Mol Gen Genet (1998) 257: 655–661 © Springer-Verlag 1998

ORIGINAL PAPER

T. Winckler · C. Tschepke · E. L. de Hostos
A. Jendretzke · T. Dingermann

Tdd-3, a tRNA gene-associated poly(A) retrotransposon from *Dictyostelium discoideum*

Received: 28 July 1997 / Accepted: 30 October 1997

Abstract The full-length 5218-bp sequence of the mobile genetic element Tdd-3 from *Dictyostelium discoideum* is described. Tdd-3 encodes two overlapping open reading frames (ORFs) flanked by non-redundant, untranslated regions. The deduced amino acid sequence of ORF2 is homologous to reverse transcriptases (RTs) encoded by the class of poly(A) retrotransposons. ORF2 also encodes a putative protein domain related to the family of apurinic/apyrimidinic (AP) endonucleases, whose retroelement-encoded homologs have recently been proposed to represent the integrase function of poly(A) retrotransposons. Comparison of several genomic Tdd-3 copies revealed that element insertion is orientation specific and occurs about 100 bp downstream of tRNA genes in the *D. discoideum* genome. These properties of Tdd-3 suggest that the element is a tRNA gene-associated poly(A) retroelement present in the *D. discoideum* genome. Analysis of several cloned cDNAs derived from Tdd-3-specific plus strand RNAs indicate that the element is transcribed and polyadenylated during the growth of *D. discoideum* cells.

Key words *Dictyostelium* · tRNA genes · Reverse transcription · Retrotransposon · Repetitive DNA

Communicated by D. J. Finnegan

T. Winckler (✉) · C. Tschepke · A. Jendretzke · T. Dingermann
Institut für Pharmazeutische Biologie,
Universität Frankfurt/M. (Biozentrum), Marie-Curie-Strasse 9,
D-60439 Frankfurt/M, Germany
Fax: +49-69-798-29662;
E-mail: Winckler@em.uni-frankfurt.de

E. L. de Hostos
Department of Biochemistry and Cell Biology, MS 140,
Rice University, 6100 Main St., Houston, TX 77005-1892, USA

Introduction

Retroelements are genetic entities originally derived by reverse transcription of RNA into DNA catalysed by enzymes that possess RNA-directed DNA polymerase activity. Retrotransposons inhabit a large variety of eukaryotic genomes and make up several percent of the host genome size (Voytas 1996). Retroelements have been subdivided into viral and nonviral superfamilies (Boeke and Stoye 1996). The group of retroviruses and retrovirus-like retrotransposons is defined by long terminal repeats (LTRs) located at both ends of the elements. All retroelements have in common an open reading frame (ORF) that encodes a reverse transcriptase (RT), the enzyme that converts RNA intermediates of the retroelements into cDNAs. Non-LTR retrotransposable elements are a large and diverse group lacking LTR structures. The poly(A) retroelements are a subgroup of the non-LTR elements that share structural homologies, e.g. they encode one or two open reading frames flanked by non-redundant untranslated regions (UTRs), and usually terminate with poly(A) or A-rich tails (Hutchison III et al. 1989).

In the retrotransposition mechanism of retrovirus-like retroelements the LTRs are important for both reverse transcription and integration of the elements (Boeke and Corces 1989; Boeke and Chapman 1991). Since poly(A) retroelements lack LTRs, they must use a replication mechanism significantly different from that of LTR retrotransposons. Within the past few years the retrotransposition mechanism of poly(A) retroelements has become clearer, thanks to the discovery of an element-encoded protein domain that acts as an endonuclease on the target DNA (Luan et al. 1993; Martín et al. 1995; Feng et al. 1996). The endonuclease functions to nick the target DNA and produces free 3′-OH ends, which can be used as primers for reverse transcription of the plus strand RNA (Luan et al. 1993; Feng et al. 1996). Thus, in the current model for transposition of poly(A) retroelements, plus

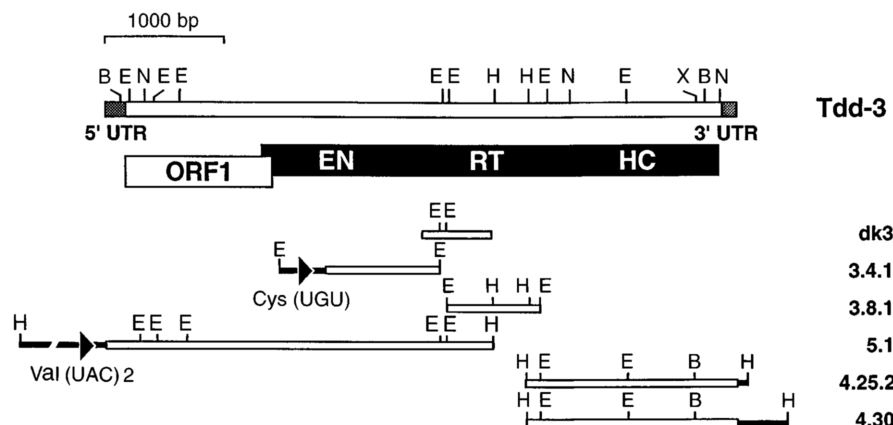

Fig. 1 Genomic organization of the Tdd-3 element. The physical maps of Tdd-3 and a collection of representative genomic clones used to determine the full-length sequence of Tdd-3 are shown. Element-encoding regions are shown as *open boxes*. Flanking genomic sequences are indicated by *thick black lines*. *Black triangles* represent tRNA genes with their transcription orientations. Note that only two (4.25.2 and 4.30) out of 10 individual clones carrying 3′ ends of Tdd-3 are shown. ORF2 is shown in *black* with the position of the endonuclease (EN), reverse transcriptase (RT), and histidine/cysteine-rich (HC) domains indicated. Abbreviations of restriction enzyme recognition sites are as follows: B, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; N, *Nde*I; X, *Xba*I

strand RNAs are reverse transcribed and integrated in a coupled reaction catalysed by an element-encoded multifunctional protein.

Transfer RNA genes of the eukaryotic microorganism *D. discoideum* have been studied in detail (Marschalek and Dingermann 1991). Like their counterparts in higher eukaryotes (Geiduschek and Tocchini-Valentini 1988; Willis 1993), *D. discoideum* tRNA genes contain gene-internal promoters termed "A box" and "B box". About 80% of *D. discoideum* tRNA genes analysed so far contain in their flanking regions one or two extra B box consensus motifs (exB boxes) located about 40 and 80 bp downstream of the tRNA gene coding regions (Hofmann et al. 1991; Marschalek and Dingermann 1991). During the analysis of transfer RNA genes in *D. discoideum*, mobile genetic elements were discovered that are associated with tRNA genes. The most remarkable of these is the "*Dictyostelium* repetitive element" DRE, a poly(A) retrotransposon which integrates in a highly position-specific manner 50 ± 4 bp upstream of tRNA genes (Marschalek et al. 1989, 1992; Hofmann et al. 1991). About 25% of the tRNA genes analysed are associated with DRE elements (Hofmann et al. 1991). The "transposable elements of *Dictyostelium*", Tdd-2 and Tdd-3, are repetitive DNA elements which were originally identified as a source of restriction length polymorphisms in the discoidin I gene locus (Poole and Firtel 1984). During the analysis of *D. discoideum* tRNA genes, Tdd-3 elements were identified which are inserted about 100 bp downstream of the coding regions of *D. discoideum* tRNA genes (Marschalek et al. 1990).

With the exception of DRE, whose full-length DNA sequence has been determined, only partial sequences are so far available for the other *D. discoideum* repetitive elements listed above, and little is known of the biology of these elements. Here we report the full-length sequence of Tdd-3 and show that this element shares structural characteristics with the class of poly(A) retrotransposons. Our data underline the importance of tRNA genes as landmarks for integration of mobilized retroelements. This has also been observed in other eukaryotic organisms, and possibly reflects a coevolution process involving retroelements and their host cells that tends to minimize the frequency of knock-out mutations caused by the activity of retroelements.

Materials and methods

Oligonucleotides

The following Tdd-3-specific primers were employed: CT-20 (5′-CCTGTTCTTCTCATGCAAG-3′); CT-24 (5′-GAAGTCTCCTG CTATGATG-3′); CT-30 (5′-CACATAAACGCAAGAACGAG G-3′) and CT-32 (5′-GGATTAACAAGATTGATATATGTC-3′).

Construction and screening of genomic and cDNA libraries

Genomic libraries were constructed by ligating 25 μg of genomic DNA from *D. discoideum* AX2 cells, digested with *Eco*RI or *Hin*d III, into the corresponding vector-encoded restriction sites of pUC19 or pGEM7Zf(–).

A cDNA library was constructed from RNA isolated from vegetatively growing *D. discoideum* AX2 cells. Some 500 μg of total RNA was purified as described (Schumann et al. 1994). Poly(A)$^+$ RNA was isolated from total RNA using the Qiagen Oligotex kit using the procedure described in the manual provided. Double-stranded cDNA was synthesized from 5 μg poly(A)$^+$ RNA primed with oligo(dT) and cloned into the *Eco*RI and *Xho*I sites of λHybriZAP using the HybriZAP^{TM} Two-Hybrid cDNA Gigapack cloning kit (Stratagene). The primary library, consisting of 6.5×10^5 pfu, was amplified in XL1Blue MRF′ cells (Stratagene) on 30 agar plates (9 cm diam.). The phages were eluted from each plate into 5 ml of SM buffer (Sambrook et al. 1989) and stored in 7% dimethylsulfoxide at −80° C.

PCR screening of the entire phage library was performed as follows. 100 μl aliquots of the 30 phage stocks (1.3×10^7 pfu/ml) were adjusted to 10 mM EDTA and incubated at 70° C for 15 min. Samples (5 μl) of phage lysate were subjected to PCR reactions

5' UTR

```
   1  CCGTACCGCGATCAAGAGGATACAAGATACACGTGAAAAGTAATTCATCCTTTTCTATCT
  61  TTAATCTCGGTCAATTTTAACCAATCTTTCAAAAAAAATCACCAATCCACCACGATTTAC
 121  AGATCTAACATACCAAATCGCCGATTCAAAGGAATCCCATCGAGATGAATTCCTACGCAG
 181  ACGCGGCTAAAATGGCCAACCCCTCCACCCATATGGACGAGGCATCGGAGAACTTCATTA
```

3' UTR

```
4998  TCAAAGATCCTAAGATCCTTCTCACCTCAACGATCAAGCTAAGAAAAACAACAGCTAATT
5058  ACTATTGTATTCCAGAATCATCTCTCCCAAATATTATATCTTTTGATCAATTCATATGAT
5118  TAAAATAACCCTTCAGCTTAAATGATAAGCCTTTAAATAAATATTAAATAAAACTCGTAT
5178  TAACACAGATGGACATATATCAATCTTGTTAATCCAATATTaaaaaaaaaaaaaaaaaa
```

Fig. 2 TR sequences of the Tdd-3 element. The +1 nucleotide of Tdd-3 is the same as that given by Marschalek et al. (1990). A complete Tdd-3 element is 5218 bp long; in addition, homopolymeric (dA) tracts of variable lengths are present at the 3′ ends of the elements (shown in *lower case letters*). In the 5′ UTR the 22-bp DNA motif that is homologous to Tdd-2 elements (Poole and Firtel 1984) is *underlined*. The putative ATG initiation codon for translation of ORF1 is *highlighted*. In the 3′ UTR in-frame translation stop codons are shown as *black boxes*, and putative AATAAA polyadenylation sites (Proudfoot 1991) are shown in *open boxes*

using Tdd-3-specific oligonucleotides as outlined in Fig. 4. Phage aliquots yielding PCR products were amplified and screened as described above.

Results

Genomic organization of Tdd-3 elements

An attempt to isolate cDNAs encoding *D. discoideum* kinesin by PCR using degenerate oligonucleotides yielded a 585-bp DNA fragment (clone dk3, Fig. 1) that encoded a single ORF. The deduced amino acid sequence of dk3 was not related to kinesin, but instead showed a high degree of homology to RTs of the subgroup of poly(A) retrotransposons, especially DRE from *D. discoideum* and human L1Hs (discussed below). The DNA sequence contained in clone dk3 was extended by cloning of overlapping genomic DNA fragments. The full-length 5218 bp DNA sequence of the Tdd-3 element was obtained with the cloning strategy depicted in Fig. 1 (GenBank accession number AF002669). Sequence information obtained from both ends of the cloned retroelement showed that stretches extending 126 bp from the 5′ end and 301 bp from the 3′ end were identical to the corresponding segments of available Tdd-3 sequences (Poole and Firtel 1984).

Several short DNA sequences from the 5′ ends of Tdd-3 elements were isolated as a result of their association with tRNA genes (Marschalek et al. 1990). The exact position of the +1 nucleotide of the Tdd-3 element has been determined by comparing several genomic clones encoding 5′ ends of Tdd-3 (Marschalek et al. 1990). The consensus start of Tdd-3 5'-CCGTACCG...-3′ (Fig. 2) was found in clone 5.1, suggesting that this DNA fragment derived from a genomic Tdd-3 copy with an intact 5′ end. Clone 5.1 carried a 5.2-kb DNA fragment in which the 5′ end of Tdd-3 was preceded by a $tRNA^{Val}$(UAC) gene in the 5′-flanking genomic region. Clone 3.4.1 was truncated at a point corresponding to position +2001 of the full-length Tdd-3 sequence and contained a $tRNA^{Cys}$(UGU) gene upstream of the Tdd-3 coding region. Two partial (5′-deleted) Tdd-3 sequences have deposited in the EMBL/GenBank that were identified downstream of a $tRNA^{Ile}$(AAU) and a $tRNA^{Lys}$(UUU) gene, respectively (Hofmann et al. 1991; Vithalani et al. 1996). Several other genomic clones have been shown to encode tRNA genes upstream of (partially 5′-truncated) Tdd-3 elements (Hofmann et al. 1991) (summarized in Table 1). Analysis of the distances of several Tdd-3 elements from their corresponding tRNA genes suggests that the elements inserted downstream of either the internal B box or an exB box motif. The average distance between Tdd-3 elements and the nearest B box equivalent is about 100 bp.

The location of the 3′ end of Tdd-3 was confirmed by sequence analysis of various independent clones encoding the 3′ end of the element and flanking genomic sequences (Fig. 1). All clones analysed showed identical Tdd-3-specific DNA sequences that ended with the consensus 5′-...CCAATATT-3′ (Fig. 2), followed by homopolymeric (dA) tracks of up to 70 bp in length and diverse 3′ flanking genomic sequences.

Experiments were performed in order to determine the genomic copy number of Tdd-3 elements. Southern blots that were hybridized with a 3′-specific probe showed about 20 distinct bands. Identical blots hybridized with a 5′-specific probe showed only about 10 bands, indicating that half of the Tdd-3 elements in *D. discoideum* AX2 lack intact 5′ ends (data not shown).

Open reading frames encoded by Tdd-3

Tdd-3 encodes two putative polypeptides in overlapping reading frames (Fig. 1). The 5′ UTR is unrelated to

Table 1 Summary of Tdd-3::tRNA gene associations detected in the *D. discoideum* genome

tRNA gene	Distance (bp) to		5′ deletion[a]	Reference
	B box	exB box		
Ile (AAU)1	111	56	+ 3945	Vithalani et al. (1996)
Cys (UGU)1	84	–	+ 2001	This study
Val (UAC)2	104	–		This study
Lys (UUU)2	124	71	+ 4492	Hofmann et al. (1991)
Val (UAC)1	106	25		Hofmann et al. (1991)
Glu (UUC)2	105	54	+ 2	Hofmann et al. (1991)
Glu (UUC)3	131	82	+ 4	Hofmann et al. (1991)
Val (AAC)1	89	41		Hofmann et al. (1991)
Tyr (GUA)1	115	67		Hofmann et al. (1991)
Ala (AGC)1	82	34		Hofmann et al. (1991)
Asn (GUU)1	149	97		Hofmann et al. (1991)
Arg (ACG)1	98	42	+ 4	Hofmann et al. (1991)
Met (CAU)1	116	66		Poole and Firtel (1984)

[a] Where no entry is given, the element has a complete 5′ end

DNA sequences of other retrotransposable elements, except for a 22-bp sequence that is nearly identical to the 5′ end of Tdd-2 (Poole and Firtel 1984) (Fig. 2). The first ATG codon is located at DNA position 165. However, according to the ribosome scanning theory (Kozak 1989) a downstream ATG codon may also serve as translation start. We assume that translation of ORF1 starts at DNA position 192, which fits well with the translation start sequence (AAA/ATG/G) frequently found in *D. discoideum* structural genes. Hence Tdd-3 ORF1 probably encodes 408 amino acids and does not share significant homology with any protein available in the EMBL/Genbank databases (Altschul et al. 1990). No consensus AATAAA polyadenylation signal (Proudfoot 1991) is present at the end of ORF1, suggesting that ORF1 may be translated from longer plus strand RNAs. ORF1 terminates with four adjacent TAA translation stop codons, providing a strong stop signal to avoid translational read-through and generation of nonsense fusion proteins (not shown).

ORF1 and ORF2 overlap by 75 bp. ORF2 is 3771 bp long and encodes a putative protein of approximately 147 kDa. ORF2 does not start with an ATG translation initiation codon. The first methionine is located at amino acid position 202 of the ORF2-encoded protein, suggesting that ORF2 may be translated as a multifunctional protein fused to ORF1 due to a −1 frame shift during protein translation. Within the ORF1/ORF2 overlap region two adjacent codons (5′-CCG/CAG-3′) are located in ORF1 that are used with very low frequency (<3%) in other *D. discoideum* genes. This may delay protein translation at the ribosome, such that a –1 ribosomal frame shift is favored. The putative RT domain of Tdd-3 element is situated in the central portion of ORF2. It shares significant homologies with various poly(A) retroelements, including DRE and human L1Hs (Fig. 3). Tdd-3 RT is also highly homologous to the corresponding domain of RED, another recently characterized poly(A) retroelement from *D. discoideum* (Fig. 3; T. Winckler, unpublished results).

The N-terminal one-third of ORF2 shows significant homology to the endonuclease (EN) domains of various poly(A) retroelements and to a *D. discoideum* AP endonuclease homolog (Fig. 3). Although the overall homology among EN domains of poly(A) retroelements and AP endonucleases is low (Feng et al. 1996), several stretches of sequence identity to human L1Hs and DRE from *D. discoideum* are present in Tdd-3. In particular, amino acids important for the DNA-nicking activity of L1Hs endonuclease (Feng et al. 1996) are conserved in Tdd-3.

The C-terminal one-third of ORF2 encodes a protein of unknown function, whose amino acid sequence was identified in the databases (Altschul et al. 1990) as being similar to a previously identified, but not further characterized, RED element from *D. discoideum* (Marschalek et al. 1990; Hitt et al. 1994). Figure 3 shows an amino acid sequence alignment of Tdd-3 with the previously reported RED-encoded ORF. RED shows high homology to Tdd-3 in a distinct domain that encodes a motif rich in histidine and cysteine residues (HC domain) with the consensus sequence $H\text{-}X_2\text{-}C\text{-}X_6\text{-}H\text{-}X_3\text{-}C\text{-}X_2\text{-}C\text{-}X_{10}\text{-}H\text{-}X_4\text{-}C$ (where X is any amino acid). The HC domain of Tdd-3 also shows low homology to the corresponding part of ORF2 from DRE (see Fig. 4). However, in DRE ORF2 the consensus HC motif is $C\text{-}X_2\text{-}C\text{-}X_7\text{-}H\text{-}X_4\text{-}C$.

Expression of Tdd-3 elements

RNA transcripts specific for Tdd-3 could not be detected on Northern blots of total RNA prepared from *D. discoideum* cells. However, the following observations suggest that plus strand RNA of Tdd-3 elements is produced and polyadenylated during vegetative growth of *D. discoideum* cells. Tdd-3 contains a 102-bp 3′ UTR preceding the poly(dA) tracts that terminate all elements. Two putative polyadenylation sites (AATAAA) (Proudfoot 1991) are present in the 3′ UTR about 50 bp upstream of the genomic poly(dA) ends of the elements (Fig. 2). Thus Tdd-3 copies present in the *D. discoideum*

EN domain

```
Tdd-3   (14) AHQLFPGSIISATNRGNG (12) LSPIFIIEGR (11) TTRIIAIYAP (13) NKHFNNQYHNLTS (10)
DRE     (05) KTIKKNTIRIGVWNVQGS (17) LDAALLTETN (27) QGVSQIIINT (23) QIKCTTIYAPAKS (17)
L1Hs         MTGSNSHITILTLNINGL (17) PSVCCIQETH (25) KAGVAILVSD (27) ELTIILNIYAPNTG (18)
DdApeA  (98) KNVEENQMKIISWNVAGF (16) PDVLCLQETK (28) GTGVITKKKP (24) QFYIVNTYIPNAG (29)

Tdd-3   GDFNCLDFNDNHTSND (33) KRPTFSRTIHNTNNNLTRILERRLDRIYLN (22) LSDHNFLSTTFTL
DRE     SDIITGDFNVDCSVDN (19) NGITFPR------------NKSTIDRVFVS (17) KSDHNMVIIELKI
L1Hs    HTLIMGDFNTPLSTLD (34) TEYTFFSAPHHT--------YSKIDHIVGS (16) LSDHSAIKLELRI
DdApeA  PIIWCGDLNVAHTEID (41) GSYTFWSYLGGGRSKNVGW---RLDYFVVS (18) GSDHCPIGVVVDL
```

RT domain

```
Tdd-3  TTIPFEIKRGVKQGDPLSPTLFVLVIEALARKILQDDRITGLPLNNSNHREKFQSFADDSASMVPDSQQFNSFCKATP
RED    TTRKFDIKRGVKQGDPISATLF--VIEILARTINADNTIIGLPQIKI----KFTQFADDSTTYNITYEQQQSIKHFK
DRE    LSKSFTSKRGTKQGDPISPTIFALVVECMATTIINDRCINGVTKETI----KILQFADDTATIAYNFMDHFLMNEWIK
L1hs   KLEAFPLKTGTRQGCPLSPLLFNIVLEVLARAIRQEKEIKGIQLGKE--EVKLSLFADDMIVYLENPIVSAQNLLKLI
```

HC domain

```
Tdd-3  LFNNILKIKDPKTRDTMFRFHARCLPINYLHNKQCPLCKEDMSKDPYGHLFFSCKKTKQFIKINKLKNFIYITTGKGK
RED    LFVKIQKLKDPKGRDTMQRFHARCLPINHLHNKVCPICNNEMNNDPYGHLFFNCQHTINFINHDKLKYFIV-KNCNGN
DRE    PFKEIKKIINMKGRDLLWRYTLKALPKIY--NMPCQQCGED---ETSEHIFFNCKAHI----KNTQEIFNVTLTKSGH
```

HC motif:

Fig. 3 Amino acid sequence alignments of Tdd-3 ORF2 domains with related proteins. Selected amino acid sequences of the Tdd-3-encoded endonuclease domain (EN), the reverse transcriptase domain (RT), and the histidine/cysteine-rich domain (HC) were aligned by visual comparison with the corresponding domains of DRE (Marschalek et al. 1992), human LINE-1 (L1Hs) (Feng et al. 1996), RED (Hitt et al. 1994), and the *D. discoideum* AP endonuclease homolog DdApeA (Freeland et al. 1996). Computer searches were then carried out using the BLAST algorithm (Altschul et al. 1990). Positions that are occupied by identical amino acids in at least two of the aligned sequences are shown against a *black background. Dashes* indicate gaps introduced to optimize the alignment. In the EN domain the numbers of amino acids connecting the homology domains are shown in *parentheses*

genome may be transcribed into polyadenylated mRNA species. PCR analysis was performed on single-stranded cDNA synthesized from oligo(dT)-primed poly(A)$^+$ RNA. One PCR primer was designed to bind 6 bp upstream of the oligo(dA) end of Tdd-3 elements. In RT-PCRs a specific DNA fragment was detected that was identical to the PCR fragment obtained using genomic DNA as template (data not shown). In order to exclude false positive results in RT-PCR reactions, poly(A)$^+$ RNA was purified and a cDNA library was generated. Amplification from these phages should exclude the possibility of contamination with genomic *D. discoideum* DNA. The phage library consisted of 30 different stocks, each representing a mixture of individual cloned cDNAs present in the primary library. Phage DNA was extracted from the individual phage stocks and subjected to PCR reactions using the primer pairs CT-20/CT-32 and CT-24/CT-30. Figure 4 summarizes the results of such a screening approach. A specific PCR fragment was detected in each phage stock using the primer pair CT-20/CT-32, which detects cDNAs cloned from the 3′ ends of full-length Tdd-3 transcripts. About half of the phage stocks were positive with the primer pair CT-24/CT-30, which detects cDNA fragments extending at least to the ORF1/ORF2 overlap (Fig. 4).

Discussion

Tdd-3 is a poly(A) retrotransposon

The work reported in this study describes the isolation of a new retrotransposable element from *D. discoideum*. Database searches revealed that the new retroelement was identical to partial sequences of Tdd-3, a previously identified repetitive DNA element from *D. discoideum*. However, the exact structure and transposition mechanism of Tdd-3 elements remained unclear. The data obtained in the present study enable various conclusions

to be drawn, which suggest that Tdd-3 is a tRNA gene-associated poly(A) retrotransposon.

Various cloned Tdd-3 elements show 5′ deletions. For poly(A) retroelements the existence of correctly integrated, 5′-truncated copies is characteristic and indicates that integration of such elements depends on an intact 3′ terminus. The most characteristic feature of retroelements is an encoded RT domain. The amino acid sequences of RTs have been used to establish phylogenetic relationships among retroelements (Xiong and Eickbush 1990). Computer searches revealed that the putative RT domain of Tdd-3 belongs to the class of poly(A) retroelements and shares little homology with retrovirus-like RTs. The Y/FXDD motif, which is important for RT activity, is also present in the putative Tdd-3 RT. In addition, the FADD motif (Fig. 3) is characteristic for RTs of poly(A) retroelements (Xiong and Eickbush 1990). The C-terminal part of Tdd-3 ORF2 has a significant homolog only in the RED element that was first identified through its association with Tdd-3 elements (Marschalek et al. 1990). RED has recently been identified in our laboratory as a poly(A) retroelement that is also targeted to sequences downstream of tRNA genes (T. Winckler, unpublished observation). Thus, it appears that Tdd-3 and RED belong to a new multimember family of tRNA gene-associated poly(A) retroelements in *D. discoideum* cells. Insertion of multiple retroelements adjacent to tRNA genes seems to be common in eukaryotes. It has been assumed that clustering of retroelements near tRNA gene loci may be the consequence of a coevolution process that enables host cells to avoid harmful mutations resulting from retroelement insertion into structural genes (Voytas and Boeke 1993; Voytas 1996).

Integration of poly(A) retroelements is catalyzed by enzymes that exhibit DNA-nicking activities and cleave both strands of the target DNA sequence in a staggered fashion. DNA repair mechanisms fill-in the single-stranded target sequences after insertion of the retroelements, resulting in target site duplications (TSD) of variable lengths. TSDs characteristic for Tdd-3 insertion were not determined in this study for two reasons. First, no genomic DNA fragments could be isolated that contained flanking regions on both sides of a Tdd-3 element. Second, it proved quite difficult to sequence the Tdd-3 termini which end in poly(dA) tracts of up to 70 bp. A 9–10 bp TSD was reported to flank a Tdd-3 element that was found inserted into Tdd-2 (Poole and Firtel 1984).

Expression of Tdd-3 elements

There is increasing evidence that poly(A) retroelements are actively transcribed from promoters located in the 5' UTR. It has been proposed that only full-length LINE-1 elements containing intact ORFs are able to transpose (Boeke 1997). Clone 5.1 isolated in this study contains about 3.3 kb of Tdd-3 coding region, including the complete ORF1 and half of ORF2. All genomic DNA clones containing 3' ends of Tdd-3 elements had an intact C-terminal half of ORF2. Thus it is likely that *D. discoideum* strain AX2 contains at least one functional Tdd-3 element. In our study we showed that Tdd-3 elements are most probably transcribed into full-length RNAs including both ORFs, since the PCR screen described in Fig. 4 identified DNA fragments both at the 3' end of putative Tdd-3-specific transcripts and at the end of ORF1. Thus, it can be assumed that Tdd-3 elements have the capacity to transpose actively in the *D. discoideum* cells. This agrees with the data of Poole and Firtel (1984), who showed that the activity of Tdd-3 elements caused restriction length polymorphisms in several wild-type and laboratory strains.

Tdd-3 elements are associated with tRNA genes

We have shown that Tdd-3 elements are targeted to sites about 100 bp downstream of *D. discoideum* tRNA

Fig. 4 Bulk screening of a cDNA library. A cDNA library was constructed from poly(A)$^+$ RNA prepared from vegetatively growing *D. discoideum* cells as described in the Materials and methods section. The library was represented in 30 aliquots, each containing approximately 20 000 individual phages. Phage lysates were subjected to PCR using the primer pairs indicated in the scheme. As positive controls PCR was performed with clone 5.1 (probe CT-24/30) or 4.25.2 (probe CT-20/32) (compare Fig. 1). PCR fragments were size-separated in 1% agarose gels, blotted onto Hybond-N membranes, and hybridized with the DNA fragments obtained in the positive controls

genes. Obviously the integration of Tdd-3 elements occurs independently of certain DNA sequences at the sites of integration. The data presented in Table 1 suggest that B box promoter equivalents may function as facilitating elements in determining the sites of Tdd-3 integration. TFIIIC is a B box-specific DNA-binding protein that initiates transcription by RNA polymerase III by recruiting additional factors to the tRNA gene (Geiduschek and Tocchini-Valentini 1988). A *D. discoideum* TFIIIC homolog has been purified that binds to B boxes and exB boxes of isolated tRNA genes with high affinity (Bukenberger et al. 1991, 1994). Integration of Tdd-3 elements may be guided by TFIIIC bound to B box or exB box motifs within the *D. discoideum* genome. In yeast, position-specific integration of the LTR retrotransposon Ty3 depends on protein-protein interactions with RNA polymerase III transcription factors (Kirchner et al. 1995). Genomic Tdd-3 copies identified downstream of exB boxes suggest that exB boxes can also be bound by TFIIIC in vivo. The physiological role of exB motifs in the transcription of *D. discoideum* tRNA genes is unclear. However, it is known that sequence-specific binding of factors to *cis*-acting elements may reorganize the nucleosomes in the chromatin to allow regulation of gene transcription of nearby genes by other transcription factors (Pazin et al. 1997).

Acknowledgements We thank D. Pätsch and S. Richter for expert technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft to T.D. and T.W., and from the American Heart Association to E. L. de H.

References

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Boeke JD (1997) LINEs and Alus - the polyA connection. Nature Genet 16:6–7

Boeke JD, Chapman K (1991) Retrotransposition mechanisms. Curr Opin Cell Biol 3:502–507

Boeke JD, Corces V (1989) Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol 43:403–434

Bukenberger M, Marschalek R, Salzburger W, Dingermann T (1991) Nuclear factors which bind to *Dictyostelium discoideum* transfer RNA genes. Curr Genet 20:129–135

Bukenberger M, Dingermann T, Meissner W, Seifart KH, Winckler T (1994) Isolation of transcription factor IIIC from *Dictyostelium discoideum*. Eur J Biochem 220:839–846

Feng QH, Moran JV, Kazazian HH, Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916

Freeland TM, Guyer RB, Ling AZ, Deering RA (1996) Apurinic/apyrimidinic (AP) endonuclease from *Dictyostelium discoideum*: cloning, nucleotide sequence and induction by sublethal levels of DNA-damaging agents. Nucleic Acids Res 24:1950–1953

Geiduschek E, Tocchini-Valentini G (1988) Transcription by polymerase III. Annu Rev Biochem 57:873–914

Hitt AL, Lu TH, Luna EJ (1994) Ponticulin is an atypical membrane protein. J Cell Biol 126:1421–1431

Hofmann J, Schumann G, Borschet G, Gosseringer R, Bach M, Bertling WM, Marschalek R, Dingermann T (1991) Transfer RNA genes from *Dictyostelium discoideum* are frequently associated with repetitive elements and contain consensus boxes in their 5'-flanking and 3'-flanking regions. J Mol Biol 222:537–552

Hutchison III C, Hardies S, Loeb D, Shehee W, Edgell M (1989) LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington DC, pp. 593–617

Kirchner J, Conolly C, Sandmeyer S (1995) Requirement of RNA polymerase III transcription factors for *in vitro* position-specific integration of a retroviruslike element. Science 267:1488–1491

Kozak M (1989) The scanning model for translation: an update. J Cell Biol 108:229–241

Luan D, Korman M, Jakubczak J, Eickbush T (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595–605

Marschalek R, Dingermann T (1991) Structure, organization and function of transfer RNA genes from the cellular slime mold *Dictyostelium discoideum*. Adv Gene Technol 2:103–143

Marschalek R, Brechner T, Amon-Böhm E, Dingermann T (1989) Transfer RNA genes: landmarks for integration of mobile genetic elements in *Dictyostelium discoideum*. Science 244:1493–1496

Marschalek R, Borschet G, Dingermann T (1990) Genomic organization of the transposable element Tdd-3 from *Dictyostelium discoideum*. Nuc Acids Res 18:5751–5757

Marschalek R, Hofmann J, Schumann G, Gösseringer R, Dingermann T (1992) Structure of DRE, a retrotransposable element which integrates with position specificity upstream of *Dictyostelium discoideum* tRNA genes. Mol Cell Biol 12:229–239

Martín F, Marañón C, Olivares M, Alonso C, López C (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the Ape family of DNA repair enzymes. J Mol Biol 247:49–59

Pazin M, Bhargava P, Geiduschek E, Kadonaga J (1997) Nucleosome mobility and the maintenance of nucleosome positioning. Science 276:809–812

Poole SJ, Firtel RA (1984) Genomic instability and mobile genetic elements in regions surrounding two discoidin I genes of *Dictyostelium discoideum*. Mol Cell Biol 4:671–680

Proudfoot N (1991) Poly(A) signals. Cell 64:671–674

Sambrook J, Fritsch E, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Schumann G, Zündorf I, Hofmann J, Marschalek R, Dingermann T (1994) Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the *Dictyostelium* Repetitive Element, from *Dictyostelium discoideum*. Mol Cell Biol 14:3074–3084

Vithalani K, Shoffner J, DeLozanne A (1996) Isolation and characterization of a novel cytokinesis-deficient mutant in *Dictyostelium discoideum*. J Cell Biochem 62:290–301

Voytas D (1996) Retroelements in genome organization. Science 274:737–738

Voytas D, Boeke JD (1993) Yeast retrotransposons and tRNAs. Trends Genet 9:421–427

Willis I (1993) RNA polymerase III: genes, factors and transcriptional specificity. Eur J Biochem 212:1–11

Xiong Y, Eickbush T (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9:3353–3362