Short interspersed repeats from *Xenopus* that contain multiple octamer motifs are related to known transposable elements

Garry T.Morgan* and Kim M.Middleton
Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, UK

Received June 1, 1990; Revised and Accepted September 5, 1990 EMBL accession no. Y00713+

ABSTRACT

We have identified in an intron of an *X. laevis* α-tubulin gene a member of a novel family of short (226 – 431 bp) interspersed repetitive elements. We have isolated other members of this family, which we term Ocr, from ovary cDNA and genome libraries and have identified another two in the published sequences of an H1B histone gene cluster and an actin gene intron. The termini of the Ocr elements are formed by a 19 bp inverted repeat that has clear sequence homologies to those of certain large transposable elements, such as 1723 (*Xenopus*) and *Ac* (maize). However, the Ocr elements do not appear to be deletion derivatives of larger transposons. The internal regions of the Ocr elements contain multiple copies of the octamer motif (ATTTGCAT) arranged as divergently-orientated dyads. We have shown by a gel mobility shift assay that these octamer dyads specifically bind what is presumably an OTF-type activator protein in oocyte nuclear extracts. We speculate that short interspersed repetitive families of this type may be generated by a mechanism of replicative transposition that uses a DNA intermediate and involves the interaction of DNA-binding proteins also utilised in other cellular processes.

INTRODUCTION

The short-period pattern of genome organization in which short repetitive sequences alternate with longer stretches of unique sequence DNA was first described in *Xenopus laevis* (1) and has since been shown to be a characteristic of most eukaryotes. The generation and dispersal of families of short repeats is thought to result from replicative transposition, although a clear idea of the mechanisms responsible is limited to those mammalian SINES like the Alu I family that appear to have arisen *via* self-primed reverse transcription (2). However, the by-products of reverse-transcription such as processed pseudogenes seem to be less common in other vertebrate genomes (3). Accordingly many of the families of short repeats known in any detail in *Xenopus* do not exhibit the characteristics of reverse transcripts but instead the organization of their terminal regions resembles those of transposable elements that are thought to move via DNA intermediates (4). However, beyond this general resemblance the nature of the relationship between short interspersed repeats and DNA-based transposons is unclear as is the molecular mechanism underlying their replicative transposition.

We describe here a novel family of short repetitive elements from *Xenopus* that has two features that may help to clarify these questions. First the sequence and organization of the repeats' termini are so similar in detail to those of a group of known transposons that it would appear that they must demonstrate the existence of some sort of evolutionary or functional relationship between the two types of element. However, in spite of the terminal similarities the internal regions of the two types of element appear unrelated. Secondly, these short repeats contain multiple copies of the octamer sequence motif (ATTTGCAT). This motif is an esential component of a variety of transcriptional promoters and enhancers (5, 6) and operates *via* the binding of activator proteins (7) that are also involved in DNA replication (8, 9). We have shown by a gel mobility shift assay that the octamer motifs in the *o*ctamer *c*ontaining *r*epeat (Ocr) elements specifically bind what is presumably such a protein from oocyte nuclear extracts. Thus the sequence characteristics of this family suggests that replicative transposition of short repeats does occur *via* a DNA intermediate and could involve the interaction of at least two types of DNA-binding protein that bind to conserved terminal and internal motifs.

MATERIALS AND METHODS

Recombinant clones and probes

The isolation and sequence of XαT14 has been described (10; GenBank Y00713) and its organization is shown in Fig. 1. The intron subclone mp.Tha140 was constructed by inserting a 141 bp *Tha* I fragment (Fig. 1) into the polylinker *Hinc* II site of

* To whom correspondence should be addressed

+Genbank accession no. M21286

M13 mp18 and the insert was verified by DNA sequencing. pSPT.Tha140 was derived by recloning this insert into pSPT19 (Boehringer) using flanking polylinker sites. The octamer-containing clone U4.pBLΔCAT was constructed by *Eco*R I digestion and vector religation of the plasmid U4.pBLCAT2 (a gift from Dr. D. Latchman), which comprises an oligonucleotide containing a U4 snRNA consensus octamer cloned into the polylinker of pBLCAT2 (11).

Hybridization probes were obtained either by nick-translation of the 880 bp *Acc* I-*Eco*RI fragment (Fig. 1) using standard methods (12) or by transcription of linearised pSPT.Tha140 with T7 RNA polymerase as described (13). The labelled 174 bp probe used in gel retardation assays was produced by fill-in with [^{32}P]-dCTP of the *Hin*d III and *Bam*H I termini created by excision of the insert fragment from the pSPT.Tha140 polylinker.

Blot hybridizations

For Southern blots five μg samples of genomic DNAs were digested with restriction enzymes using conditions recommended by the supplier and run on 1% agarose gels. After blotting to nitrocellulose, filters were hybridised at 65°C with $3-4\times10^7$ dpm of the RNA probe in 6×SSC, 5× Denhardt's solution, 25 mM sodium phosphate pH 6.5, 250 μg/ml *E. coli* tRNA and 0.1% SDS. Filters were washed at 65°C in 2×SSC, 0.1% SDS, then in 0.5×SSC, 0.1% SDS (i.e. allowing about 15% mismatch) and autoradiographed with intensifying screens.

For slot blots the indicated amounts of either *X. laevis* genomic DNA or single-stranded mp.Tha140 DNA were applied to Zeta-Probe (BioRad) in 0.4M NaOH and the filters then hybridised and washed as for Southern blots, though at salt concentrations calculated to give the required stringencies. Autoradiographs were quantified by densitometry and the conversion of signal strength to genomic copy number assumed that 80% of Ocr elements contained one copy of the region used as probe and that 20% possessed two copies (see Fig. 4a).

Isolation of Ocr recombinants

Recombinants from either an *X. laevis* genomic library constructed in EMBL4 by Dr. E. Jonas (supplied by Dr. T. Sargent) or an *X. laevis* ovary cDNA library constructed in λgt10 by Dr. R. Harland were screened by standard techniques (12) using the *Acc* I-*Eco*R I hybridization probe. Plaque lifts were washed under conditions requiring 80% sequence homology to the probe. Insert subfragments that hybridised with the RNA probe were subcloned into pTZ 18/19, subjected to unidirectional deletion with exonuclease III and their sequences determined on one strand. Sequence alignments were carried out using various routines from DNASTAR (Chiswick, London).

Nuclear extracts and DNA binding assays

X. laevis oocyte nuclei were isolated manually into 0.1M KCl, 20mM HEPES pH7.7, 5mM DTT, 1mM EDTA, 10% glycerol, and gently homogenised by aspiration (14). Binding reactions for gel mobility shift assays (15) contained 0.1−0.2 ng probe (about 2,000 dpm) that had been mixed with appropriate competitors and 1μg poly(dI-dC).poly(dI-dC), and 2.5 nuclear equivalents of extract in a volume made up to 15 μl with nuclear isolation buffer. After incubation at room temperature for 30 min, samples were loaded onto 5% polyacrylamide gels (60:1 acrylamide:*bis*), electrophoresed at 10 V/cm in 0.25×TBE and the gels autoradiographed undried.

RESULTS

During work aimed at locating transcriptional control signals of the *X. laevis* α-tubulin gene XαT14 (13), we noticed that a region of the second intron contained an overabundance of CpG dinucleotides (Fig. 1) and a number of copies of the octamer motif (see later). Since the former sequence characteristic can be a

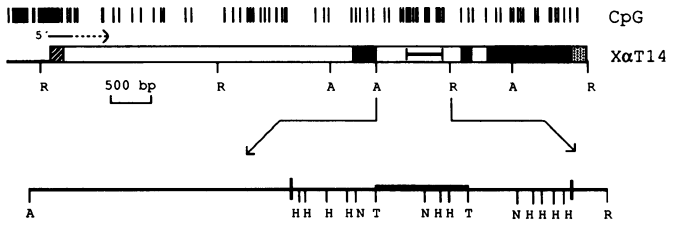

Figure 1. Organization of the *X. laevis* α-tubulin gene XαT14 and derivation of probes. On the top line the position of each CpG dinucleotide in the published XαT14 sequence (10) is shown as a vertical bar. Underneath is drawn the exon/intron organization of the gene (diagonal lines, 5' leader; open boxes, introns; black boxes, coding exons; dots, 3'untranslated region) and the position of Ocr1 in the second intron is indicated. On the bottom line are shown relevant restriction sites in the 880 bp *Acc* I-*Eco*R I fragment, with the extent of Ocr1 indicated by cross-bars. The thickened part of the line corresponds to the 141 bp *Tha* I fragment cloned into mp.Tha140 and pSPT.Tha140. Restriction sites: A, *Acc* I; H, *Hin*f I; N, *Nde* I; R, *Eco*R I; T, *Tha* I.

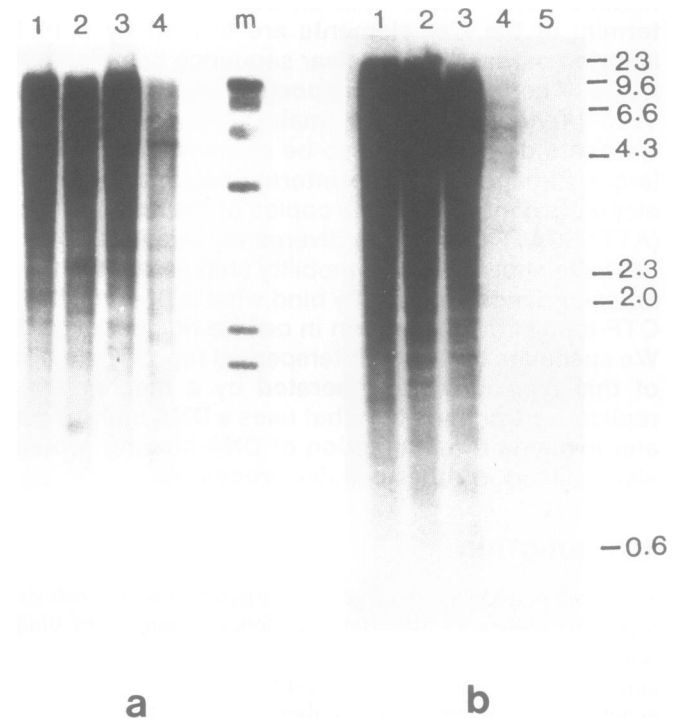

Figure 2. Genomic distribution of Ocr sequences (a) Southern blots of *X. laevis* genomic DNA digested with *Eco*R I (lane 1), *Pst* I (lane 2), *Hin*d III (lane 3) or *Nde* I (lane 4) and hybridised with an RNA probe made from the 141 bp *Tha* I fragment shown in Fig. 1. Note, there is an *Nde* I site within the probe fragment. m indicates marker lane containing λ/*Hin*dIII fragments. (b) Southern blots of *Eco*R I-digested genomic DNAs from *X. laevis* (lane 1), *X. laevis sudanensis* (lane 2), *X. borealis* (lane 3), *X. tropicalis* (lane 4) and *Ascaphus truei* (lane 5) probed as in a). Sizes of λ/*Hin*dIII marker fragments indicated in kb at the right.

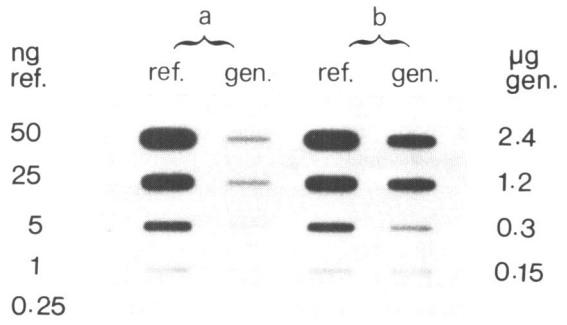

Figure 3. Estimation of genomic copy number of Ocr elements. Slot blots showing hybridization of the Tha140 antisense RNA probe to the amounts of mp.Tha140 sense-strand reference DNA (ref.) and *X. laevis* genomic DNA (gen.) indicated on the left and right, respectively. After hybridization the blots were washed at a criterion allowing either about 15% (a) or about 20% (b) mismatch.

feature not only of regulatory regions (16) but also of some repeated sequences (17), we probed Southern blots of *X. laevis* DNA with fragments containing the CpG-rich region. The resultant pattern of hybridization (Fig. 2a) showed that the intron did contain an interspersed repeat that we have called Ocr1. DNA from the closely-related species *X. laevis sudanensis* and *X. borealis* contained a similar number of repeats, but there were far fewer hybridizing bands in the more distantly related *X. tropicalis* and hardly any in the primitive frog *Ascaphus truei* (Fig. 2b). We estimate from DNA slot blots washed at a criterion allowing either about 15% or about 20% mismatch that there are 1,200 or 5,000 copies, respectively, of this repeat per haploid *X. laevis* genome (Fig. 3).

We screened 10^5 *X. laevis* /λ recombinants from genomic and ovary cDNA libraries with a probe containing the second intron and detected hybridization to about 2% and 0.2% of the clones, respectively. (The former figure predicts a number of repeats per genome closely in agreement with that determined from blots). The hybridizing regions from one genomic and two cDNA

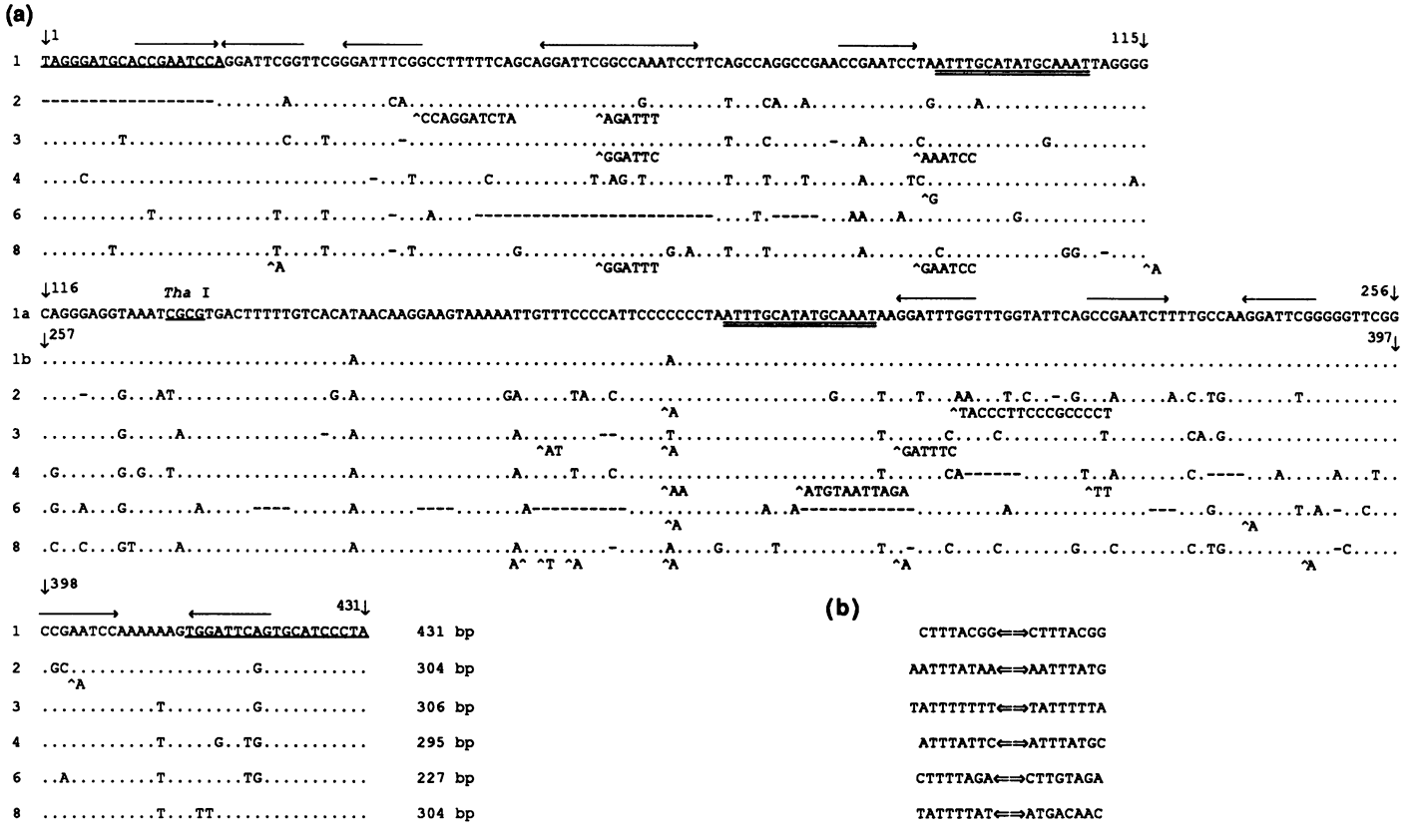

Figure 4. DNA sequences and organization of Ocr elements. (a) The sequence of repeat Ocr1 is shown as the top line of each group, with the lines being broken to indicate the internal organization of the Ocr elements. The 19 bp sequence forming the left hand inverted terminal repeat (single underline) is followed by a left flank containing multiple *Hinf* I motifs (horizontal arrows) and a 24 bp region of dyad symmetry containing at its centre two octamer motifs orientated divergently (double underline). Nucleotides 116–256 constitute a central region repeated in Ocr1 from 257–397 (1a and 1b) but present only once in the other elements. Central regions contain *Hinf* I motifs and an octamer dyad. Remaining nucleotides comprise a short right flank with a single *Hinf* I motif followed by the 19 bp right hand inverted repeat. Aligned below Ocr1 the sequences of the other Ocr copies are represented either by a dot for the same nucleotide or by an alternative nucleotide; dashes indicate missing nucleotides, and ^ the insertion of the indicated nucleotide(s). The size of each element is indicated at the end of each sequence. Ocr1 is nucleotides 4776–5206 of the XαT14 sequence obtained by Smith (10; GenBank Y00713), Ocr2 is nucleotides 9646–9949 (complementary strand) of the Xlh1 histone gene cluster sequenced by Perry *et al* (34; GenBank M21286) and Ocr4 is nucleotides 2678–2972 (complementary strand) of the sarcomeric α3I-actin gene sequenced by Mohun *et al* (35; GenBank X12525). Other sequences were obtained by us from clones isolated either from an *X. laevis* genomic library (Ocr3) or from an ovary cDNA library (Ocr6 and Ocr8). (b) Apparent target site duplications immediately flanking the inverted repeats (represented by open arrows) of the Ocr elements. Note that the duplications contain continuous runs of A/T base pairing and that although Ocr2 lacks a left inverted terminal repeat, it is flanked by an apparent target site duplication.

clones were subcloned and their DNA sequences determined. We also identified in the GenBank database regions similar in sequence to part of the tubulin intron about 1kb upstream of an *X. laevis* H1B histone gene and in the second intron of an *X. laevis* actin gene. Alignment of the six similar sequences demonstrated that they were all members of the same family of short (226 bp−430 bp) interspersed repeats, which we have termed Ocr elements, and made clear the organization of the repeating unit (Fig 4a).

The termini of the Ocr elements are formed by a 19 bp inverted repeat, a structure characteristic of prokaryotic and eukaryotic transposons thought to move via a DNA intermediate (18). Surprisingly, the consensus sequence of the Ocr inverted repeat is the same at 13 out of the first 15 nucleotides as that of the *X. laevis* transposon 1723 (4, 19; Fig.5). It has been noted previously that a number of transposons including 1723 have similar terminal sequences (20), and accordingly the Ocr terminus appears closely related to those of plant transposons such as *Ac* (Fig.5) and more weakly to the *Drosophila* transposons P and hobo. This relatedness is further suggested by the sequences flanking most of the Ocr copies, which appear as the perfect or near-perfect eight bp target site duplications characteristic of this group of transposons (Fig.4b). However, apart from the terminal regions there appear no other extensive sequence similarities between Ocr and 1723.

Ocr	TAGGGATGCACCGAATCCA
1723 (*Xenopus*)	TAGGGATGTAGCGAACGT
Ac/Ds (maize)	TAGGGATGAAA
Ips-*r* (pea)	TAGGGGTGGCAA
Tpc1 (parsley)	TAGGG-TGTAAA

Figure 5. Comparison of consensus sequence of Ocr terminal inverted repeat with transposons or transposon-like elements (20, 36). Regions of identical sequence are underlined.

The internal sequences of the Ocr elements possess a further identifiable element in 24 bp regions of dyad symmetry that when perfect, as in Ocr 1, contain at their centre ATTTGCATAT-GCAAAT. This dyad contains in divergent orientation two copies of the octamer motif, each of which also matches the stricter decanucleotide consensus (5). Ocr1 therefore contains six octamers and, although in the other Ocr sequences some of the dyads have been altered by mutation and there is only one copy of the entire central region, all contain one or more perfect octamers. In addition to the octamers, the internal Ocr sequences contain multiple copies in either orientation of a different eight base pair motif that usually contains a *Hinf* I site and that is also found in the terminal inverted repeats (Fig. 4a).

In order to test whether the octamer dyad in Ocr1 is a substrate for octamer-binding proteins such as those present in *X. laevis* oocytes (21), we asked whether a gel mobility shift could be induced by incubation of the 141 bp *Tha* I fragment of Ocr1 (Fig. 1) with isolated oocyte nuclear extract (Fig 6a). The specific band shift detected was competed by the addition of an increasing molar excess of a competitor plasmid, pSPT.Tha140, containing the *Tha* I fragment but not by addition of the same excess of vector DNA. When pSPT.Tha140 was first digested with *Nde* I (CA↓TATG), which cleaves the octamer dyad and leaves incompletely-double stranded octamer motifs, its competitive strength was drastically reduced almost to that of vector alone. However, competitive strength was unaffected by prior digestion of pSPT.Tha140 with *Hinf* I, which cuts twice within the insert but not in the octamer dyad. A plasmid containing a single octamer consensus from a U4 snRNA promoter also competed the band shift (Fig. 6b) though much less effectively than a comparable molar excess of pSPT.Tha140, perhaps because octamer-binding proteins can bind cooperatively to a divergently arranged octamer dyad (22). It appears from these results that a protein present in oocyte nuclei that behaves operationally as an octamer-binding protein can bind to the octamer dyads of the Ocr family.

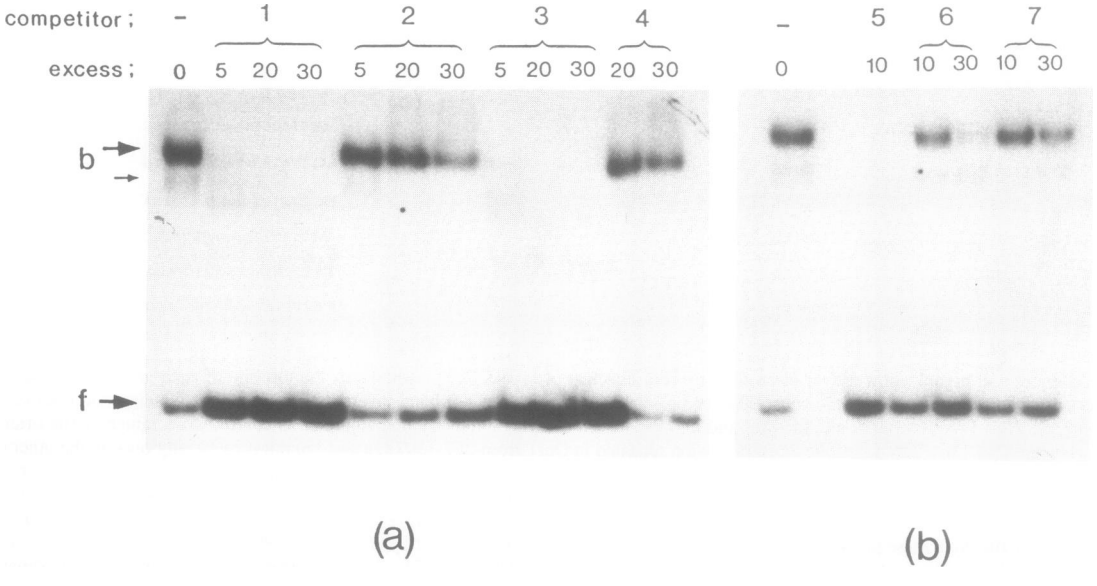

Figure 6. Gel mobility shift assays of protein binding by Ocr1 octamers. (a) Left hand lane shows the band shift induced by binding of oocyte nuclear proteins to the *Tha* I fragment of Ocr1. Large arrows indicate free (f) and bound (b) probe, small arrow a minor bound species. Other lanes show the effects on binding of the indicated molar excesses of the following competitors; (1) pSPT.Tha140 linearised in polylinker with *Eco*R I and *Bam*H I (2) *Nde* I-digested pSPT.Tha140 (3) *Hinf* I-digested pSPT.Tha140 (4) *Nde* I-digested pSPT19. (b) As (a), except undigested competitor plasmids were; (5) pSPT.Tha140 (6) U4.pBLΔCAT (7) pTZ18.

DISCUSSION

We have found that a region of the second intron of an *X laevis* α-tubulin gene contains multiple octamer motifs. This motif is an essential component of a variety of transcriptional control regions including the intragenic enhancer of immunogobulin genes (5). However, in apparent contradiction to our initial assumption that part of the tubulin intron might therefore represent a transcriptional enhancer, we have shown that in fact this region contains a member of a family of widely dispersed repetitive elements all of which contain octamers. Although we cannot rule out the possibility that the ability shown here of Ocr elements to interact with octamer binding proteins *in vitro* could indicate that they have a role in, or effect on, the *in vivo* control of cellular activities such as transcription, because of the limited phylogenetic distribution of Ocr elements (Fig.2b) we believe that the presence of octamers is more likely to be explained by their involvement in the family's mode of replicative transposition (see below).

The detailed resemblance in sequence and structure of the termini of the Ocr elements and the *Xenopus* mobile element 1723 provides evidence for an evolutionary or functional relationship between short interspersed repetitive families and DNA-based transposons that is more compelling than the more general similarities previously noted (e.g. 4, 23). It is apparent also that the Ocr family comprises a set of closely-related elements and not simply a collection of deletion derivatives of 1723. Deletions of 1723 have apparently occurred in the *X. laevis* genome (24) just as they have amongst P-elements in *Drosophila* and the maize *Ac* transposons (25, 26). However, Ocr and 1723 appear as two distinctive types of element, their relatedness being confined to the possession of similar termini in much the same way as the so-called aberrant or type-I *Ds* elements appear related to *Ac* (27, 28). It seems possible that rather than representing defective deletion derivatives, short repeats such as Ocr actually could represent the primitive progenitors (18) from which mature transposable elements may have evolved.

The presence of sequences closely related to TAGGGATG in the inverted terminal repeats of diverse transposons (Fig. 5) is a surprising feature from which it has been inferred that a highly-conserved DNA-binding protein(s) encoded by 'host' rather than transposon genes may be involved in transposition (20). In agreement with this it has recently been shown that the analogous region of the P-element inverted terminal repeat specifically binds a protein that is produced in cells lacking P-elements and thus presumably also operates in basic cellular processes other than P-element transposition (29). Our results suggest that such a protein may also have been co-opted in the generation and maintenance of families of short interspersed repeats in a wide range of organisms.

The evolutionary conservation of multifunctional DNA-binding proteins and the motifs they recognise is a common feature of eukaryotic transcriptional activators. Some of these activators, such as CTF (NF-I), Mcm1 and the octamer-binding protein OTF-1 (NF-III) are important in the initiation of replication as well as transcription (30, 31, 9). The ability of the Ocr family to interact with octamer-binding protein(s) could be interpreted as reflecting the involvement of such activators in additional processes such as replicative transposition. Perhaps the assembly on short repeats of multiprotein complexes that include activators like OTF-1 and the proteins binding to conserved terminal sequences could initiate a mechanism of DNA-based replicative transposition. Such a mechanism would involve proteins normally playing other roles in the cell and would not require dedicated transposases encoded by mobile elements nor, obviously, a reverse transcribed intermediate. Since other short interspersed repeats have been shown to bind the transcriptional activators Sp1 or NF-I (32,33), it seems possible that the evolution and maintenance of different repeat families of this type may eventually prove to be driven by distinctive combinations of various activators and other proteins.

ACKNOWLEDGEMENTS

We are grateful for gifts of clones or libraries to Drs. Richard Harland, David Latchman, Tom Sargent and Desmond Smith, and to Evelyn Gurd for technical assistance. We thank the National Westminster Bank Research Fund and University of Nottingham Research Fund for financial support.

REFERENCES

1. Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973) J. Mol. Biol. **77**, 1–23
2. Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) Cell **26**, 141–142
3. Wagner, M. (1986) Trends Genet. **2**, 134–137
4. Carroll, D., Knutzon, D.S. and Garrett, J.E. (1989) in 'Mobile DNA' (D.E Berg and M.H. Howe, eds.), 567–574 (American Society for Microbiology, Washington D.C.).
5. Falkner, F.G. and Zachau, H.G. (1984) Nature **310**, 71–74
6. Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) Proc. natn. Acad. Sci. USA **81**, 2650–2654
7. Schaffner, W. Trends Genet. (1989) **5**, 37–39
8. Pruijn, G.J.M., van Driel, W. and van der Vliet, P.C. (1986) Nature **322**, 656–659
9. O'Neill, E.A., Fletcher, C., Burrow, C.R., Heintz, N., Roeder, R.G. and Kelly, T.J. (1988) Science **241**, 1210–1213
10. Carroll, D., Knutzon, D.S. & Garrett, J.E. in Mobile DNA, 567–574 (American Society for Microbiol., Washington D.C., 1989). Smith, D.J. (1988) Biochem.J. **249**,465–472
11. Luckow, B. and Schütz, G. (1987) Nucl. Acids Res. **15**, 5490
12. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in 'Molecular Cloning. A Laboratory Manual' (Cold Spring Harbor Laboratory)
13. Middleton, K.M. and Morgan, G.T. (1989) Nucl. Acids Res. **17**, 5041–5055
14. Middleton, K.M. and Morgan, G.T. (1990) Mol. Cell. Biol. **10**, 727–735
15. Fried, M. and Crothers, D.M. (1981) Nucl. Acids Res. **9**, 6505–6525
16. Bird, A.P. (1986) Nature **321**, 209–213
17. Cooper, D.N., Gerber-Huber, S., Nardelli, D., Schubiger, J.-L. and Wahli, W. (1987) J. Mol. Evol. **25**, 107–115
18. Finnegan, D.J. (1985) Int. Rev. Cytol. **93**, 281–325
19. Kay, B.K. and Dawid, I.B. (1983) J. Mol. Biol. **170**, 583–596
20. Streck, R.D., McGaffey, J.E. and Beckendorf, S.K. (1986) EMBO J. **5**, 3615–3623
21. Mocikat, R., Falkner, F.G., Mertz, R. and Zachau, H.G. (1986) Nucl. Acids Res. **14**, 8829–8844
22. LeBowitz, J.H., Clerc, R.G., Brenowitz, M. and Sharp, P.A. (1989) Genes Devel. **3**, 1625–1638
23. Flavell, R.B. (1985) in 'Genetic Flux in Plants' (B.Hohn and E.S. Dennis, eds.), 139–156 (Springer-Verlag, Wien New York)
24. Martens, G.J.M. and Herbert, E. (1984) Nature **310**, 251–254
25. O'Hare, K. and Rubin, G.M. (1983) Cell **34**, 25–35
26. Fedoroff, N.V. (1989) in 'Mobile DNA' (D.E Berg and M.H. Howe, eds.), 375–411 (American Society for Microbiology, Washington D.C.)
27. Sutton, W.D., Gerlach, W.L., Schwartz, D. and Peacock, W.J. (1984) Science **223**, 1265–1268
28. Döring, H.-P. and Starlinger, P. (1986) Ann. Rev. Genet. **20**, 175–200
29. Rio, D.C. and Rubin, G.M. (1988) Proc. natn. Acad. Sci. USA **85**, 8929–8933
30. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelley, T.J. and Tjian, R. (1987) Cell **48**, 79–87
31. Passmore, S., Maine, G.T., Elble, R., Christ, C. & Tye B.K. (1988) J. Mol. Biol. **204**, 593–606
32. Dynan, W.S., Saffer, J.D., Lee, W.S. and Tjian, R. (1985) Proc. natn. Acad. Sci. U.S.A. **82**, 4915–4919

33. Farache, G., Razin, S.V., Targa, F.R. and Scherrer, K. (1990) Nucl. Acids Res. **18**, 401–409
34. Perry, M., Thomsen, G.H. and Roeder, R.G. (1985) J. Mol. Biol. **185**, 479–499
35. Mohun, T., Garrett, N., Stutz, F. and Spohr, G. (1988) J. Mol. Biol. **202**, 67–76
36. Bhattacharyya, M.K., Smith, A.M., Ellis, T.H.N., Hedley, C. and Martin,C. (1990) Cell **60**, 115–122