DNA binding. The putative RT domain of DRE (ORF2$_{446–745}$) shares 26 identical and 4 conserved-group amino acids of the 33 highly conserved amino acids among RT homologs from a wide variety of biological sources. A recent evolutionary analysis of 82 RTs and 15 RNA polymerases (51) suggests that with respect to the RT domain, DRE exhibits stronger homology to non-LTR-containing retrotransposons (jockey group) than to LTR-containing retrotransposons (gypsy group and Ty1-copia group).

Verification of DRE in the genome of *D. discoideum* AX-3. To confirm the DRE structure, *D. discoideum* genomic DNA was digested with various enzymes which cut DRE elements internally, and fragments were probed with selected oligonucleotides. In Fig. 5, the complete structure of the 6.4-kb DRE element (Fig. 5A) and representative Southern blot analyses (Fig. 5B) are shown.

Two types of signals were observed in lanes a to f of Fig. 5B: a strong signal of the predicted size (indicated next to the signal), apparently resulting from many identical fragments, and signals of lower intensity which might represent low-copy-number or single-copy fragments. The high-copy-number signals confirm that DRE is at least a middle repetitive element in AX-3 cells. Additionally, these results indicate that the vast majority of DRE elements are entirely intact at least up to the first C module of the 3′ LTR, exhibiting a structure which is schematically presented in Fig. 5A. Low-copy-number signals (Fig. 5B, lanes a to f) most likely reflect restriction fragment length polymorphism for one or both of the selected restriction sites or, less likely, disrupted elements.

Analysis of the 3′ end of DRE elements was more complicated (Fig. 5B, lane g). A *Bgl*II digest probed with the C-module-specific probe Rep15 was expected to show a signal in the range of 990 nucleotides as well as a 760-bp fragment if the proposed –BC · ABC structure at the 3′ end was correct. The 990-bp fragment was indeed found, representing the *Bgl*II fragment which carries a part of the second ORF and the internal BC module (see also lane f). However,

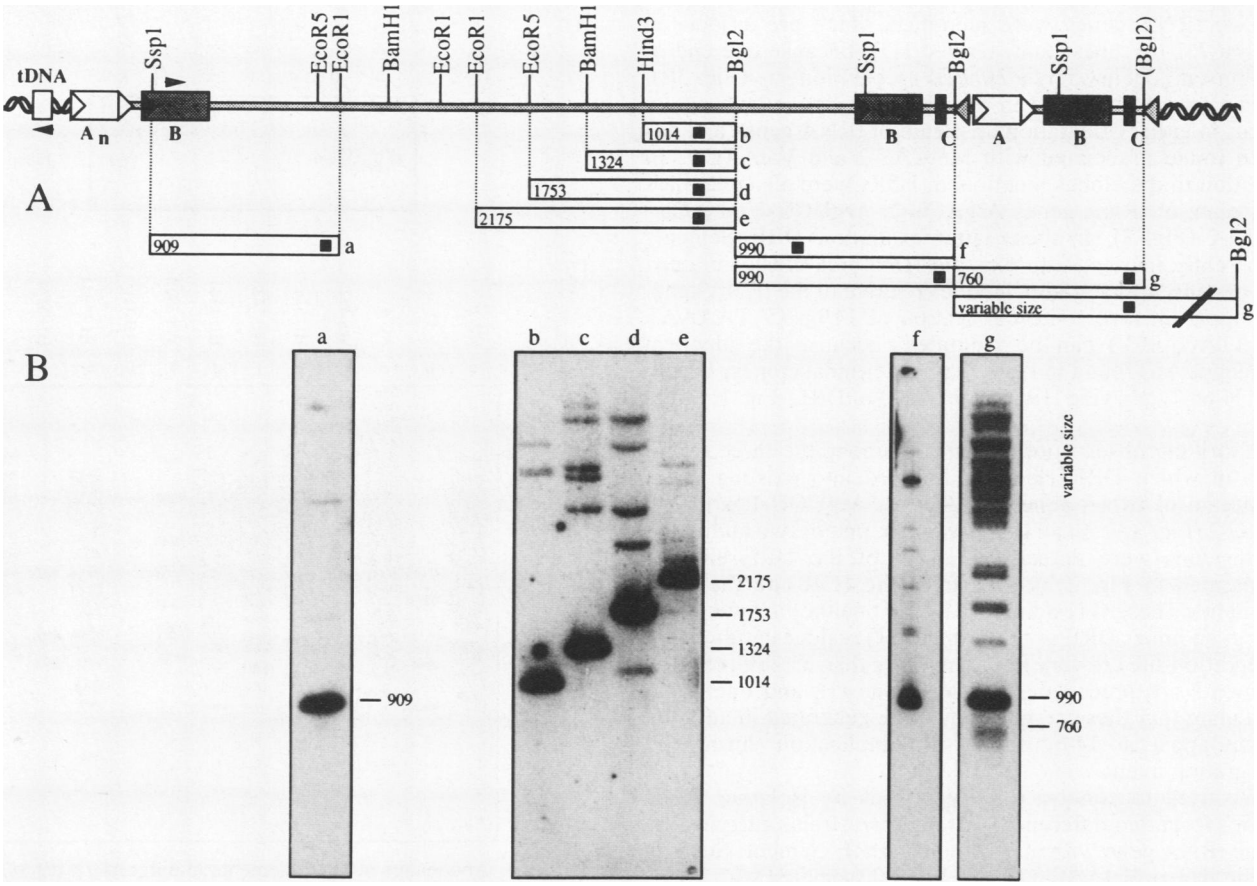

FIG. 5. Genomic mapping of DRE elements in AX-3. Genomic DNA from *D. discoideum* AX-3 was digested with defined sets of restriction enzymes which liberate DRE-internal restriction fragments (indicated by the width of the rectangles underneath the schematic drawing of a consensus DRE element). Fragments were size fractionated and examined by Southern blot analysis with end-labeled oligonucleotides Rep9 (a), Rep18 (b to e), Rep17 (f), and Rep15 (g). Positions of the oligonucleotides on the DRE element are indicated by black squares. Each lane contains at least one strongly hybridizing restriction fragment, indicating that most of the elements are entirely intact. The few low-intensity signals in each lane indicate elements which either are truncated or exhibit a restriction fragment length polymorphism for one of the restriction enzymes used. Lane g shows unexpectedly many low-intensity signals in addition to one strong signal of a 990-bp *Bgl*II fragment. This fragment is identical to that observed in lane f, since the C-module-specific oligonucleotide Rep15 recognizes the same *Bgl*II fragment as does Rep17. The presence of so many low-intensity signals instead of one expected strongly hybridizing fragment of 760 bp indicates that the terminal *Bgl*II site (indicated in parentheses on the schematic drawing of the consensus DRE element) is not present or is not accessible for the enzyme (see text).

the expected 760-nucleotide fragment can hardly be detected. Instead, a ladder of single-copy or low-copy-number signals was observed. As a reason for this result, one might assume that the *Bgl*II site within the terminal C module is highly polymorphic, although all eight isolated clones (group 4a and 4c) show no point mutation at this site.

Analysis of DRE::tRNA gene associates. One of the most striking features of DRE elements in *D. discoideum* is the fact that all elements analyzed so far are found exclusively 50 ± 3 nucleotides upstream of different tRNA genes, always organized in such a way that tRNA genes and elements are transcribed in opposite orientations. The yeast Ty3 element, which also integrates in front of tRNA genes, differs from DRE with respect to integration specificity in that it integrates 16 to 19 nucleotides upstream of tRNA genes in either orientation.

Among the 22 analyzed DRE::tRNA gene associates, only 14 DREs are integrated correctly, indicated by the fact that the first A module of DRE starts with nucleotide 1 (AG-ATCGA$_{7/8}$. . .) (Fig. 3A).

Misintegrated forms can be subdivided into three categories. (i) ΔDRE::tRNA gene associates (Fig. 3B) may reflect incompletely replicated DRE elements containing small deletions (5 to 10 nucleotides) of one of the left-terminal A-module core units (ValAAC-6, LysCUU-1, LeuUAA-2, GlnUUG-1, and LysUUU-1). (ii) DRE(+)::tRNA gene associates (Fig. 3C) contain, at the tRNA-proximal side, a few additional A-module nucleotides (10 to 12 nucleotides) in front of an intact A_n module (LeuUAA-1 and SerGCU-1). These variants may also be viewed as ΔDRE::tRNA gene associates that have lost an almost entire A-module core unit of the left-terminal A_nB structure. (iii) Misintegrated DRE::tRNA gene associates (Fig. 3D) contain C-module sequences (ValUAC-1) instead of an A module at the tRNA gene-proximal side. This clone, which was isolated twice from two different libraries, may be an example of the integration of a solo C module in front of a preexisting DRE element. The C-module sequence is flanked by 14 identical nucleotides (5′-AATCGCAATTTTAA-3′) which might reflect a target site duplication.

Two tRNA genes were identified which are associated with a 261-bp repeat unit, termed H3R because one end of the repeat is defined by a *Hin*dIII recognition site (Fig. 3E). This element (identified in Fig. 3 by lowercase letters) can occur in either orientation upstream of tRNA genes and has been found associated with MetCAT-2 and ValAAC-1. In addition to the clones mentioned, H3Rs were also identified upstream of tRNA genes ArgACG-2, ArgUCU-1, and Gln-UUG-2 (Fig. 3). In these latter examples, DRE elements apparently inserted into existing H3Rs but still maintained the distance of 50 ± 3 nucleotides relative to the tRNA gene. An apparent target site duplication of 14 bp (5′-TAAAAT TAATAGAT-3′) can be identified, because the identical sequence was found in the two right-terminal clones, H.37-1 and H.64-1, carrying HR3 at the 3′ end of DRE (Fig. 1, group 4a).

A striking observation was made among the three examples in which DRE elements inserted into existing H3Rs upstream of tRNA genes ArgACG-2, ArgUCU-1, and Gln UUG-2 (Fig. 3A). In all of these cases, one or two additional G residues were detected in front of the DRE element (separated in Fig. 3 by dots from the H3R and the DRE element). These G residues are present neither in intact H3R nor in an intact DRE at this position. G residues at the DRE integration site are very frequent (more than 65% of the DRE elements are preceded by at least one G), and one could speculate that these residues might be generated in addition to an apparent 14-bp target site duplication during the integration event.

We tried to confirm this hypothesis by isolating allelic clones from two different *D. discoideum* strains carrying the same tRNA gene with and without a DRE element. Such an allelic pair is ValAAC-13 from AX-2 (associated with DRE) and from the wild-type strain V12 (no DRE). An oligonucleotide derived from the sequence upstream of position −50 of the V12 allele served as probe for screening an AX-2 library. The isolated clone contains the 3′ LTR of a 5.7-kb DRE element flanked by the 13 nucleotides 5′-CTTTTTTTCA AAA-3′ which also flank the 5′ LTR (A_2B–) of ValAAC-13 from AX-2.

Frequency of DRE in *D. discoideum* strains. Genomic DNA of strain AX-2 was digested with the endonucleases *Eco*RI and *Cla*I to release an element-internal DNA fragment of 183 nucleotides from all DRE copies. Plasmid DNA containing a cloned part of DRE was used in equimolar amounts to 500, 250, 125, 67, 33, and 16 genome copies, respectively (Fig. 6A, lanes a to f). For AX-2, a copy number of 150 to 200 DRE elements was estimated from signal intensities on the autoradiogram (Fig. 6A, lane g). The axenic *D. discoideum* strain AX-2 is therefore a high-copy-number DRE strain.

A variety of other *D. discoideum* strains, laboratory strains as well as wild-type isolates, were also analyzed. Genomic DNA was digested with *Eco*RI and probed with Rep9, an oligonucleotide binding to the N-terminal coding region of DRE. According to this analysis, NC4 and its derivatives are high-copy-number DRE strains (Fig. 6B, lanes a, c, e, f, g, k, and m). Other wild-type strains (e.g., V12, WS472, WS380B, and OHIO) apparently contain many fewer intact DRE elements (Fig. 6B, lanes b, d, h, i, and l). These strains are called low-copy-number DRE′ strains.

Expression properties of DRE. Considering the high copy number of DRE elements in strain AX-3, expression is very low. In Northern blot experiments, RNA transcripts of distinct but various sizes were detected by hybridization either with synthetic oligonucleotides or with subcloned DNA fragments (Fig. 7). From these data, we conclude that

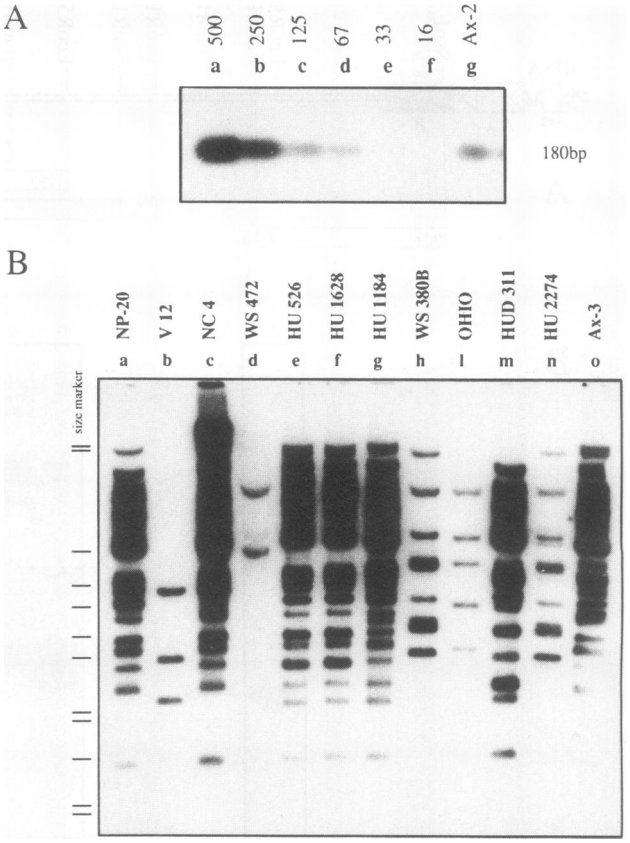

FIG. 6. Abundance of DRE elements in different *D. discoideum* isolates. (A) Genomic AX-2 DNA as well as 500, 250, 125, 67, 33 and 16 genome equivalents of pB3, as indicated, were completely digested with *Eco*RI and *Cla*I, size fractionated, and probed with Rep9. Approximately 150 to 200 copies of the liberated fragment (coordinates 1145 to 1324) can be estimated for the NC4 derivative AX-2. (B) Genomic DNA of several *D. discoideum* strains was digested with *Eco*RI and probed with Rep9. This primer binds to nucleotides 1280 to 1298 and should recognize only intact or nearly intact elements. Strains with many DRE elements (NP20, NC4, HU526, HU1628, HU1184, HUD311, and AX-3 [high-copy-number DRE strains]) can be distinguished from strains with significantly fewer DRE elements (V12, WS472, WS380B, OHIO, and HU2274 [low-copy-number DRE strains]).

DRE is expressed and that a minor fraction of the derived RNAs contains a putative primary transcript of about 5.7 kb.

RNAs corresponding to the 3.1-, 2.6-, 2.3-, and 2.1-kb signals cannot be identified with oligonucleotides complementary to the major part of ORF2 (nucleotide positions 2229 to 5054; data not shown), thus representing processed RNAs which have lost most of ORF2.

DISCUSSION

DRE represents a retrotransposable element in *D. discoideum*. In addition to DIRS-1 (6, 7) (also called Tdd-1 [39]), DRE constitutes a retrotransposable element in *D. discoideum*. Although direct proof for a retrotransposition event involving DRE is not available, a whole set of convincing circumstantial evidence suggests that DRE is a retrotransposon (31). The element is composed of a central coding part and LTRs which are differently organized at each end. This is not unusual for retrotransposons, since DIRS-1 (6) and

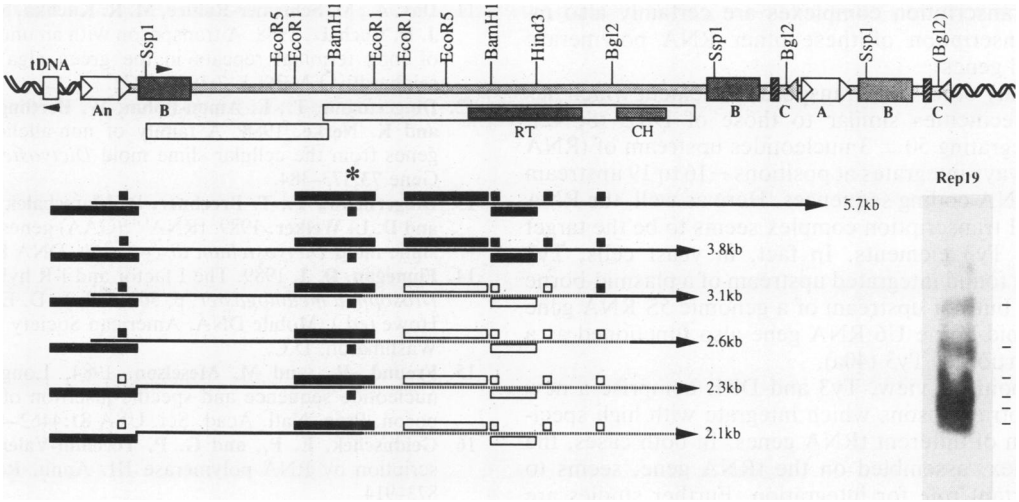

FIG. 7. Analysis of DRE transcripts. DRE transcripts were mapped by hybridization with strand-specific oligonucleotides (indicated by filled or open squares) or with labeled restriction fragments (indicated by filled or open rectangles). Filled symbols correspond to positive hybridization signals, whereas open symbols indicate probes which failed to hybridize. Below the schematic drawing of the 6.4-kb consensus DRE element is shown one Northern blot experiment involving hybridization with Rep19 (marked with an asterisk).

TOC.1 (11), a retrotransposon from *Chlamydomonas reinhardii*, also contain asymmetric LTRs. Two main subforms have been repeatedly isolated. The more extended 6.4-kb version contains a consensus structure which can be formally described as A_nB–BC · ABC, whereas a smaller, 5.7-kb version lacks the 3′-terminal ABC structure. Neither subform is constant in length, since the A module of the left LTR is frequently redundant and organized in tandem arrays. This feature is reminiscent of LINE elements. LINE similarity is also supported by lack of an apparent tRNA primer binding site and by a significant homology of parts of ORF2 to LINE RT.

The interior coding part comprises two ORFs which overlap by 22 nucleotides. Since there is no methionine start codon within the first 105 amino acids of ORF2, it is assumed that both ORF1 and ORF2 are translated into a single polyprotein. This might be achieved by a −1 frameshift event, by ribosomal hopping, or even as a consequence of splicing. Some form of translational read-through (19)—nonsense suppression (29), +1 frameshifting (2, 33), or −1 frameshifting (4, 24, 25)—occurs for most retroviruses or retrotransposons. Frameshifting could be achieved through a small stem-loop structure (1692-GGTAGATAAAAATC TACC-1709). However, we have no evidence yet for the functional significance of this structure.

We do not know the reason for the extremely low levels of DRE-derived mRNA, despite the fact that in NC4-derived strains such as AX-2 and AX-3, intact DRE elements occur at such high copy number. Whether transcription efficiency is generally very low or whether only a few DRE elements are transcriptionally competent has not yet been determined. The low level of DRE transcripts has made transcript characterization very difficult.

Translation of the DRE message is likely to start within the B module of the left LTR. The sequence of the B module also provides a 34-nucleotide-long putative heat shock response sequence as determined from sequence similarity with corresponding elements in *S. cerevisiae* and *D. melanogaster* (44). Functional evidence for this element has not yet been obtained, although much higher mRNA levels can be detected in cold-shocked *D. discoideum* cells (unpublished observation). This may, however, be due to enhanced stability of the mRNA under these conditions and may not necessarily reflect an elevated transcription rate.

The internal region codes for two extended ORFs formally of 448 and 1,152 amino acids. ORF2 shows strong similarity to known LINE-type RT domains (51), whereas ORF1 does not significantly match any known polypeptide. The only notable features of the putative peptide encoded by ORF1 are some proline clusters and several basic amino acids which are typical for retroviral *gag* proteins.

The sequence of DRE is remarkably constant, although isoforms have been detected. Thus far, three subtypes carrying mutations at distinct positions generally leading to conserved amino acid exchanges can be defined (31).

Integration model of DRE. The strict position- and orientation-specific association of DRE elements with tRNA genes remains remarkable. Because sequence-specific integration can be excluded, other properties in common among the targets must be responsible for this position specificity.

Eukaryotic tRNA genes are transcribed from stable complexes formed by two different transcription factors, TFIIIC and TFIIIB, and RNA polymerase III. Internal promoter elements (A and B blocks) are involved in TFIIIC binding (16). Stable binding of TFIIIC mediates stable TFIIIB binding to the 5′-flanking region of tRNA genes. Transcription is initiated after binding of RNA polymerase III to the tDNA-protein complex. This mechanism was most clearly demonstrated in yeast cells (26, 27). Evidence for a similar mechanism was also obtained for *D. discoideum* (unpublished results). The stable tRNA gene transcription complex might be responsible for a position-specific integration of DRE upstream of various tRNA genes. Assuming that TFIIIB in *D. discoideum* protects about 35 nucleotides of the 5′-flanking sequence (−15 to −50), as it does in yeast cells, the 50 ± 3-nucleotide-long spacing between the integrated DRE and a particular tRNA gene could be rationalized.

An association of DRE with other RNA polymerase III transcription units such as 5S RNA genes, U6 RNA genes, or 7SK RNA genes (36) has not been observed yet in *D. discoideum*. One could postulate such associations, since at least some transcription factors which are part of active

tRNA gene transcription complexes are certainly also involved in transcription of these other RNA polymerase III-transcribed genes.

Ty3 is the only other retrotransposable element which has integration specificities similar to those of DRE (8, 42). Instead of integrating 50 ± 3 nucleotides upstream of tRNA genes, Ty3 always integrates at positions −16 to 19 upstream of different tRNA-coding sequences. Here as well, the RNA polymerase III transcription complex seems to be the target for mobilized Ty3 elements. In fact, in yeast cells, Ty3 elements were found integrated upstream of a plasmid-borne 5S RNA gene but not upstream of a genomic 5S RNA gene (40a). A plasmid-borne U6 RNA gene also functioned as a target for insertion of Ty3 (40a).

From this point of view, Ty3 and DRE comprise a new family of retrotransposons which integrate with high specificity upstream of different tRNA genes. In both cases, the protein complex, assembled on the tRNA gene, seems to play an important role for integration. Further studies are required to confirm the notion that the RNA polymerase III transcription complex is not only sufficient but also necessary for DRE integration.

ACKNOWLEDGMENTS

We thank W. Kersten and H. Kersten for continuous support and encouragement. J. Williams and H. Ennis provided genomic *D. discoideum* libraries, and M. Bach provided excellent technical help. We are particularly grateful to S. B. Sandmeyer, who communicated results prior to publication, and to L. Schweizer, D. Söll, and S. B. Sandmeyer, who critically read the manuscript.

R. Marschalek is a recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. This work supported by grants from the Deutsche Forschungsgemeinschaft, Ria Freifrau von Fritsch Stiftung, and Fond der Chemischen Industrie to T.D.

REFERENCES

1. **Aksoy, S., S. Williams, S. Chang, and F. F. Richards.** 1990. SLACS retrotransposon from *Trypanosoma brucei gambiense* is similar to mammalian LINEs. Nucleic Acids Res. **18**:785–792.
2. **Belcourt, M. F., and P. J. Farabaugh.** 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNA induced slippage on a 7 nucleotide minimal site. Cell **62**:339–352.
3. **Boeke, J. D., and V. G. Corces.** 1989. Transcription and reverse transcription of retrotransposons. Annu. Rev. Microbiol. **43**: 403–434.
4. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell **57**:537–547.
5. **Burke, W. D., C. C. Calalang, and T. H. Eickbush.** 1987. The site-specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. Mol. Cell. Biol. **7**:2221–2230.
6. **Cappello, J., K. Handelsman, S. Cohen, and H. Lodish.** 1985. Structure and regulated transcription of DIRS-1: an apparent retrotransposon of *Dictyostelium discoideum*. Cold Spring Harbor Symp. Quant. Biol. **50**:759–767.
7. **Cappello, J., K. Handelsmann, and H. F. Lodish.** 1985. Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. Cell **43**:105–115.
8. **Chalker, D. L., and S. B. Sandmeyer.** 1990. Transfer RNA genes are genomic targets for *de novo* transposition of the yeast retrotransposon Ty3. Genetics **126**:837–850.
9. **Clare, J., and P. Farabaugh.** 1985. Nucleotide sequence of a yeast Ty element: evidence for an usual mechanism of gene expression. Proc. Natl. Acad. Sci. USA **82**:2829–2833.
10. **Clark, D. J., V. W. Bilanchone, L. J. Hywood, S. L. Dildine, and S. B. Sandmeyer.** 1988. A yeast sigma composite element, TY3, has properties of a retrotransposon. J. Biol. Chem. **263**:1413–1423.
11. **Day, A., M. Schirmer-Rahire, M. R. Kuchka, S. P. Mayfield, and J. D. Rochaix.** 1988. A transposon with an unusual arrangement of long terminal repeats in the green alga Chlamydomonas reinhardtii. EMBO J. **7**:1967–1972.
12. **Dingermann, T., E. Amon-Böhm, W. Bertling, R. Marschalek, and K. Nerke.** 1988. A family of non-allelic $tRNA^{Val}$(GUU) genes from the cellular slime mold *Dictyostelium discoideum*. Gene **73**:373–384.
13. **Dingermann, T., T. Brechner, R. Marschalek, E. Amon-Böhm, and D. L. Welker.** 1989. $tRNA^{Glu}$(GAA) genes from the cellular slime mold *Dictyostelium discoideum*. DNA **8**:193–204.
14. **Finnegan, D. J.** 1989. The I factor and I-R hybrid dysgenesis in *Drosophila melanogaster*, p. 503–518. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
15. **Freund, R., and M. Meselson.** 1984. Long terminal repeat nucleotide sequence and specific insertion of the gypsy transposon. Proc. Natl. Acad. Sci. USA **81**:4462–4464.
16. **Geiduschek, E. P., and G. P. Tocchini-Valentini.** 1988. Transcription by RNA polymerase III. Annu. Rev. Biochem. **57**: 873–914.
17. **Grandbastien, M. A., A. Spielmann, and M. Caboche.** 1989. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature (London) **337**:376–380.
18. **Hansen, L. J., D. L. Chalker, and S. B. Sandmeyer.** 1988. Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. Mol. Cell. Biol. **8**:5245–5256.
19. **Hatfield, D., and S. Oroszlan.** 1990. To where, what and how of ribosomal frame-shifting in retroviral protein synthesis. Trends Biochem. Sci. **15**:186–190.
20. **Hutchison, C. A., III, S. C. Hardies, D. D. Loeb, W. R. Shehee, and M. H. Edgell.** 1989. LINEs and related retrotransposons: long interspersed repeated sequences in the eucaryotic genome, p. 593–618. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
21. **Ikenaga, H., and K. Saigo.** 1982. Insertion of a movable genetic element, 297, into the TATA box for the H3 histone gene in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **79**:4143–4147.
22. **Inouye, S., S. Yuki, and K. Saigo.** 1984. Sequence specific insertion of the *Drosophila* transposable genetic element 17.6. Nature (London) **310**:332–333.
23. **Inouye, S., S. Yuki, and K. Saigo.** 1986. Complete nucleotide sequence and genome organization of a *Drosophila* transposable genetic element, 297. Eur. J. Biochem. **154**:417–425.
24. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell **55**:447–458.
25. **Jacks, T., and H. E. Varmus.** 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science **230**:1237–1242.
26. **Kassavetis, G. A., B. R. Brown, L. H. Nguyen, and E. P. Geiduschek.** 1990. *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. Cell **60**:235–245.
27. **Kassavetis, G. A., D. L. Riggs, R. Negri, L. H. Nguyen, and E. P. Geiduschek.** 1989. Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. Mol. Cell. Biol. **9**:2551–2566.
28. **Kimmel, B. E., O. K. ole-Moiyoi, and J. R. Young.** 1987. Ingi, a 5.2-kb dispersed sequence element from *Trypanosoma brucei* that carries half of a smaller mobile element at either end and has homology with mammalian LINEs. Mol. Cell. Biol. **7**:1465–1475.
29. **Kuchino, Y., H. Beier, N. Akita, and S. Nishimura.** 1987. Natural UAG suppressor glutamine tRNA is elevated in mouse cells infected with Moloney murine leukemia virus. Proc. Natl. Acad. Sci. USA **84**:2668–2672.
30. **Marschalek, R., G. Borschet, and T. Dingermann.** 1990. Genomic organization of the transposable element Tdd-3 from *Dictyostelium discoideum*. Nucleic Acids Res. **18**:5751–5757.
31. **Marschalek, R., T. Brechner, E. Amon-Böhm, and T. Dingermann.** 1989. Transfer RNA genes: landmarks for integration of

mobile genetic elements in *Dictyostelium discoideum*. Science **244**:1493–1496.
32. **Marschalek, R., and T. Dingermann.** Structure, organization and function of transfer RNA genes in the cellular slime mold *Dictyostelium discoideum*. *In* P. Greenaway (ed.), Advances in gene technology, in press. JAI Press Ltd., London.
33. **Mellor, J., S. M. Fulton, M. J. Dobson, W. Wilson, S. M. Kingsman, and A. J. Kingsman.** 1985. A retrovirus-like strategy for expression of a fusion protein encoded by the yeast transposon Ty1. Nature (London) **313**:243–246.
34. **Michel, F., and B. F. Lang.** 1985. Mitochondrial call II introns encode proteins related to the reverse transcriptase of retroviruses. Nature (London) **316**:641–643.
35. **Mount, S. M., and G. M. Rubin.** 1985. Complete nucleotide sequence of the *Drosophila* transposable element copia and retroviral proteins. Mol. Cell. Biol. **5**:1630–1638.
36. **Murphy, S., C. Di Liegro, and M. Melli.** 1987. The in vitro transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. Cell **51**:81–87.
37. **Panganiban, A. T., and H. M. Temin.** 1984. Circles with two LTRs are precursors to integrated retrovirus DNA. Cell **36**:673–679.
38. **Poole, S., and R. Firtel.** 1984. Genomic instability and mobile genetic elements in regions surrounding two discoidin I genes of *Dictyostelium discoideum*. Mol. Cell. Biol. **4**:671–680.
39. **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.
40. **Saigo, K., W. Kugimiya, Y. Matsuo, S. Inouye, K. Yoshioka, and S. Yuki.** 1984. Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*. Nature (London) **312**:659–661.
40a. **Sandmeyer, S. B.** Personal communication.
41. **Sandmeyer, S. B., V. W. Bilanchone, D. J. Clark, P. Morcos, G. F. Carle, and G. M. Brodeur.** 1988. Sigma elements are position-specific for many different yeast tRNA genes. Nucleic Acids Res. **16**:1499–1515.
42. **Sandmeyer, S. B., L. J. Hansen, and D. L. Chalker.** 1990. Integration specificity of retrotransposons and retroviruses. Annu. Rev. Genet. **24**:491–518.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Tuite, M. F., P. Bossier, and I. T. Fitch.** 1988. A highly conserved sequence in yeast heat shock gene promoters. Nucleic Acids Res. **16**:11845.
45. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 53–108. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
46. **Voytas, D. F., and F. M. Ausubel.** 1988. A copia-like transposable element family in Arabidopsis thaliana. Nature (London) **336**:242–244.
47. **Warmington, J. R., R. B. Waring, C. S. Newlon, K. J. Indge, and S. G. Oliver.** 1985. Nucleotide sequence characterization of Ty1-17, a class II transposon from yeast. Nucleic Acids Res. **13**:6679–6693.
48. **Watts, D. J., and J. M. Ashworth.** 1970. Growth of myxamoebae of the cellular slime mould *Dictyostelium discoideum* in axenic culture. Biochem. J. **119**:171–174.
49. **Welker, D. L., and K. L. Williams.** 1982. A genetic map of *Dictyostelium discoideum* based on mitotic recombination. Genetics **109**:691–710.
50. **Xiong, Y., and T. H. Eickbush.** 1988. The site-specific ribosomal DNA insertion element R1Bm belongs to a class of non-long-terminal-repeat retrotransposons. Mol. Cell. Biol. **8**:114–123.
51. **Xiong, Y., and T. H. Eickbush.** 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. **9**:3353–3362.
52. **Yuki, S., S. Inouye, S. Ishimaru, and K. Saigo.** 1986. Nucleotide sequence characterization of a *Drosophila* retrotransposon 412 element. Eur. J. Biochem. **158**:403–410.