Conserved Sequence and Functional Domains in Satellite 2 from Three Families of Salamanders

Barbara Green, * Lil M. Pabón-Peña, † Teresa A. Graham, ‡ Sharon E. Peach, ‡ Stephen R. Coats, * and Lloyd M. Epstein ** †

*Institute of Molecular Biophysics, †Department of Biological Science, and ‡Department of Chemistry, Florida State University

Thirteen satellite 2 elements from Ambystoma talpoideum and 16 from Amphiuma tridactylum were cloned, sequenced, and compared with the satellite 2 consensus from Notophthalmus viridescens. These elements have maintained a high degree of similarity during the 65–200 Myr that the salamander families, represented by the three species, have been separated. The DNA sequences of the consensus elements from A. talpoideum and A. tridactylum are 81% similar, and both are ~65% similar to the N. viridescens consensus. In addition to its DNA sequence, the functional properties of satellite 2 have also been conserved. By selecting and analyzing clones that closely mimicked the consensus of each species, we were able to demonstrate that satellite 2 from each species was capable of promoting transcription after injection into Xenopus laevis oocytes and that synthetic transcripts of satellite 2 from each species were capable of catalyzing their own site-specific cleavage. These properties may be related to the process of retroposition, which was previously proposed to be responsible for the genomic proliferation of satellite 2. Each of these functional properties also has general biological interest.

Introduction

The prototypical satellite DNA consists of short tandemly repeated units that are localized to heterochromatic regions of the chromosomes (Brutlag 1980). While there are only minor sequence differences between the repeats within a species, there is little, if any, conservation of satellites between species. Finally, satellites are transcriptionally inert, and proposed functions for satellites have concentrated on aspects of DNA or chromosome structure (Brutlag 1980). It is not surprising that exceptions to this generalized description exist and are especially prominent in amphibians. Various examples of transcribed satellites have been reported (Diaz et al. 1981; Ackerman 1983; Kay and Dawid 1983; Hummel et al. 1984), and high-copy-number tandem repeats that are conserved between species and found in both heterochromatin and euchromatin are common (Lam and Carroll 1983; Epstein et al. 1986; Macgregor and Sessions 1986; Vignali et al. 1991). For these reasons amphibians have been popular subjects for studying the function and evolution of highly repetitive DNAs, as well as for studying the role that these DNAs have had in shaping the amphibian genome (Macgregor and Sessions 1986).

Our efforts have concentrated on an unusual element called “satellite 2.” Ap-
approximately 100,000 copies of this 330-bp sequence are organized in tandem arrays that are dispersed throughout the genome of the newt, *Notophthalmus viridescens* (Epstein et al. 1986). In addition, this element is highly conserved in all species of the family Salamandridae that have been investigated (Epstein et al. 1986; Macgregor and Sessions 1986; Cremisi et al. 1992). For example, satellite 2 elements from *N. viridescens* and the European newt, *Triturus cristatus*, are 85% similar at the nucleotide level (Epstein et al. 1986), even though these species separated ≥50 Mya (Duellman and Trueb 1986, pp. 477–492). The selective forces that maintain the sequence of satellite 2 probably act in part on the strand-specific satellite 2 transcripts found in all tissues examined. The predominant transcripts in the *N. viridescens* ovary are the size of the basic DNA repeat unit; all other tissues are enriched for larger, multimeric forms (Epstein et al. 1986). The function of these transcripts is not known, but it is possible that they are involved in the genomic propagation of satellite 2. We previously proposed a model in which multimeric satellite 2 transcripts serve as templates for reverse transcription and the production of cDNA copies (Epstein et al. 1986). Random genomic integration of these cDNAs would then generate dispersed tandem arrays of satellite 2. This process is analogous to the retroposition process believed to be responsible for the generation of primate Alu elements, rodent B1 elements, and a variety of other SINES and processed pseudogenes in the vertebrate genome (Weiner et al. 1986; Deininger 1989). Evidence for retroposition of satellite 2 is indirect and includes the finding that the consensus of eight monomeric satellite 2 DNA clones from *N. viridescens* was almost identical to the sequence of the homogeneous population of ovarian satellite 2 transcripts. This indicated that a subset of satellite 2 DNA was used to produce both the cellular transcripts and the bulk of the genomic DNA repeats, which are diverging randomly from their common progenitor. The simplest mechanism to account for this is that the repeats are made by using transcripts of the progenitor as intermediates in retroposition.

On the premise that they are involved in the propagation and evolution of satellite 2 DNA, we have extensively studied these transcripts and the mechanism of their production. Two properties of satellite 2 are directly related to transcript formation. First, the satellite 2 DNA repeat contains an octamer and a proximal sequence element (PSE), which promote accurate and efficient transcription of cloned satellite 2 DNA after injection into *Xenopus laevis* oocytes (Cremisi et al. 1992; S. R. Coats and L. M. Epstein, unpublished data). These elements were originally identified by their similarity to the core promoters of small nuclear RNA genes transcribed by RNA polymerase II (Dahlberg and Lund 1988). Transcription promoted by these elements in satellite 2 initiates at a site that corresponds to the major 5' end of the ovarian transcripts. Therefore these endogenous promoter elements are likely involved in ovarian transcript formation.

Transcription alone cannot account for the production of satellite 2 transcripts in nonovarian tissues, because the major 5' end of these transcripts is 46–47 nt downstream from the transcription initiation site (Epstein and Gall 1987; Epstein and Coats 1991). These transcripts appear to be generated by a processing event that is catalyzed by the RNA itself. We previously demonstrated that synthetic satellite 2 transcripts undergo self-catalyzed cleavage in vitro at a single site within each repeat (Epstein and Gall 1987). The reaction requires a divalent cation and generates 5'-hydroxyl and 2',3'-cyclic phosphate groups. The position and chemical composition of the ends
of the nonovarian satellite 2 transcripts are identical to those produced by in vitro self-cleavage (Epstein and Coats 1991).

Interest in self-cleavage extends beyond its role in the molecular biology of satellite 2. Self-cleavage occurs within a domain that is similar to the "hammerhead" motif found in a number of infectious plant RNAs that utilize self-cleavage during their replication cycle (Branch and Robertson 1984; Symons 1989, 1992). The consensus hammerhead consists of three base-paired stems surrounding a single-stranded central core (fig. 1A). The satellite 2 cleavage domain has many of the conserved features of the hammerhead but also has some unique modifications (fig. 1B). First, stem III is atypically unstable, consisting of only 2 bp connected by a 2-nt loop. Second, stem I has an internally looped extension that is required for cleavage of the satellite 2 transcripts but not for cleavage of the plant RNAs (Epstein and Pabón-Peña 1991; Pabón-Peña et al. 1991). As a result of these novel structural variations, the satellite 2 hammerhead, the only hammerhead found in an animal, is making unique contributions to studies of the mechanism and evolution of this catalytic RNA motif.

The strong conservation of satellite 2 within the Salamandridae suggested that we might find related elements in other families of salamanders. We anticipated that a comparative sequence analysis of these elements would augment our functional studies by providing information about the constraints and requirements in the various functional domains. In this report we describe the isolation and characterization of satellite 2 elements from Ambystoma talpoideum (family Ambystomatidae) and Amphiuma tridactylum (family Amphiumidae). Regions of extended nucleotide similarity are interspersed between more variable regions in these elements. Many of the conserved regions correspond to known components of the transcriptional promoter or self-cleavage domains. As predicted, differences exist in these regions that will be useful in our analysis of the functional domains. At least two other conserved regions were found that may be previously unidentified accessory elements for transcription or self-cleavage or that may be involved in unrelated functions. The comparative sequence analysis has therefore delineated conserved regions of satellite 2 that are appropriate targets for further investigation.

**Material and Methods**

**Salamanders**

Adult specimens of Ambystoma talpoideum and Ambystoma opacum were collected in Tallahassee. Amphiuma tridactylum were collected in Louisiana and pur-
chased from a local distributor. Laboratory-raised *Xenopus laevis* were purchased from Xenopus I.

Genomic DNA

Techniques for preparing genomic DNA and tissue RNA were modified from our previous protocols (Epstein et al. 1986). For DNA, whole livers were frozen in liquid nitrogen and were ground with a mortar and pestle. One hundred milligrams of the powdered tissue was incubated in 1.2 ml of digestion buffer [100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 25 mM ethylenediaminetetraacetate (EDTA), 0.5% sodium dodecyl sulfate (SDS), 0.1 mg proteinase K/ml] at 50°C for 7 h. DNA was then purified from RNA and residual proteins by centrifugation on a CsCl cushion consisting of 1.67- and 1.74-g/cm³ layers. The DNA was recovered from the interface and dialyzed against 1 × TE [10 mM Tris-HCl (pH 8.0), 1 mM EDTA] overnight at 4°C.

For the preparation of total RNA, 0.5 g of tissue was homogenized in 5 ml of lysing medium [4 M guanidine isothiocyanate, 1% diethyl pyrocarbonate, 0.1 M potassium acetate (pH 5.0)] on ice. The homogenate was extracted twice with a 4:1 mixture of phenol:chloroform and twice with a 1:1 mixture of phenol:chloroform. Nucleic acids were precipitated by adding 1/20 vol of 4.5 M sodium acetate and 1/2 vol of 95% ethanol to the final aqueous phase. RNA was then purified from DNA and proteins by centrifugation over a CsCl cushion as described above. The clear RNA pellet was resuspended in 1 × TE, and residual CsCl was removed by a final round of precipitation.

Satellite 2 Clones

pGM1 and pG47:293 are monomeric satellite 2 clones from *Notophthalmus viridescens*. These clones have different permutations of the basic satellite 2 repeat, and their derivation from the genomic dimer clone pSP6D6 has been described elsewhere (Epstein and Pabón-Peña 1991).

To clone satellite 2 from *Ambystoma talpoideum* and *Amphiuma tridactylum*, DNA from each species was digested with AluI and electrophoresed on a 1.75% agarose gel. After a staining with ethidium bromide, a segment of the gel containing DNA fragments of ~330 bp was excised. The DNA was purified from the gel slice by electroelution and was ligated into the SmaI site of the phage vector M13mp10 (Messing 1983).

Initially, plates of recombinant phage containing *Ambystoma talpoideum* inserts were screened by the method of Benton and Davis (1977), using gel-purified 330-bp genomic DNA from *Ambystoma opacum* as a probe. Preliminary Southern blots indicated that both *Ambystoma* species contained satellite 2-related sequences that migrated with 330-bp DNA after digestion with AluI. We reasoned that, by using the nonhomologous size-selected DNA probe, the number of non-satellite 2 clones that screened positive would be reduced. A single satellite 2 clone was obtained in this manner, (pAtal-7), and synthetic transcripts representing both strands of this clone were used to find the remaining *Ambystoma talpoideum* clones. Seven *Amphiuma tridactylum* clones (pAtri-A, -B, -C, -D, -H, -K, and -N) were then obtained by using probes derived from the *Ambystoma talpoideum* clones pAtal-2 and -3. The remaining *Amphiuma tridactylum* clones were found by using probes derived from pAtri-A and -D.

The inserts and some flanking vector sequences were excised from the M13 clones
as HindIII/EcoRI fragments and inserted into the phagemid vector pGEM3zf(+) (Promega). Single-stranded DNAs corresponding to both strands of each insert were obtained from the M13 and pGEM clones, and both strands were sequenced in their entirety by the dideoxynucleotide method (Sanger et al. 1980). The pGEM clones also served as templates for the production of synthetic transcripts of the cloned inserts, with use of either the SP6 or T7 RNA polymerase promoters of the vector. These transcripts were used as probes in the various filter hybridizations or for substrates in the in vitro self-cleavage assays.

Since the AluI restriction site in the Ambystoma talpoideum and Amphiuma tridactylum clones dissected the satellite 2 repeat within the self-cleaving hammerhead domain, clones used to analyze self-cleavage were modified to restore continuous hammerhead domains. This was accomplished by site-directed mutagenesis using single-stranded uracil-containing templates and variant oligonucleotide primers (Kunkel et al. 1987). The primers used for mutagenesis were 5'-TGAGGAGTAGCTTGGGGATCCTC-3' and 5'-CAGTGAGGTGCAGCTTAAGGGGGATCCTCTA-3', for Ambystoma talpoideum and Amphiuma tridactylum, respectively. In addition, 31 bp of the vector located between the SP6 RNA polymerase promoter and the 5' end of the satellite 2 insert were removed, as follows: DNA was linearized by digestion with BamHI and HindIII, the 5' overhangs were removed with S1 nuclease, and the blunt ends were ligated. Transcripts from the resulting clones (pAtal-4-5'H and pAtri-D-5'H) contained five vector-derived nucleotides preceding a complete hammerhead domain at their 5' ends.

DNA Sequence Analysis

Sequence comparisons were performed by using the Alignment program in the Microgenie Software Package (Beckman) and by realigning the output manually to minimize the number of gaps. For computing the similarity between two sequences, gaps were counted as single characters, irrespective of their length. Multiple sequence alignments were performed by using the Pileup and Pretty programs in the GCG Software Package (Genetics Computer Group).

Blots and Filter Hybridizations

Southern blots of genomic DNA and northern blots of tissue RNA have been described elsewhere (Epstein et al. 1986). Blots were processed and hybridized at 42°C also as described, by using synthetic transcripts of the indicated satellite 2 clones as probes. The probes were generated by transcribing linearized template DNA with either SP6 or T7 RNA polymerase at 37°C for 60 min in the presence of 40 mM Tris-HCl (pH 7.5), 6 mM MgCl₂, 5 mM spermidine, 10 mM DTT, 250 μM each of ATP, CTP, and GTP, 25 μM UTP, and 15 μCi [α-32P]UTP.

In Vitro Self-Cleavage Reactions

Synthetic transcripts of cloned satellite 2 were prepared by using the same reactions described above for probe preparation, except that the concentration of unlabeled UTP was lowered to 15 μM. Transcripts were purified from polyacrylamide gels and were analyzed for self-cleavage according to a method described elsewhere (Epstein and Gall 1987; Pabón-Peña et al. 1991). Self-cleavage reaction mixtures consisted of 0.002–0.004 ng/μl of transcripts, 133 mM morpholineethanesulfonic acid (pH 6.9), 30 mM MgCl₂, 10 mM NaCl, and 0.3 mM EDTA. Cleavage products were separated
Conserved Domains in Satellite 2 737

on denaturing polyacrylamide gels and were quantified with a Betascope 603 Blot
Analyzer (Betagen Corporation) according to a method described elsewhere (Pabón-

Transcription of Satellite 2 Clones in Xenopus Oocytes

Ovaries were dissected from mature X. laevis, and oocytes were manually removed
from the ovaries in modified Barth’s solution [88 mM NaCl, 1 mM KCl, 2.4 mM
NaHCO3, 0.82 mM MgSO4, 7.5 mM Tris-HCl (pH 7.6), 0.33 mM Ca(NO3)2, 0.41
CaCl2, 100 U penicillin/ml, 100 μg streptomycin/ml, 0.25 μg amphotericin B/ml].
Each oocyte nucleus was injected with ~50 nl of water containing 40 ng/μl of cloned
satellite 2 DNA and 0.3 ng/μl of cloned X. laevis 5S maxigene (Bogenhagen and

For the preparation of RNA from injected oocytes, batches of 15–20 oocytes
were homogenized at room temperature in 500 μl of homogenization buffer [100 mM
sodium acetate (pH 5.0), 5 mM EDTA, 0.5% SDS]. The homogenate was extracted
twice with phenol and once with chloroform and was precipitated by the addition of
1/30 vol of 4.5 M sodium acetate and 2 vol of 95% ethanol. The precipitate was
pelleted, resuspended in 75 μl water, and reprecipitated with an equal volume of 8 M
lithium chloride. A final round of precipitation with ethanol was performed to remove
residual lithium chloride. The final RNA preparation was analyzed by primer exten-
sions with radioactively labeled oligonucleotide primers according to a method de-
scribed elsewhere (Epstein et al. 1986).

Results

Sequence Analysis

As described in Material and Methods, 13 genomic clones from Ambystoma
talpoideum and 16 from Amphiuma tridactylum were obtained and sequenced. The
sequences of the Ambystoma talpoideum clones are shown in the left-hand panel of
figure 2. The satellite 2 segments of these clones are 79%–99% similar to their common
consensus. Six of the clones are complete monomers, indicating that they were derived
from internal segments of tandem arrays. pAtal-5, -8, and -11 are truncated because
of mutations that generated internal AluI sites. Three of the remaining clones (pAtal-
7, -9, and -12) are mixed clones consisting of partial satellite 2 repeats flanked on
either end by non–satellite 2 sequence. In clones pAtal-7 and -9 the position where
satellite 2 is joined to non–satellite 2 sequence is similar (approximately position 255
of the consensus). Clone pAtal-13 has three large internal deletions. Moreover, the
first 57 positions of this clone seem to be derived from the inversion of an internal
segment (positions 249–306 of the consensus) of the satellite 2 repeat.

The sequences of the Amphiuma tridactylum clones are shown in the right-hand
panel of figure 2. The satellite 2 segments in these clones are 88%–98% similar to their
common consensus. Five clones are complete monomers, which are arranged as in-
dicated in the figure. Three other clones (pAtri-B, -C, and -K) are monomers that
begin and end at a secondary AluI site at position 227 of the consensus. Several more
clones (pAtri-E, -F, -G, and -I) are incomplete monomers, which are bounded by the
primary AluI site on one end and by the secondary AluI site on their other end. These
results suggest that there are two potential AluI sites within any given repeat unit, and
these sites may or may not be present in the same tandem array. Further support for
this conclusion was obtained from the Southern analysis (see below). The remaining
FIG. 2.—Nucleotide sequences of satellite 2 elements from *Ambystoma talpoideum* and *Amphiuma tridactylum*. Above, Sequences of the 13 *A. talpoideum* clones, shown aligned to their common consensus. The clones were ordered and named according to their overall similarity to the consensus. A consensus nucleotide is one present in >50% of the clones. Dashes represent identity to the consensus, dots represent gaps, arrowheads indicate the first nucleotide and orientation of the original M13 clones, and N's represent either undetermined nucleotides in the clones or unspecified nucleotides in the consensus. Facing page, Sequences of the 16 *A. tridactylum* clones, aligned to their common consensus.
FIG. 2 (Continued)
clones (pAtri-H, -J, -O, and -P) are partial repeats flanked on either end by non-satellite 2 sequences. The satellite 2 segment in pAtri-J ends at position 254 of the consensus, which is analogous to the endpoint of satellite 2 in the Ambystoma talpoideum clones pAtal-7 and -9. The satellite 2 segment of pAtri-H ends just a few nucleotides downstream of this same position, making a total of four clones, from the two species, that have junctions in the same vicinity. Of the two remaining mixed clones (pAtri-O and -P), the junction in pAtri-P at position 142 is very similar to the junction at position 146 in the Ambystoma talpoideum clone pAtal-12. In contrast to the similarities between the satellite 2 portions of the various junctions, there are no similarities between the non-satellite 2 segments of any of the mixed clones from either Ambystoma talpoideum or Amphiuma tridactylum.

The consensus sequences of the Ambystoma talpoideum and Amphiuma tridactylum clones are compared with the Notophthalmus viridescens satellite 2 consensus, in figure 3. The Ambystoma talpoideum and Amphiuma tridactylum consensuses are 81% similar to each other but are only 65% and 64% similar, respectively, to the N. viridescens consensus. Several conserved regions correspond to portions of the hammerhead self-cleaving domain or elements of the transcriptional promoter. Other regions exist that are conserved but that have not yet been assigned a function, including CR1 and CR2 in figure 3.

Genomic Organization

To investigate the genomic organization of satellite 2, Southern blot hybridizations were performed on a series of partial AltI digests of DNA from Ambystoma talpoideum and Amphiuma tridactylum. The progressive decrease in the size of multimeric products

---

**Fig. 3.—Alignment of the consensus sequence of satellite 2 elements from three families of salamanders.**

The consensus sequences for Ambystoma talpoideum (*A. tal*) and Amphiuma tridactylum (*A. tri*) are from fig. 2. The Notophthalmus viridescens (*N. vir*) consensus is from Epstein et al. (1986). This last sequence was permuted such that position 1 corresponds to position 9 of the previously published sequence. Regions where three or more nucleotides are conserved in all three consensuses are in boldface. The location of known components of the transcriptional promoter (Octamer and PSE) and the self-cleaving hammerhead domain (Stem I and Core) are indicated above the sequences. The locations of two conserved regions with unknown function are also indicated (CR1 and CR2).
with increased times of digestion indicates that a substantial portion of satellite 2 in *Ambystoma talpoideum* is organized in tandem arrays (fig. 4, lanes 1–5). Arrays of 10 monomeric units were visible in lane 3 of the original autoradiogram, and it is likely that larger arrays exist but, because of the decrease in the ratio of band to background intensity with increasing molecular weight, were not differentiable.

Complete digestion of *Amphiuma tridactylum* DNA with *Alu*I generated the expected 325-bp monomeric fragment, but an equivalent amount of a 250-bp fragment was also formed (fig. 4, lane 10). This fragment was most likely derived from repeats that have the primary *Alu*I site at position 1 and the secondary *Alu*I site at position 227 as numbered in the consensus (fig. 2, right). The larger fragments generated by partial *Alu*I digestion correspond to the fragments expected from different combinations of primary and secondary *Alu*I sites in adjoining repeats. For example, the 650-bp fragment represents a perfect dimer generated when a repeat with neither *Alu*I site is

---

**FIG. 4.**—Southern analysis. Samples (2.5 µg) of genomic DNA from either *Ambystoma talpoideum* or *Amphiuma tridactylum* were digested with the restriction enzyme *Alu*I for the indicated times. After electrophoresis on 1.5% agarose gels, the DNA was blotted to nitrocellulose filters and was probed with synthetic transcripts from pAtal-4 (lanes 1–5) or pAtri-D (lanes 6–10). The sizes of the major bands, as determined by using *Hind*III-digested λ DNA and *Hae*III-digested φX DNA as standards, are indicated on the right.
flanked by repeats that have the same AluI sites (either the primary or secondary sites but not both). The 550-bp fragment is derived from digestion at the primary site in one repeat and at the secondary site in the adjacent upstream repeat. Each of the bands visualized in lanes 6–10 of figure 4 can be explained in similar manners. Thus, despite the complex pattern of fragments generated by digestion with AluI, we conclude that satellite 2 in Amphiiuma tridactylum is arranged in simple tandem arrays.

Cellular Satellite 2 Transcripts

Ovary and liver RNA from Ambystoma talpoideum and Amphiiuma tridactylum was probed for the presence of transcripts that hybridize to satellite 2 by northern blot hybridizations (fig. 5). For comparison, parallel hybridizations were performed with

---

**Fig. 5.—**Northern analysis of satellite 2 transcripts. Ovary (lanes O) and liver (lanes L) RNA from the indicated salamanders was transferred to nitrocellulose filters and was probed with synthetic RNA probes derived from pGM1 (lanes 1 and 2), pAtri-D (lanes 3 and 4), or pAtal-4 (lanes 5 and 6). Even-numbered lanes contained 20 µg of total liver RNA. Odd-numbered lanes contained 10 µg of total ovary RNA, except for lane 5, which had 20 µg of Ambystoma talpoideum ovary RNA. Exposure times of 3.5 h (lanes 1 and 2), 93 h (lanes 3 and 4), and 132 h (lanes 5 and 6) were required to obtain the comparable autoradiographic signals seen here. From previous studies of Notophthalmus viridescens (Epstein et al. 1986), the transcripts labeled “monomer,” “dimer,” and “trimer” have sizes of ~330, ~660, and ~990 nt, respectively.
RNA from N. viridescens. In each case, probes were derived from clones that closely mimicked the consensus of the species being investigated.

As seen elsewhere (Epstein et al. 1986; Epstein and Coats 1991), ovary RNA from N. viridescens has an abundant transcript that corresponds in size to the monomeric satellite 2 DNA unit, while liver RNA is enriched for larger transcripts that are perfect multiples of the repeat unit (fig. 5, lanes 1 and 2). Similar transcript patterns were found in the ovary and liver from Amphiuma tridactylum (lanes 3 and 4) and Ambystoma talpoideum (lanes 5 and 6), but there were differences between the three species. First, the transcripts were more abundant in N. viridescens than in either of the other two species. In light of the different quantities of RNA analyzed and of the different exposure times of the autoradiograms shown in figure 5, the ovary monomers were ~25 and ~750 times more abundant in N. viridescens than in Amphiuma tridactylum and Ambystoma talpoideum, respectively. Ambystoma talpoideum liver RNA was further differentiated by the appearance of a second multimeric series of satellite 2 transcripts slightly larger than the predominant series. Repeats >325 bp were not evident in either the cloning or the Southern analysis, and the source of these larger transcripts is not presently known.

Transcriptional Properties of Satellite 2 Clones

The comparison of consensus sequences in figure 3 indicated that the PSE and octamer promoter elements, important for transcription of N. viridescens satellite 2 by RNA polymerase II, were conserved in Ambystoma talpoideum and Amphiuma tridactylum. Nevertheless, there are differences in the PSEs from the three species. To determine whether these differences correlate to a loss or decrease in transcriptional activity, we subjected selected clones to a transcriptional analysis. Clone pAtal-2, which is identical to the Ambystoma talpoideum consensus in the octamer and PSE regions, was chosen as the Ambystoma talpoideum representative. Of the Amphiuma tridactylum clones with permutations suitable for this analysis (see fig. 6A), none were identical to their consensus in both the octamer and PSE regions. Two clones, pAtri-N and -D, which have one and four differences from their consensus in the octamer and PSE, were therefore chosen for analysis, by virtue of having the least divergent promoters of the available Amphiuma tridactylum clones. These clones, as well as the N. viridescens clone pG47:293 (Epstein and Pabón-Peña 1991), were injected into X. laevis oocytes, and the resulting transcripts were analyzed by primer extensions. A cloned X. laevis 5S maxigene was coinjected with each satellite 2 clone as an internal control (Bogenhagen and Brown 1981).

Three primers were used in the primer-extension analysis. The first primer was specific for the 5S maxigene transcript and gave the 136-nt extension product seen near the top of all lanes in figure 6B. The other two primers were complementary to the universal forward and reverse primer sites in the vector. As illustrated in figure 6A, if transcription from the satellite 2 clones initiated at the expected sites, then the forward primer (but not the reverse primer) would be complementary to these transcripts and would result in discretely sized extension products. Figure 6B shows that the major extension products obtained with the forward primer correspond well to the expected products (lanes 3, 5, 7, and 9) and that these products are absent in extensions from the reverse primer (lanes 4, 6, 8, and 10). These results indicate that each of the satellite 2 elements have active promoter elements and that each initiates transcription in the vicinity of the expected site. The decreased activity of pAtri-D is...
Fig. 6.—Transcription of cloned satellite 2 DNA in *Xenopus laevis* oocytes. A, Generalized map of the clones used in this analysis. Each clone consisted of a monomeric unit of satellite 2 DNA permuted so that the promoter elements (octamer and PSE) would cause transcription to initiate at a site near the right end of the insert and to continue into the adjacent vector sequences. If transcription initiated at the expected site, the reverse primer would not be complementary to the transcripts, while the forward primer would hybridize and give 76-nt extension products with the *Notophthalmus virideescens* transcripts and 83-nt products with *Ambystoma talpoideum* and *Amphiuma tridactylum* transcripts. B, Primer extensions of RNA extracted from *X. laevis* oocytes injected with DNA from the indicated satellite 2 clones. The forward (f) primer was used in lanes 1, 3, 5, 7, and 9, while the reverse (r) primer was used in lanes 2, 4, 6, 8, and 10. As an internal control, all injections included DNA from a cloned *X. laevis* 5S maxigene, and the primer-extension reactions included a primer specific for the 5S maxigene transcript. The 136-nt product from this primer is evident near the top of each lane. Sizes (in nt) of coelectrophoresed size markers (i.e., pBR322 *Hae*III fragments) are indicated to the right.
most likely due to the divergence of its promoter elements from the *Amphiuma tridactylum* consensus and illustrates the importance of using the consensus as a prototype for the functional element.

**Self-Cleavage Abilities**

The ability to form a modified hammerhead structure is also conserved in satellite 2 from the three families of salamanders (figs. 1B and 7A). Each structure has the conserved central core of the consensus hammerhead (fig. 1A) but differs from the consensus by having an abbreviated stem III and an internally looped extension to stem I. Despite these similarities, nucleotide differences occur in the hammerhead domains of the three satellite 2 elements. To determine the effect that these species-specific nucleotide differences have on self-cleavage, we tested the cleavage properties of transcripts containing each of these domains. Clone pAtal-4, which has a hammerhead domain identical to its species consensus, and, pAtri-D, which has a hammerhead domain with a single nucleotide different from its species consensus, were chosen for this analysis. The *AluI* sites in these clones interrupted the hammerhead motifs in the lower portion of the stem I extension, and, as a result, their transcripts were incapable of self-cleavage (data not shown). We therefore reconstructed these clones (pAtal-4-5'H and pAtri-D-5'H) so that they would produce transcripts with complete and continuous hammerhead domains at their 5′ ends. Figure 7B illustrates that these transcripts were capable of self-cleavage and compares their cleavage to the

![Diagram](image)

**Fig. 7.**—Self-cleavage of satellite 2 transcripts from the different salamanders. A, Potential extended hammerhead structures in transcripts from the *Ambystoma talpoideum* and *Amphiuma tridactylum* clones used in this analysis (pAtal-4-5'H and pAtri-D-5'H). Conventions used are as in fig. 1, with additional boxes in stem I enclosing nucleotides that are conserved in these species and in *Notophthalmus viridescens*. Vector-derived nucleotides at the 5′ end of these transcripts are denoted by lowercase letters, and numbers in brackets indicate the number of nucleotides at the 3′ ends that are not shown here. B, Synthetic transcripts were prepared in vitro by using EcoRI-digested DNA from the indicated clones as templates for transcription with SP6 RNA polymerase. Transcripts were incubated under self-cleavage conditions for the indicated times and electrophoresed on a 7.5%-acrylamide, 7-M urea gel. Expected sizes (in nt) of the transcripts and their cleavage products are indicated on the right. The small 5′ cleavage products of the pAtal-4-5'H and pAtri-D-5'H transcripts were electrophoresed off the gel in order to resolve the larger RNAs.
self-cleavage of transcripts of the *N. viridescens* clone pGM1. As previously shown, the 390-nt pGM1 transcripts cleaved in the presence of Mg$^{2+}$ to generate 216- and 174-nt products (lanes 1–4). Incubating the 347-nt pAtal-4-5'H transcripts (lanes 5–8) and the 352-nt pAtri-D-5'H transcripts (lanes 9–12) under identical conditions resulted in the appearance of fragments that had mobilities consistent with the sizes expected for the 3' products of self-cleavage. (The shorter 5' cleavage products were electrophoresed off the gel, in order to resolve the larger RNAs.) Enzymatic RNA sequencing of these products confirmed that cleavage occurred at the sites indicated by the arrows in figure 7A (data not shown). Although it appears from figure 7B that the pAtri-D-5'H transcripts were the least active of the three transcripts, a more extensive series of time points indicated that the initial rate of cleavage of the pAtri-D-5'H transcripts was nearly double the initial rates of the pGM1 and pAtal-4-5'H transcripts (data not shown). Despite this rapid initial rate, cleavage of pAtri-D-5'H transcripts plateaued early, and >75% of the transcripts failed to cleave, even after 3 h. Similar kinetics have been found for other hammerhead RNAs and were attributed to individual RNA molecules folding into inactive configurations prior to the start of the reactions (Sheldon and Symons 1989; Fedor and Uhlenbeck 1990; Pabón-Peña et al. 1991). The present results indicate that the ability to self-cleave has been conserved in the three salamander families, although slight modifications to the hammerhead design have resulted in different kinetics of self-cleavage.

**Discussion**

Satellite 2 elements from *Ambystoma talpoideum* and *Amphiuma tridactylum* were cloned and compared with homologous elements from *Notophthalmus viridescens*. While there are quantitative and qualitative differences at both the DNA and RNA levels in the three species, the basic functional properties of satellite 2 have been maintained for the 65–200 Myr that these species have been separated (Estes 1965; Salthe and Kaplan 1966; Maxson and Wilson 1979; Larson 1991). These properties include the possession of functional transcriptional promoter elements, the ability to produce stable cellular transcripts with sizes that correspond to monomers and multimers of the DNA repeat unit, and the ability of these transcripts to catalyze their own site-specific cleavage. The significance of these properties is not known, but they can be related to our previous model for the proliferation of satellite 2 by retroposition. Together these properties result in the production of cellular transcripts with tissue-specific configurations. Some of these transcripts, by virtue of their end groups or overall tertiary structure, may be particularly good substrates for germ-line retroposition. The transcripts with other configurations may have other cellular functions or may simply be the products of mechanisms that have evolved to inhibit retroposition and to protect the genome from the damaging effects of excessive satellite 2 movement.

Of the 29 genomic clones analyzed in the present study, 7 were mixed, in that they had junctions between satellite 2 and non–satellite 2 sequences. In our previous analysis of *N. viridescens*, no mixed clones were found among 12 monomeric and dimeric clones (Epstein et al. 1986; Epstein and Gall 1987). The recovery of mixed clones in the present study might indicate that repeats at the ends of tandem arrays represent a significant fraction of the total population and that the tandem arrays in *Ambystoma talpoideum* and *Amphiuma tridactylum* are shorter than those in *N. viridescens*. While this is supported by Southern blots of partial genomic digests from the three species (fig. 4; also see Epstein et al. 1986), it is also possible that differences
in the cloning procedures used in these studies differentially biased the types of clones recovered. Regardless of the source of these clones, they provide important information that might relate to the putative retroposition process. In three of the mixed clones the junction is at the 5' end of satellite 2, with respect to the strand shown in figure 3, and in two of these (pAtal-12 and pAtri-P) the junctions' actual positions in satellite 2 are nearly identical. The remaining four mixed clones (pAtal-7 and -9 and pAtri-H and -J) have junctions at the 3' end of the satellite 2 sequence, and these four junctions' positions in satellite 2 are also nearly identical. Possibly these junctions represent the 5' and 3' ends of the original units of retroposition. While the putative 3' borders lack the track of A residues characteristic of most retroposons, some retroposons or retroposed elements have been found that lack this poly-A track (Bernstein et al. 1983; Spence et al. 1985). If retroposition really was the mechanism of satellite 2 proliferation, then the ends of the retroposed elements would correspond to the ends of the RNA intermediates. The junctions observed in the present study do not correspond to the RNA ends produced by either transcription or by self-cleavage. This does not rule out the possibility that (a) reverse transcription begins and ends at positions within the mature transcripts or (b) undiscovered processing events occur to generate retroposable transcripts. On the other hand, the common junctions may represent positions in satellite 2 that are hot spots for illegitimate recombination with non-satellite 2 sequences. The most direct way to distinguish between these possibilities would be to clone complete tandem arrays of satellite 2 and to examine the ends of these arrays, for direct repeats or other hallmarks of retroposition. This would also enable one to examine the possibility that in these three species there are significant differences in the lengths of tandem arrays that might also be pertinent to the retroposition process.

The conservation of satellite 2 extends beyond the known functional domains discussed above. Regions such as CR1 and CR2 (fig. 3) may modulate either transcription or self-cleavage or may be involved in totally unrelated functions. This can be investigated by targeting these regions for mutation and testing the mutants for transcription and self-cleavage. Further investigations will involve data-base searches for similar regions in other genetic elements. Satellite 2 seems to be a composite of functional domains that individually occur in other genes or RNAs. The hammerhead self-cleaving domain is characteristic of infectious plant RNAs (Symons 1989), while the transcriptional promoter is similar to the promoter of RNA polymerase II–transcribed small nuclear RNA genes (Cremisi et al. 1992). The region labeled "CR1" in figure 3 is similar to the sequence 5' TGGCTTTGGCAGTGAGG 3' found in the pTvm13 family of repeats (Vignali et al. 1991). This family, consisting of dispersed single copies of a 419-bp sequence, is present in a number of Salamandridae species but is absent in N. viridescens. The conservation of the CR1 sequence in two otherwise unrelated elements suggests that it has a novel role in the function or evolution of repetitive DNA in salamanders. As more satellite 2 sequences are obtained, the sequences of other conserved regions will be defined more precisely and used as probes in data-base searches for similarities to other genes or RNAs.

The most immediate value of the sequence information obtained in this study is for the mechanistic analysis of the known functional domains. This is evident for the self-cleavage domain. Each of the salamanders investigated has active hammerhead domains with internally looped extensions to stem I. Despite our previous demonstration that single nucleotide substitutions in the N. viridescens hammerhead were detrimental to cleavage (Pabón-Peña et al. 1991), the nucleotide identities of the stem
I loops from the three species are not identical. Multiple substitutions during the divergence of these three species have apparently resulted in alternative designs for active hammerhead domains. Furthermore, it is possible that these loops are evolving in parallel with other regions of the hammerhead that also vary between the species. This is presently being examined by testing the activity of transcripts with artificially constructed chimeric hammerhead domains. The sequence information obtained in this study has therefore led to the design of experiments that will provide insights into the structural constraints and interactions in the modified satellite 2 hammerhead.

Finally, the sequences obtained in this study have unexpected evolutionary implications. Similarity relationships between the satellite 2 elements do not correlate to phylogenetic relationships derived from the comparison of morphological characteristics (Edwards 1976), ribosomal RNA gene sequences (Larson and Wilson 1989; Larson 1991), or other biochemical data (Salthe and Kaplan 1966; Maxson and Wilson 1979). While there is considerable disagreement between the various salamander phylogenies that have been proposed, none of the phylogenies suggest that Ambystomatidae and Amphiumidae are more similar to each other than either is to Salamandridae. Nevertheless, the satellite 2 elements from these families are related in precisely this manner. While it may be premature to propose a specific pattern for the evolution of satellite 2, the same relationships between these families were obtained by comparing chromosome structure (Morescalchi 1975). Specifically, representatives from Ambystomatidae and Amphiumidae have similar karyotypes consisting of either 26 or 28 chromosomes, most of which are metacentric. The Salamandridae have what is considered to be the most evolutionarily advanced karyotype of 22 metacentric chromosomes. The sequences of satellite 2 elements from additional salamander families are needed to further investigate the possibility that satellite 2 evolution parallels chromosome structure rather than phylogeny.

Sequence Availability

All sequences have been submitted to GenBank under the following accession numbers: AMPH1, L05837; AMPH14, L05838; AMPH16, L05839; AMPH19, L05840; AMPH2, L05841; AMPH23, L05842; AMPH27, L05843; AMPH28, L05844; AMPH29, L05845; AMPH30, L05846; AMPH31, L05847; AMPH33, L05848; AMPH34, L05849; AMPH36, L05850; AMPH37, L05851; AMPH6, L05852; AMPHCON, L05853; ATAL1, L05854; ATAL2, L05855; ATAL8, L05856; ATALA, L05857; ATALBB, L05858; ATALCON, L05859; ATALD, L05860; ATALL, L05861; ATALL, L05862; ATALM, L05863; ATALN, L05864; ATALO, L05865; ATALU, L05866; AND ATALY, L05867.

Acknowledgments

This work was supported by NIH grant GM 38791 to L.M.E. and by a James R. Fisher fellowship from the American Cancer Society to T.A.G.

LITERATURE CITED


ALAN M. WEINER, reviewing editor

Received November 1, 1992; revision received January 27, 1993

Accepted February 12, 1993