Gene, 172 (1996) 183–190

GENE 09723

Cloning and characterization of extended hammerheads from a diverse set of caudate amphibians

(Recombinant DNA; RNA self-cleavage; satellite 2; salamander)

Yi Zhang* and Lloyd M. Epstein

Institute of Molecular Biophysics, and Department of Biological Science, Florida State University, Tallahassee, FL 32306-3050, USA

Received by M. Belfort: 18 August 1995; Revised/Accepted: 2 December/21 December 1995; Received at publishers: 15 February 1996

SUMMARY

Satellite 2 (*sat2*) elements from six species representing four caudate amphibian families were cloned and analyzed. Despite differences in the abundance of this element, the tandemly repeated genomic organization found for *sat2* in earlier investigations was conserved in the newly investigated species. The ability of *sat2* transcripts to catalyze their own site-specific self-cleavage using an extended hammerhead (HH) motif had also been conserved. This motif is similar to the HH found in a number of infectious plant RNAs, but has an abbreviated stem III region and an internally looped stem I extension that distinguish it from the prototypical HH. The extended HH analyzed in this survey can be further organized into two groups on the basis of shared sequences and structural details in the peripheral stem-loop structures. HH derived from the family Salamandridae constitute one group, whereas HH from all other salamanders that have been investigated belong to the other group. Although many of the natural variants did not cleave efficiently in vitro, examples of relatively active variants for each design were found. This survey has therefore identified two designs that are suitable starting models for the structural and functional analysis of the extended HH.

INTRODUCTION

Satellite 2 (*sat2*) is an approx. 330-bp repetitive DNA that is highly conserved in salamanders (Green et al., 1993). Tandem arrays of this repeat are dispersed throughout the genome of the newt, *Notophthalmus viridescens*, where *sat2* was first characterized (Epstein et al., 1986). Some of these arrays are transcribed using internal snRNA-like promoter elements (Coats et al., 1994), and *sat2* transcripts are capable of site-specific self-cleavage (Epstein and Gall, 1987). A cellular function for these transcripts has not been found, and we proposed that they are used during the propagation of *sat2* by retroposition (Green et al., 1993). With the goal of better defining the nature of *sat2*, we have devoted considerable effort to characterizing its properties and have paid particular attention to the self-cleavage reaction.

Self-cleavage occurs within a domain that is similar to the hammerhead (HH) domain used during the replication of a number of infectious plant RNAs (Fig. 1; Bruening, 1989). In the presence of a divalent cation, a single phosphodiester within this domain is cleaved to generate products with 5′-hydroxyl and 2′,3′-cyclic phosphate groups (Symons, 1992). The self-cleavage domain in *sat2* transcripts from *N. viridescens* has the same core and secondary structure as the prototypical HH but has

Correspondence to: Dr. L.M. Epstein, Department of Biological Science, Florida State University, Tallahassee, FL 32306-3050, USA. Tel. (1-904) 644-4560; Fax (1-904) 644-0481; e-mail: epstein@bio.fsu.edu
*Present address: Department of Biochemistry, University of Medicine and Dentistry of New Jersey, 663 Hoes Lane, SRB, Piscataway, NJ 08854-5635, USA. Tel. (1-908) 235-4194.

Abbreviations: bp, base pair(s); CR1 and CR2, conserved regions 1 and 2; HH, hammerhead(s); N, A/C/G/T; nt, nucleotide(s); oct (OCT), octamer; oligo, oligodeoxyribonucleotide; PA, polyacrylamide; PCR, polymerase chain reaction; PSE, proximal sequence element; *sat2*, satellite 2; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; ss, single strand(ed).

SSDI 0378-1119(96)00126-6

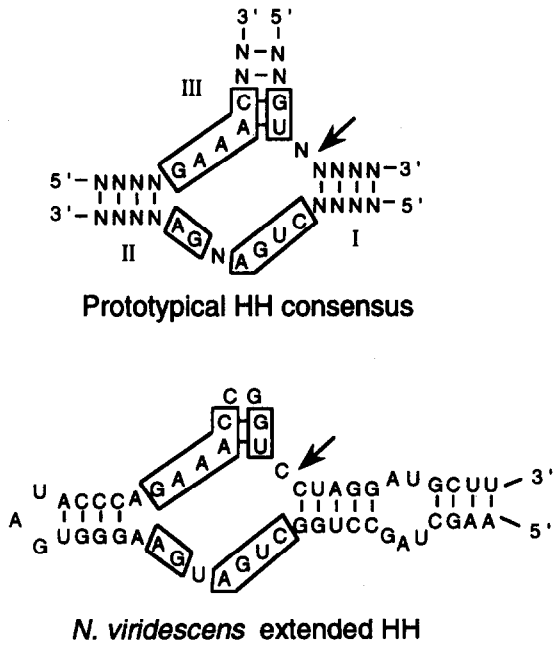

Fig. 1. HH structures. On top is the consensus sequence and secondary structure of the HH domain required for self-cleavage of viroid, virusoid, and viral satellite RNAs. Stems are labeled I, II, and III; conserved nt are boxed; and the arrow denotes the site of self-cleavage according to the conventions of Forster and Symons (1987). The *Notophthalmus viridescens* extended HH domain is from Pabón-Peña et al. (1991).

two unusual features (Fig. 1). First, stem III consists of only 2 bp and a 2-nt loop. The stability of this 2-nt hairpin loop was questioned, and it was proposed that *sat2* uses an alternative structure, called the double HH, for self-cleavage (Forster et al., 1988). The double HH consists of two cleavage domains that interact to form two HH connected by a single and more stable stem III. Despite evidence that interdomain interactions are important for some transcript configurations (Forster et al., 1988; Epstein and Pabón-Peña, 1991), cleavage by *sat2* can use single-domain structures (Epstein and Pabón-Peña, 1991). This cleavage depends on the second novel feature of the *sat2* HH, which is the internally looped extension to stem I (Pabón-Peña et al., 1991). Cleavage activity was abolished when this extension was replaced with a nonhomologous stem or when the nt in the internal loop were individually mutated (Pabón-Peña et al., 1991). Similarly, the identities of the nt in the external stem II loop were also found to be important for self-cleavage (L.M.E., T.A. Graham, and N. Gokaldas, unpublished data).

The plant HH do not have comparable requirements in their peripheral structures. For these HH, any stable stems support cleavage as long as they do not promote folding into inactive RNA conformations (Fedor and Uhlenbeck, 1990; Heus et al., 1990; Tuschl and Eckstein, 1993). It has even been shown that the entire stem II region can be replaced with a short ss loop without a dramatic loss of activity (McCall et al., 1992; Hendry et al., 1994). The absence of specific requirements in the peripheral stems, and the clustering of important functional groups in the conserved core (Heidenreich et al., 1993), strongly suggest that the catalytic center is located in the central core. Because this core is conserved in the *sat2* HH, *sat2* probably uses the same active site for cleav-

TABLE I

Sources of *sat2* clones

Order	Family	Species	Number of clones sequenced[a]	Length of consensus (bp)[b]	Range of similarities[c] (%)	Median similarity (%)
Anura	Pipidae	*Xenopus laevis*	n.a.[d]	n.a.[d]	n.a.[d]	n.a.[d]
Gymnophiona	Typhlonectidae	*Typhlonectes natans*	n.a.[d]	n.a.[d]	n.a.[d]	n.a.[d]
Caudata	Sirenidae	*Siren lacertina*	n.a.[d]	n.a.[d]	n.a.[d]	n.a.[d]
Caudata	Salamandridae	*Cynops pyrrhogaster*	8 (type 1)	105	92.4–98.1	96.7
			9 (type 2)	105	86.7–95.2	91.4
Caudata	Proteidae	*Necturus maculosus*	17	108	77.1–97.2	94.4
Caudata	Plethodontidae	*Plethodon glutinosus*	17	112	68.8–99.1	95.5
		Eurycea longicauda	21	110	81.8–100	96.4
		Desmognathus apalachicolae	20	114	77.8-98.2	95.7
Caudata	Cryptobranchidae	*Cryptobranchus alleganiensis*	19	107	82.2–95.3	90.7

[a] Partial *sat2* clones were obtained and sequenced as follows. The products of PCR amplification with the CR1 and CR2 primers were blunt-ended by treatment with S1 nuclease, phosphorylated by treatment with T4 polynucleotide kinase, purified from 5% PA gels, and ligated to the plasmid vector pGEM3Zf(–) that was blunt-end linearized with *Hin*cII. Double-stranded DNAs from positive clones were sequenced using the dideoxynucleotide method (Sanger et al., 1980). Primers complementary to the Sp6 and T7 RNA polymerase promoters, which flank the MCS polylinker region of pGEM3Zf(–), were used to sequence both strands of the 145–154 bp inserts in their entirety.

[b] Consensus sequences were generated using the multiple sequence alignment programs Pileup and Pretty in the GCG Software Package (Genetics Computer Group, Madison, WI). A consensus nt is one that is present in >50% of the clones.

[c] Similarities are those between individual clones and the consensus sequence, and were determined using the alignment program of the Microgenie Sequence Software package (Beckman, Palo Alto, CA).

[d] Not applicable. In these species, *sat2* elements were not detected by either PCR amplification or the Southern blot analyses.

age. Stems I and II of the extended HH are therefore more likely to be involved in helping to stabilize the active structure than to be components of the active site. Crystal structures for two prototypical HH provide a basis for this stabilization (Pley et al., 1994; Scott et al., 1995). In these structures, stem III is coaxially stacked with stem II. This stabilizing effect is unlikely to occur in the extended HH because of its abbreviated stem III, and other interactions may be required to stabilize the active structure. Such interactions might have developed between stems I and II, which lie roughly parallel and in close proximity in the crystal structures, explaining the requirements in these stems.

The present study was designed to generate a set of active variants for the structural and functional analysis of the extended HH. By comparing available *sat2* sequences, we previously identified several highly conserved regions that included portions of the extended HH domain, components of the transcriptional promoter, and two segments with unknown function called CR1 and CR2 (Green et al., 1993). Because CR1 and CR2 flank the extended HH domain, these sequences can be used to amplify extended HH by means of the polymerase chain reaction (PCR). Using this strategy we surveyed three amphibian orders for the presence of *sat2*. Wherever *sat2* was found, it was cloned and sequenced, and the self-cleavage activity of transcripts from these clones was analyzed.

RESULTS AND DISCUSSION

(a) Phylogenetic survey of the occurrence of *sat2*

The species screened for the presence of *sat2* by PCR amplification are listed in Table I. Fig. 2B shows that amplification with the CR1 and CR2 primers gave the expected 150-bp product when either cloned *sat2* DNA (lane 3) or genomic DNA (lane 4) from *N. viridescens* was used as a template, but not when the reaction was performed without exogenous DNA (lane 2). Similar products were obtained when 10-pg samples of DNA from five of the newly investigated species were used as templates (lanes 5–9). Although 10 pg of DNA from *Cryptobranchus alleganiensis* was not sufficient to generate a visible product (lane 10), the expected product was obtained when the amount of template was raised to either 1 ng or 100 ng (lanes 11 and 12). This result suggests that *sat2* is present in *C. alleganiensis* at a significantly lower copy number than in the other species. Specific amplification products were not obtained from *Siren lacertina*, *Typhlonectes natans*, or *Xenopus laevis* when either 10-pg or 100-ng samples of DNA were used as templates (lanes 13–15 and data not shown). Either

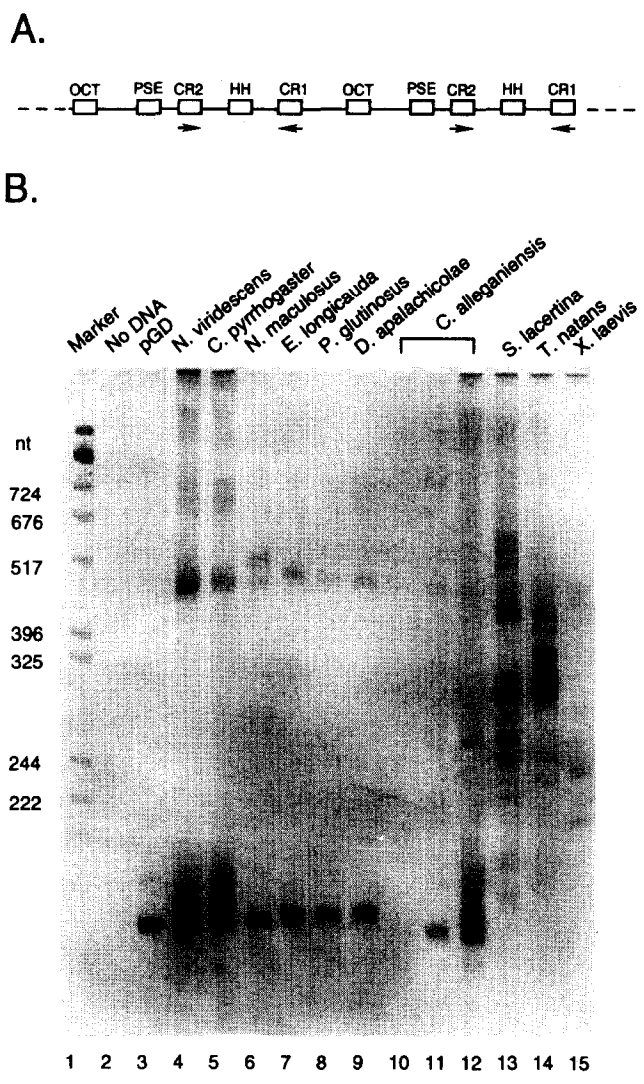

Fig. 2. PCR amplification of partial *sat2* sequences. (**A**) Strategy for PCR amplification. The diagram illustrates a segment of genomic DNA containing two tandem repeats of *sat2*. Labeled boxes indicate the relative locations of the conserved octamer (OCT) and proximal sequence element (PSE) of the transcriptional promoter (Coats et al., 1994), the extended HH domain, and blocks of conserved sequences with no known function (CR1 and CR2) (Green et al., 1993). The CR1 and CR2 oligo primers (denoted by arrows) were based on sequences of genomic *sat2* clones from *N. viridescens*, *A. talpoideum*, *A. tridactylum*, and *N. maculosus* (Epstein et al., 1986; Green et al., 1993; unpublished data), and have the partially degenerate sequences 5′-CAG TCN AGM CCG AAC TGC N-3′ (CR1) and 5′-GGG TAT GCC CAG ACN TGG GTC-3′ (CR2) (where M = A/C, N = A/C/G/T). Amplification with these primers results in an approx. 150-bp product that includes the extended HH domain. (**B**) The amplification products using genomic DNAs from the indicated species were electrophoresed on a 5% PA gel and stained with ethidium bromide. pGD is a cloned dimer of *sat2* from *N. viridescens* (Epstein and Pabón-Peña, 1991) and was used as a positive control template for amplification (lane 3). The size markers in lane 1 were generated by digestion of pGEM3Zf(−) DNA (Promega, Madison, WI) with *Ava*II, *Hin*fI, and *Rsa*I. PCR extension reactions were performed for 1 min at 72°C in 25-μl reactions consisting of 10 mM Tris·Cl (pH 8.3)/50 mM KCl/0.2 mM of each dNTP/3 mM $MgCl_2$/1 unit of *Taq* DNA polymerase/2 ng per μl (each) CR1 and CR2 oligo primers/10 pg (lanes 3–10), 1 ng (lane 11), or 100 ng (lanes 12–15) of template DNA. Prior to each extension reaction, samples were denatured at 94°C for 1 min and annealed for 1.5 min at 57°C (5 cycles), 54°C (10 cycles), and 51°C (20 cycles).

these organisms do not have *sat2* or their *sat2* repeats do not have a sequence complementary to one of the primers.

We cloned and sequenced the 150-bp PCR products to confirm that they were derived from *sat2* sequences (see section **b**). The clones were also used as probes in Southern blots to investigate the genomic organization of *sat2*. For most species that gave positive PCR results, digestion with *Nde*I or *Eco*O109I generated a ladder of hybridizing fragments that were monomers and multimers of a basic 300–350-bp unit (Fig. 3A, lanes 1–10). A similar but less intense ladder was evident in *C. alleganiensis* DNA after digestion with *Bam*HI (Fig. 3B, lanes 2 and 3). This pattern of hybridization was previously found for *sat2* in *N. viridescens*, *Ambystoma talpoideum*, and *Amphiuma tridactylum*, and indicates that the repeat units have slightly divergent sequences and are organized in tandem arrays (Epstein et al., 1986; Green et al., 1993).

Fig. 3B also shows blots of DNA from *S. lacertina*, *X. laevis*, and *T. natans* probed with the *C. alleganiensis* probe. No signals were detected even after considerable exposure times, and from these and the PCR results, we conclude that abundant *sat2*-related elements do not exist in these organisms.

(b) Cloning and sequence analysis of partial *sat2* elements

Amplification products containing the extended HH domains were cloned from the six newly identified *sat2*-containing species. Seventeen to 21 PCR-derived clones were sequenced and used to generate a consensus for each species. The clone information is summarized in Table I. The sizes of the regions between CR1 and CR2 were between 105 and 114 bp. Clones from *Cynops pyrrhogaster* were divided into two subtypes on the basis of covariations at 15 positions (Fig. 4). Distinct subtypes were not evident from the sequence data for any of the other species.

Fig. 4 shows the consensus sequences derived in this study compared to the previously determined consensus sequences from *N. viridescens*, *A. talpoideum*, *A. tridactylum*, and *Triturus vulgaris*. The sequences are organized into two groups on the basis of nucleotide similarities and similarities in the structures of the encoded HH motifs (see below). Group A includes *sat2* from the three Salamandridae species; group B includes *sat2* from the

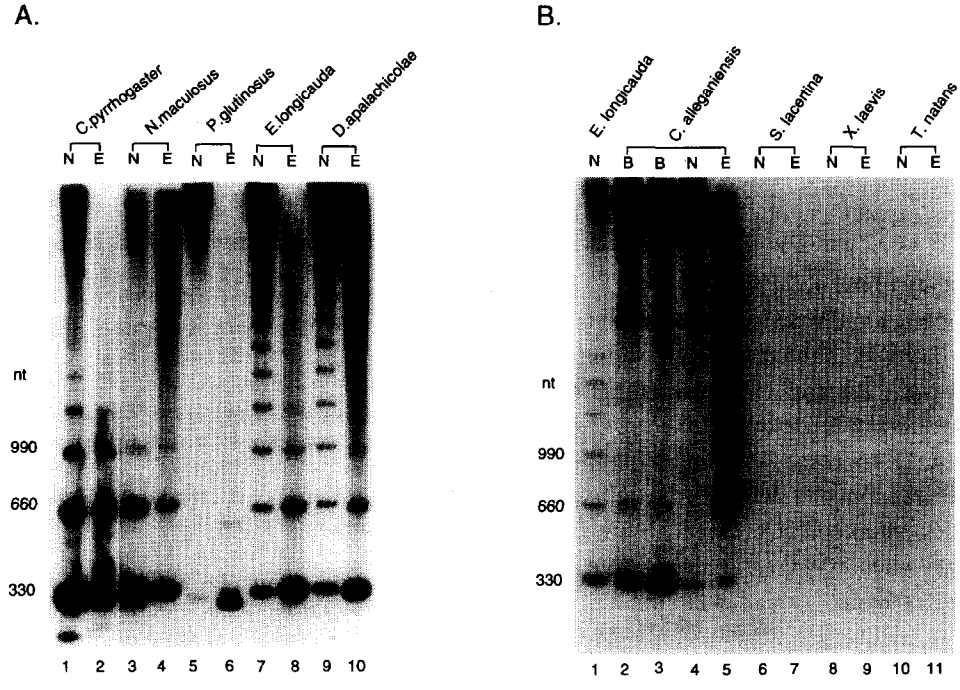

Fig. 3. Southern analysis of genomic *sat2* sequences. (**A**) Species with large quantities of *sat2*. Samples of genomic DNA (10 µg, each) from the indicated species were digested with *Nde*I (N) or *Eco*O109I (E) for 2 h. After electrophoresis on 1.5% agarose gels, the DNAs were blotted to a nitrocellulose filter and probed with radiolabeled synthetic transcripts as described previously (Epstein et al., 1986; Green et al., 1993). Clones chosen for probe preparation were those that had sequences most similar to the consensus of the species. The names of these clones were pCpyrB (lanes 1 and 2), pNmac 7 (lanes 3 and 4), pPglu 13 (lanes 5 and 6), pElon 9 (lanes 7 and 8), and pDapa 11 (lanes 9 and 10). The sequences of the HH domains in these clones are shown in Fig. 5. Exposure times were 2 h (lanes 1 and 2) and 10 h (lanes 3–10) at −80°C with an intensifying screen. (**B**) Species with few or no detectable *sat2* sequences. Samples of genomic DNA (10 µg, each) from the indicated species were digested as in panel **A** with the exception of the samples in lanes 2 and 3, which were digested with *Bam*HI (B) for 30 min (lane 2) or 2 h (lane 3). The *E. longicauda* DNA in lane 1 was included as a size marker. Samples were electrophoresed and blotted as in panel **A** and probed with transcripts derived from the *C. alleganiensis* clone pCall 5. Exposure was for 10 days at −80°C with an intensifying screen. The difference in the signal from *E. longicauda* DNA between panel **B** and panel **A** is due to the use of a heterologous probe in panel **B** and a homologous probe in panel **A**.

A. Group A Partial Satellite 2 Consensus Sequences

```
          1                                                                                                                  106
C.pyr1    ---------- tt-------- --------a- -g-------- -------t-a tg-a------ --------c- ----a-.--- ---------- ---------a ------
C.pyr2    --g------- c--------- ---------- -t-------- ---------- -a-------- ---------- ------.--- ----t----- -----g---c ------
N.vir     ---------- t-a------- ---------- -g-------- ----a----- ------t--- ---------- ------t--- ---------- ---------c ----c-
T.vul     --a-----t- cac------- --g------- -t-------- -a--t--c-- ---g--a--- ---------- ------.--- a----a---- ---a-----a ------
Consensus CCTTGCTCAC NGTGCCACTG GATTCAAGCT ANCCTGGCTG ATGAGGGGTG ATACCCCGAA ACCGGTCCTA GGATGC.TTG TTTCCGGTCC AGGGAAGACN TGGCTT
```

B. Group B Partial Satellite 2 Consensus Sequences

```
          1                                                                                                                  114
E.lon     ---------- .----tg--- ---------- ---------- ---g--.--- ---------- ------cg-- ------.ga- ------t--- ----t----- a-ag--g--- t-t.
P.glu     ---------- .--------- -----t---- ---g--ct-- ------.--- ---------- ------cg-- ------ataa ----a----- ---------- -a---ta--t tctt
A.tri     ---------- .--------- ---------- ---------- ------.t-- ----tg---- t-----gc-- t------..-- ----.gag-- ---ca----- ---------- --..
N.mac     ---a------ .--------- g--a------ -a-------- ---c--.t-- ----tg---- t-----gc-- ------..-- -----gt--- c-----a--g ----c---g- --..
A.tal     ---t--c--- .--g------ ----.----- -a-t------ ------.--- ---------- t-----at-- t------..a- -------g-- ------a--- ---------- --..
D.apa     ---------- ag-------- ----a----- -a-g--at-- ------t--- ---------- ------cg-- ---t--ac-g ct-gt----- ---------- -a--g-a--a -cag
C.all     --c------t c--------- g--ga-.--- ------cat- ---c--.--t ---gag--a- ------gc-- -at---.t-- ---ga-ac-t c-----aag- -attgtgtct ....
Consensus CCGGGCTTGC .TGTGCCTAG AGACTCAAGC TGCACCTCAC TGATGA.GGC CCAACAAGGC CGAAACNNGT CTGGGG.NTT GCTTGTGATC TCGTGCTGAA GGGATGTGAC CT..
```

Fig. 4. Partial *sat2* consensus sequences. The regions between the CR1 and CR2 primers are shown for the consensus of the partial *sat2* clones derived in this study. For comparison, the corresponding sequences from *N. viridescens* (Pabón-Peña et al., 1991), *A. tridactylum* (Green et al., 1993), *A. talpoideum* (Green et al., 1993), and *Triturus vulgaris* (Cremisi et al., 1992), are also shown. The organization of these sequences into groups A and B is based on criteria described in the text. Also shown is the overall consensus for each group. Dashes represent identity to the overall consensus, dots represent gaps, and N represents unspecified nt. The segment underlined in each group consensus corresponds to the extended HH domain. Alignments were generated by the Pileup program in the GCG Software Package (Genetics Computer Group, Madison, WI). Consensus sequences obtained in this study are available from GenBank under the following accession Nos.: *C. pyrrhogaster* type 1, U37315; *C. pyrrhogaster* type 2, U37317; *C. alleganiensis*, U37321; *D. apalachicolae*, U37320; *E. longicauda*, U37316; *N. maculosus*, U37319; *P. glutinosus*, U37318.

remaining species, which represent five other caudate families.

Potential HH structures for clones most similar to their species consensus are illustrated in Fig. 5. HH from group A sequences (in large box) are characterized by having 3-nt stem II loops, whereas group B HH have 6-nt stem II loops. Certain nt, including those in the central core region, are conserved in both groups (boxed nt), others are conserved in a group-specific manner (denoted by upper-case letters), and still others are specific to individual species (lower-case letters).

An inspection of the sequences and structures presented in Figs. 4 and 5 leads to the conclusion that the consensus sequences are representative of elements that are under selection. Twenty-three positions are conserved in the HH domains of each of the consensus sequences, including all 13 positions of the central core. The other conserved positions are involved in base-pairing interactions in stems I and II and occur in pairs, so the pairing partner of every conserved nt is also conserved. For comparison, only 68 of the 111 PCR clones have completely intact core regions, and less than half (49) have all 23 nt that are conserved in the consensus sequences (data not shown). Furthermore, mutations in the stems are typically not associated with compensatory mutations that restore base pairing. These results are consistent with our model of *sat2* evolution, which states that the bulk of the genomic repeats are inactive copies of a smaller set of active elements (Green et al., 1993). Once formed, the inactive elements diverge in sequence but maintain the sequence of the active element(s) as their consensus. These results also justify the use of clones most similar to the species consensus in the functional studies.

(c) Cleavage analysis of the newly cloned *sat2* extended HH

We assayed the self-cleavage activities of the HH shown in Fig. 5 using synthetic transcripts from the PCR clones. In addition to the extended HH domains that are illustrated, transcripts from the unmodified clones contained sequences corresponding to the CR1 and CR2 amplification primers as well as vector-derived sequences. To verify that transcripts with this configuration could fold into active structures, we analyzed transcripts from clone pGDHH for self-cleavage (Fig. 6A). pGDHH was derived by amplification of the HH domain from the *N. viridescens* clone pGD, which was previously shown to be active in self-cleavage (Epstein and Pabón-Peña, 1991). Transcripts from pGDHH therefore contain an active extended HH domain in the same configuration as in the other transcripts used in this analysis. Fig. 6A shows that these transcripts were capable of self-cleavage, and the cleavage kinetics were similar to the kinetics previously seen for *N. viridescens* transcripts with dimer, monomer, and partial monomer configurations (Epstein and Pabón-Peña, 1991; Pabón-Peña et al., 1991). Cleavages of transcripts from a *Necturus maculosus* clone derived in this study are shown in Fig. 6B and illustrate that we were successful in identifying active extended HH variants.

Not all of the newly investigated HH were as active as the *N. maculosus* HH. The HH representing *C. allega-*

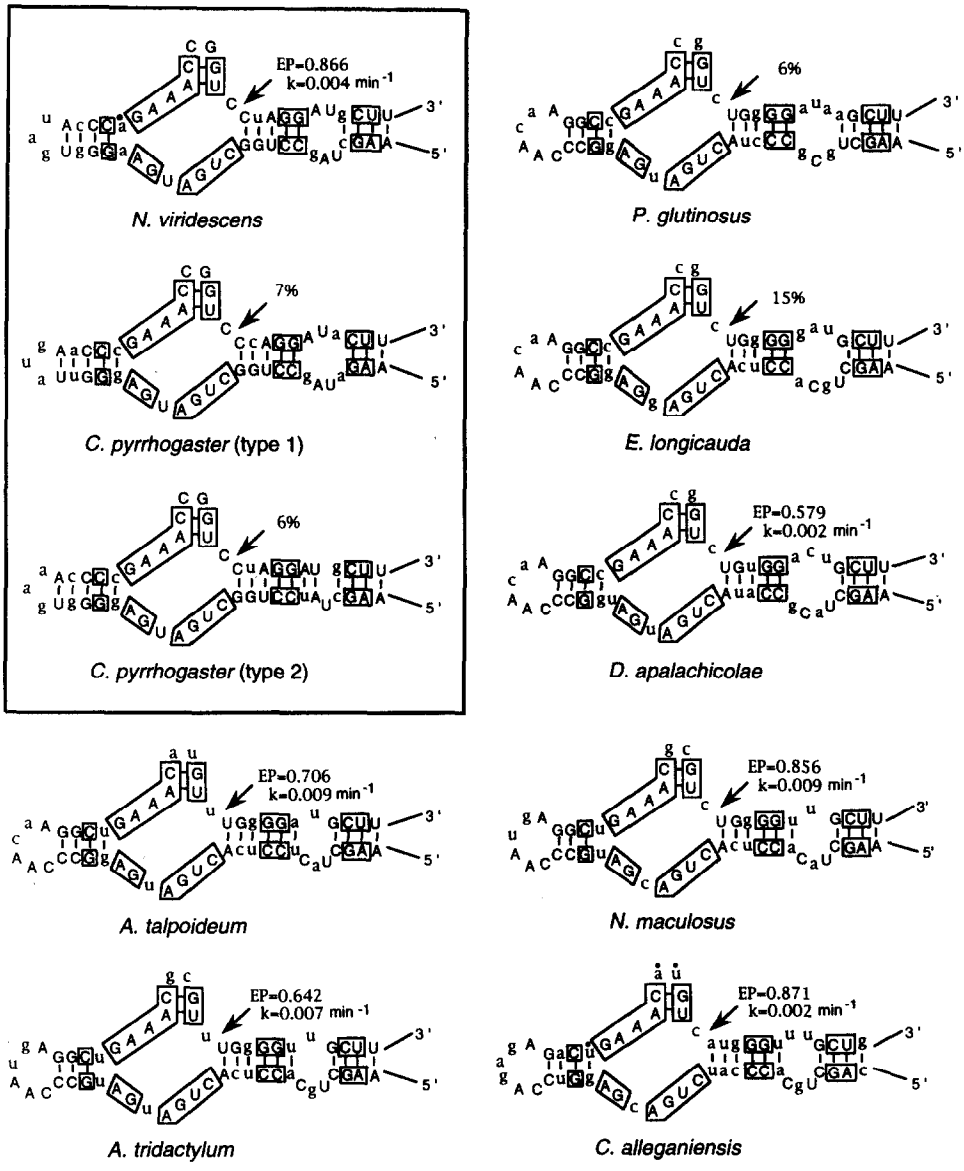

Fig. 5. Extended HH from nine caudate amphibians. For the species investigated in this study, the sequences shown correspond to the extended HH domains in the clones most similar to the species consensus. The names of these clones are given in the legend of Fig. 3. For previously investigated species, the sequences are from pG47 for *N. viridescens* (Pabón-Peña et al., 1991), pAtal-4 for *A. talpoideum* (Green et al., 1993), and pAtri-D for *A. tridactylum* (Green et al., 1993). Each nt in a clone that differs from the species consensus is marked by a dot. Group A HH are in the large rectangle; the remaining HH belong to group B. Boxed nt are conserved in all extended HH, upper-case letters in group A are conserved in each of the three representatives from this group, and upper-case letters in group B are conserved in all representatives of this group with the exception of *C. alleganiensis*, which differs at eight positions from all other members of this group. Synthetic transcripts from the original PCR clones contained flanking and vector-derived sequences in addition to the extended HH domains shown here. For *D. apalachicolae*, *C. pyrrhogaster*, *P. glutinosus*, and *E. longicauda*, modified clones that coded for transcripts that consisted mainly of the sequences shown here were used for the analysis of self-cleavage. The kinetic parameters k and EP were determined as described in Fig. 6 and are shown near the cleavage site for each HH. Because the HH representing *C. pyrrhogaster*, *P. glutinosus*, and *E. longicauda* cleaved too slowly for accurate determination of kinetic parameters, the values given are the percent cleavage after 10 h of incubation.

niensis and *Desmognathus apalachicolae* cleaved very slowly, and no cleavage was detected for the HH representing *C. pyrrhogaster*, *Plethodon glutinosus*, and *Eurycea longicauda* (data not shown). To minimize the detrimental effects of external sequences, we reconstructed the PCR clones from these species so that their transcripts would contain only the extended HH domains. As reported in Fig. 5, all of the reconstructed clones were capable of self-cleavage, although there was a great deal of variation in the actual rates of cleavage.

Because the differences between the extended HH were primarily in the peripheral stems and loops, the details of these structures were presumably responsible for the variations in the cleavage rates. The conservation of the sequences and structures of the internal stem p-I loop and the external stemp p-II loop also suggests that these

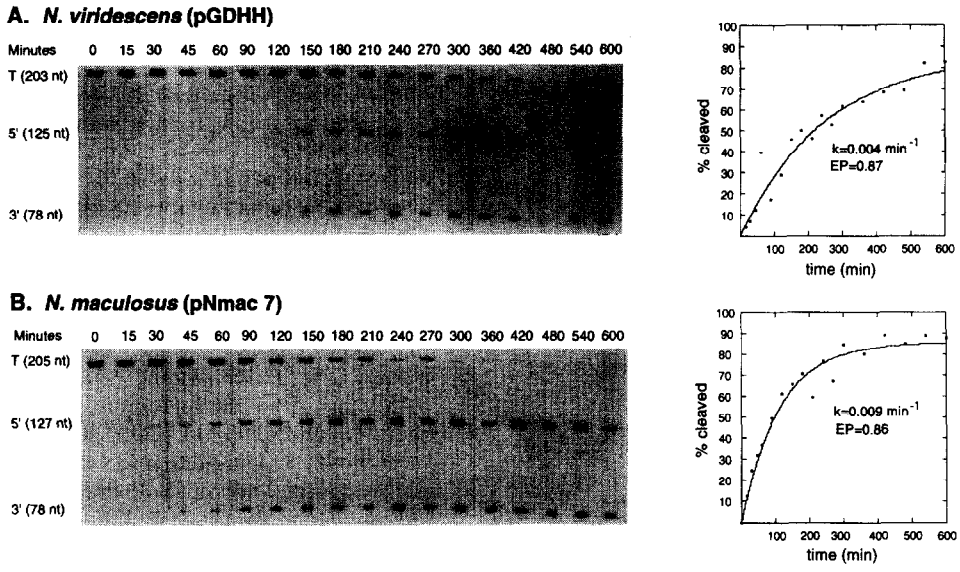

Fig. 6. Cleavage analysis of the *N. viridescens* and *N. maculosus* extended HH. Radiolabeled, synthetic transcripts of *N. viridescens* and *N. maculosus* clones were prepared with SP6 RNA polymerase as described previously (Pabón-Peña et al., 1991). Transcripts were heated at 80°C for 5 min in 1 mM EDTA, pH 8.0, quick-cooled in ice water, and combined with 2 vols. of prewarmed buffer to initiate self-cleavage. Self-cleavage reaction mixtures (6 µl) consisted of transcripts at 0.002–0.004 ng per µl/133 mM morpholineethanesulfonic acid (MES, pH 6.9)/30 mM $MgCl_2$/10 mM NaCl/0.3 mM EDTA. After incubation at 42°C for intervals up to 10 h, reactions were stopped by the addition of 6 µl of gel-loading buffer (80 mM EDTA/75% formamide/0.05% xylene cyanol/0.05% bromophenol blue). Full-length transcripts (T) and cleavage products (5′ and 3′) with the indicated sizes were separated by electrophoresis through 7 M urea 6% PA gels. Dried gels were quantified with a Betascope 603 Blot Analyzer (Betagen Corporation). The graphs on the right of each panel show the data fitted to the equation $F = EP(1 - \exp[-kt])$, where F is the fraction cleaved at time t, EP is the fraction cleaved at the end point of the reaction, and k is the first-order rate constant. Values for EP and k were determined from these plots and are shown adjacent to the graphs.

regions have important roles. However, the conserved sequence and structural elements may be important for functions, other then self-cleavage, that are known or believed to be associated with *sat2*. Furthermore, the variable sequences may affect cleavage rates by simply promoting alternative, inactive conformations of the transcripts as has been shown to occur for some variants of the prototypical HH (Fedor and Uhlenbeck, 1990; Heus et al., 1990). Further work is necessary to demonstrate that the peripheral stems have active roles during self-cleavage as indicated by our earlier studies with the *N. viridescens* extended HH (Pabón-Peña et al., 1991).

It was surprising that self-cleavage by most of the newly investigated extended HH was so inefficient. Active variants are apparently not as prevalent as our initial studies indicated (Epstein and Gall, 1987; Green et al., 1993), and methods other than the simple survey of existing salamanders are required to obtain a large pool of active variants for the analysis of the sequence and structural requirements in the extended HH. The information obtained in this study is useful in this regard in that it provides two extended HH designs that will serve as the starting point for future studies including the in vitro selection of active variants (Burke and Berzal-Herranz, 1993).

Finally, it is not known how the self-cleavage rates observed in vitro correlate with in vivo activity. The sub-optimal cleavage rates may be a direct reflection of in vivo activity, and cleavage rates may be varied to regulate some aspect of the function or evolution of *sat2*. Alternatively, by analogy to the dependence on maturase proteins by some self-splicing introns (Lambowitz and Belfort, 1993), cellular factors may be required to enhance or regulate self-cleavage in some of the salamander species. These issues will remain unresolved until systems are developed for examining *sat2* function in vivo.

(d) Conclusions

(*1*) Tandemly repeated *sat2* elements were detected in species representing four out of five families from the amphibian order Caudata. Because *sat2* was not found in *T. natans* or *X. laevis*, which represent the amphibian orders Gymnophiona and Anura, it appears that *sat2* arose after the separation of the three amphibian orders but before the separation of the majority of the caudate families.

(*2*) Sequence analyses of partial *sat2* clones indicated that, although mutations occur randomly in individual genomic repeats, core and structural elements of the extended HH domain are conserved in the consensuses of all species. These results support the idea that the consensuses are representative of active *sat2* elements and justify their use in the functional analyses.

(*3*) A comparison of the naturally occurring *sat2*

extended HH revealed the existence of two extended HH designs, each having distinct sequence and structural elements in its peripheral stem-loop structures. Although many of the natural variants were not efficient in self-cleavage, there were examples of active variants for each design. These designs will serve as models for in vitro selection experiments to investigate further the requirements in the two stems.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Institutes of Health and the American Cancer Society to L.M.E.

REFERENCES

Bruening, G.: Compilation of self-cleaving sequences from plant virus satellite RNAs and other sources. Methods Enzymol. 180 (1989) 546–558.

Burke, J.M. and Berzal-Herranz, A.: In vitro selection and evolution of RNA: applications for catalytic RNA, molecular recognition, and drug discovery. FASEB J. 7 (1993) 106–112.

Coats, S.R., Zhang, Y. and Epstein, L.M.: Transcription of satellite 2 DNA from the newt is driven by a snRNA type of promoter. Nucleic Acids Res. 22 (1994) 4697–4704.

Cremisi, F., Scarabino, D., Carluccio, M.A., Salvadori, P. and Barsacchi, G.: A newt ribozyme: a catalytic activity in search of a function. Proc. Natl. Acad. Sci. USA 89 (1992) 1651–1655.

Epstein, L.M. and Gall, J.G.: Self-cleaving transcripts of satellite DNA from the newt. Cell 48 (1987) 535–543.

Epstein, L.M. and Pabón-Peña, L.M.: Alternative modes of self-cleavage by newt satellite 2 transcripts. Nucleic Acids Res. 19 (1991) 1699–1705.

Epstein, L.M., Mahon, K.A. and Gall, J.G.: Transcription of a satellite DNA in the newt. J. Cell Biol. 103 (1986) 1137–1144.

Fedor, M.J. and Uhlenbeck, O.C.: Substrate sequence effects on 'hammerhead' RNA catalytic efficiency. Proc. Natl. Acad. Sci. USA 87 (1990) 1668–1672.

Forster, A.C. and Symons, R.H.: Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. Cell 49 (1987) 211–220.

Forster, A.C., Davies, C., Sheldon, C.C., Jeffries, A.C. and Symons, R.H.: Self-cleaving viroid and newt RNAs may only be active as dimers. Nature 334 (1988) 265–267.

Green, B.A., Pabón-Peña, L.M., Graham, T.A., Peach, S.E., Coats, S.R. and Epstein, L.M.: Conserved sequence and functional domains in satellite 2 from three families of salamanders. Mol. Biol. Evol. 10 (1993) 732–750.

Heidenreich, O., Pieken, W. and Eckstein, F.: Chemically modified RNA: approaches and applications. FASEB J. 7 (1993) 90–96.

Hendry, P., Moghaddam, M.J., McCall, M.J., Jennings, P.A., Ebel, S. and Brown, T.: Using linkers to investigate the spatial separation of the conserved nucleotides A9 and G12 in the hammerhead ribozyme. Biochim. Biophys. Acta 1219 (1994) 405–412.

Heus, H.A., Uhlenbeck, O.C. and Pardi, A.: Sequence-dependent structural variations of hammerhead RNA enzymes. Nucleic Acids Res. 18 (1990) 1103–1108.

Lambowitz, A.M. and Belfort, M.: Introns as mobile genetic elements. Annu. Rev. Biochem. 62 (1993) 587–622.

McCall, M.J., Hendry, P. and Jennings, P.A.: Minimal sequence requirements for ribozyme activity. Proc. Natl. Acad. Sci. USA 89 (1992) 5710–5714.

Pabón-Peña, L.M., Zhang, Y. and Epstein, L.M.: Newt satellite 2 transcripts self-cleave by using an extended hammerhead structure. Mol. Cell. Biol. 11 (1991) 6109–6115.

Pley, H.W., Flaherty, K.M. and McKay, D.B.: Three-dimensional structure of a hammerhead ribozyme. Nature 372 (1994) 68–74.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A.: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. 143 (1980) 161–178.

Scott, W.G., Finch, J.T. and Klug, A.: The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell 81 (1995) 991–1002.

Symons, R.H.: Small catalytic RNAs. Annu. Rev. Biochem. 61 (1992) 641–671.

Tuschl, T. and Eckstein, F.: Hammerhead ribozymes: importance of stem-loop II for activity. Proc. Natl. Acad. Sci. USA 90 (1993) 6991–6994.