MOLECULAR AND CELLULAR BIOLOGY, July 1989, p. 2828–2836
0270-7306/89/072828-09$02.00/0

Vol. 9, No. 7

Regulated Expression of Nuclear Protein(s) in Myogenic Cells That Binds to a Conserved 3′ Untranslated Region in Proα1(I) Collagen cDNA

THOMAS HERGET,† MATTHIAS BURBA, MARION SCHMOLL, KATRIN ZIMMERMANN, AND ANNA STARZINSKI-POWITZ*

Institut für Genetik der Universität zu Köln, Weyertal 121, D-5000 Cologne 41, Federal Republic of Germany

Received 30 November 1988/Accepted 29 March 1989

We describe the identification and DNA-binding properties of nuclear proteins from rat L6 myoblasts which recognize an interspecies conserved 3′ untranslated segment of proα1(I) collagen cDNA. Levels of the two proα1(I) collagen RNAs, present in L6 myoblasts, decreased drastically between 54 and 75 h after induction of myotube formation in serum-free medium. Both mRNAs contained a conserved sequence segment of 135 nucleotides (termed tame sequence) in the 3′ untranslated region that had 96% homology to the human and murine proα1(I) collagen genes. The cDNA of this tame sequence was specifically recognized by nuclear protein(s) from L6 myoblasts, as judged by gel retardation assays and DNase I footprints. The tame-binding protein(s) was able to recognize its target sequence on double-stranded DNA but bound also to the appropriate single-stranded oligonucleotide. Protein that bound to the tame sequence was undetectable in nuclear extracts of L6 myotubes that did not accumulate the two collagen mRNAs. Therefore, the activity of this nuclear protein seems to be linked to accumulation of the sequences that it recognizes in vitro. The collagen RNAs and the nuclear tame-binding proteins reappeared after a change of medium, which further suggests that the RNAs and the protein(s) are coordinately regulated.

It is generally assumed that interspecies conservation of nucleotide sequences reflects strong evolutionary selective pressure (15, 29). One type of conserved sequence element is confined to the 3′ untranslated regions (UTRs) of the mature mRNAs of a number of genes. Examples of highly conserved, gene-specific 3′ UTRs are found in the genes of the actins (15, 29, 30), c-*fos* (39), and proα1(I) collagen (2; this study). It is not clear why those 3′ UTRs contain highly conserved sequence segments of 150 nucleotides or more, since known regulatory functions 3′ UTRs are confined to much shorter sequence motifs (24, 34). Thus, 3′ UTRs have been found to control mRNA stability of, for example, c-*fos* (38), granulocyte-macrophage colony-stimulating factor (34), and tubulin (7) and appear to be required for the posttranscriptional iron-dependent regulation of the transferrin receptor gene (27). Transcription can also be regulated via sequences in the 3′ UTR, as has been shown for the β-actin gene (8). Furthermore, multiple mRNAs with identical protein-coding regions but varying 3′ UTRs are, for example, transcribed from the L-type pyruvate kinase (L-PK) gene (20). The different L-PK transcripts arise by use of different polyadenylation sites. The physiological role of the 3′-UTR variability, as well as sequence conservation, is, however, not known. One way of approaching the problem is by identifying DNA- or RNA-binding proteins that interact with 3′ UTRs on DNA or RNA.

In this investigation, we searched for nuclear protein(s) binding to conserved 3′-UTR DNA of genes that are expressed in myogenic rat L6 cells.

L6 myoblasts can be induced to form myotubes in serum-free medium containing insulin (16, 28). Morphological development from the mononucleated L6 myoblast to the multinucleated L6 myotube is accompanied by numerous alterations of gene activity (12). Muscle-specific genes are turned on in L6 cells, while certain nonmuscle genes are turned off. Housekeeping functions remain unaltered or may be modulated in degree of expression (12, 16). To study the regulatory role of interspecies conserved 3′ UTRs, we have isolated cDNA clones corresponding to mRNAs which are (i) regulated during L6 myotube formation and (ii) contain a conserved 3′-UTR sequence segment. cDNA clones containing sequences for the two proα1(I) collagen transcripts were chosen because these RNAs accumulate in L6 myoblasts, are down-regulated in L6 myotubes, and have an interspecies-conserved 3′-UTR segment (this study).

We show that this DNA element, termed tame sequence, is specifically recognized by nuclear protein(s) from L6 myoblasts. The tame-binding activity was absent from L6 myotubes in which proα1(I) collagen RNAs were also undetectable. However, these RNAs as well as the protein could be reinduced by a change of medium.

MATERIALS AND METHODS

Cells. Growth of myogenic rat L6/C15 cells in Waymouth medium with 15% fetal calf serum and initiation of myotube formation in serum-free Dulbecco modified Eagle medium containing bovine insulin (10 μg/ml) have been described elsewhere (16, 28). For preparation of cytoplasmic RNA, L6 myoblasts were used at approximately 60% confluency.

Preparation of RNA and cDNA libraries and isolation of cDNA clones. Cytoplasmic RNA from L6 cells and from primary rat myotubes was isolated as described previously (9, 16). Poly(A)$^+$ RNA was selected on oligo(dT)-cellulose (type II; Collaborative Research, Inc.; 1). cDNA production from developing L6 myotubes and cloning into plasmid vector pUR290 have been described elsewhere (16). The cDNA library prepared from primary rat myotubes (4) and cloned into lambda gt11 (3) was a gift from John Merlie.

* Corresponding author.
† Present address: Max Planck-Institut für Züchtungsforschung Egelspfad, D-5000 Cologne 30, Federal Republic of Germany.

cDNA pL6-3-1 was isolated from the L6 cDNA library by differential colony hybridization (11), using first-strand cDNA transcribed from RNA of L6 myoblasts and myotubes, respectively (16). The other cDNA clone, pmt14, was isolated from the lambda gt11 cDNA library which was made of RNA from primary rat myotubes. A synthetic oligonucleotide (20-mer; 5'-TTGAGAGATGAATGCAAAGG) derived from the sequence of clone pL6-3-1 was used for plaque hybridization (18).

Sequence analysis. cDNA pL6-3-1 isolated from the L6 cDNA library was sequenced by the method of Maxam and Gilbert (21). The insert fragment of cDNA clone pmt-14, isolated from the lambda gt11 cDNA library, was recloned into plasmid vector pUC19 and sequenced by the double-stranded dideoxy-chain termination method (32; Boehringer Mannheim Biochemicals manual). Nucleotide sequences were determined on both strands of the cDNAs.

RNA blotting, radionucleotide labeling of DNA probes, and hybridization to RNA. RNA samples were size fractionated by electrophoresis on 1% agarose–6% formaldehyde gels (37), transferred to ZetaBind nylon membranes, and dried at 80°C for 2 h. DNA was labeled with ^{32}P by random hexanucleotide priming (10) to a specific activity of 10^9 cpm/μg of DNA. Hybridization of DNA probes to RNA blots was done essentially under standard conditions (18) except for the fragment of cDNA Hf677 (6). This probe was hybridized (39°C, 50% formamide) and washed (2× SSC [SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 0.1% sodium dodecyl sulfate 50°C for 30 min) at reduced stringency. Usually, membranes were washed at a final stringency of 0.1% SSC–0.1% sodium dodecyl sulfate at 60°C for 30 min. Rehybridization with probe pAG82 was always done after the complete hybridization procedure with pL6-3-1 was finished and after a first exposure to X-ray film. For rehybridization with probe pAG82, the Northern (RNA) blots were not dehybridized.

Preparation of nuclear extracts, gel retardation assays, and DNase I footprint analysis. The experimental protocols described below are modified from published procedures (22, 26). Fifteen monolayers (15-cm diameter) of L6 myoblasts at 75% confluency were washed twice with phosphate-buffered saline and once with HB buffer (see below), followed by addition of 5 ml of HB buffer with 0.5% Nonidet P-40. Cells were lysed for 5 min on ice and harvested; nuclei were pelleted at 2,000 × *g* for 5 min, suspended in PB buffer (see below) to 10^8 nuclei per ml, and adjusted to 0.4 M KCl. Nuclei were extracted for 30 min on ice and centrifuged at 15,000 × *g* for 30 min. Proteins in the supernatant were precipitated with 45% ammonium sulfate, suspended in BB buffer (see below), and dialyzed against BB buffer for 3 h. Portions of 25 μl with a final protein concentration of 2 to 4 mg/ml were stored at −70°C. Nuclear extracts from myotubes were prepared essentially as described above except that 1% Nonidet P-40 was used and cell lysis was allowed to proceed for 10 to 15 min, until 80 to 90% of the nuclei were free (microscope check).

HB buffer consisted of 10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES)-NaOH (pH 8), 0.5 mM spermidine, 0.15 mM spermine, 1 mM EDTA, 0.25 mM ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid (EGTA), 50 mM NaCl, 2 mM dithiothreitol, 0.5 M sucrose, and 1 mM phenylmethylsulfonyl fluoride. PB buffer was the same as HB buffer except that sucrose was replaced by 10% glycerol (vol/vol). BB buffer consisted of 10 mM HEPES (pH 8), 100 mM NaCl, 0.1 mM EDTA (pH 8), 2 mM dithiothreitol, 25% (vol/vol) glycerol, and 1 mM phenylmethylsulfonyl fluoride. Buffer for binding assays (see below) was the same as BB buffer except that glycerol was replaced by 4% (wt/vol) Ficoll (Pharmacia Fine Chemicals).

For the gel retardation assays, 0.5 ng of ^{32}P-end-labeled insert fragment (fill-in reaction) of the indicated cDNA clones was added to 0.5 to 4 μg of nuclear protein in assay buffer containing 1 μg of poly(dI-dC) DNA (Pharmacia; final assay volume, 25 μl) and incubated at 25°C for 30 min. The samples were then separated by electrophoresis on a 4% polyacrylamide gel in 0.5× TBE (18) at maximal 6 V/cm for 6 h at 4°C. For retardation assays with oligonucleotides labeled by T4 polynucleotide kinase (18), binding was allowed to proceed in 50 mM instead of 100 mM NaCl. Oligonucleotide-protein complexes were run on 5% polyacrylamide gels in 6.7 mM Tris hydrochloride (pH 7.5)–3.3 mM sodium acetate–1 mM sodium EDTA (35) at 4 V/cm for 3.5 h with constant circulation of the running buffer. The gels were dried on 3MM paper (Whatman, Inc.) and exposed to Kodak XAR-5 films (Eastman Kodak Co.) at −70°C.

For the footprint analysis, DNA-protein complexes (5 ng of DNA; 20 to 60 μg of protein) were allowed to form at 25°C for 30 min and were then treated with 50 ng of DNase I for 3 min in a volume of 250 μl at 25°C. The reaction was stopped by addition of EDTA and storage on ice. The proteins were extracted from the samples with phenol-chloroform (1:1, vol/vol), and the remaining DNA was analyzed on a DNA-sequencing gel.

RESULTS

cDNA clone pL6-3-1 was detected in a cDNA library prepared from RNA of L6 cells collected 72 h after induction of myotube formation (16). The library was screened differentially with first-strand cDNA transcribed from RNA of L6 myoblasts and L6 myotubes, respectively. pL6-3-1 (0.165 kilobase pairs [kbp] hybridized with myoblast-derived but not with myotube-derived first-strand cDNA. cDNA pmt-14 (1.2 kbp) was isolated from a lambda gt11 cDNA library made from RNA of primary rat myotubes (3, 4; see Materials and Methods). Primary rat myotube cultures prepared as described previously (4) still contained proα1(I) collagen mRNAs (unpublished observation). pL6-3-1 and pmt-14 were further characterized in Northern blot experiments and by nucleotide sequence analysis as described below.

Sequence analysis and comparisons, identification of the collagen RNAs, and definition of the tame element. The nucleotide sequences of cDNAs pL6-3-1 (165 bp; Fig. 1b) and pmt-14 (Fig. 1c) were determined and compared with sequences listed in the EMBL Data Library. Both cDNAs contained sequences that were nearly identical to sequences of the proα1(I) collagen RNAs from humans (2, 5) and mice (23; Fig. 1c). This homology covered C-terminal amino acid codons as well as about 230 nucleotides downstream of the stop codon and was taken as an indication that cDNAs pL6-3-1 and pmt-14 contained rat proα1(I) collagen sequences. Frequently, 3' UTRs are the most specific probes for individual mRNAs and have, for example, been used to distinguish the different actin isoforms (29).

The proα1(I) collagen RNAs were expected to have sizes of about 5 and 6 kb, as inferred from the lengths of the human transcripts (5). The two RNAs of about 5 and 6 kb, detected in Northern blot experiments by cDNA pL6-3-1 (Fig. 1a and 2), met this expectation. An oligonucleotide derived from the coding part of cDNA pmt-14 as well as a fragment from the protein-coding region of the human proα1(I) collagen cDNA Hf677 (6) were also hybridized to

FIG. 1. (a) Northern blot analysis with RNA from L6 cells. Probes were an oligonucleotide derived from the protein-coding region of pmt-14 (5'-TAGGTCTTCAAGCAAGAGGACCAAGCT TCC-3'; lane 1), cDNA pL6-3-1 containing the tame sequence (lane 2), and the protein-coding region of the human proα1(I) collagen cDNA clone Hf677 (6) (lane 3; the 3' UTR was cut off at the *Eco*RI site); positions of the 5- and 6-kb RNAs are indicated. (b) Double-stranded nucleotide sequence of cDNA pL6-3-1. The boxed nucleotides indicate the sequences that were protected in the DNase I protection assay (see also Fig. 6). (c) Comparison of rat cDNA pmt-14 sequences (1) with murine (2) and human (3) proα1(I) collagen sequences. The boxed sequence is the 135-nucleotide-long tame sequence. The beginning of the pL6-3-1 sequence relative to the pmt-14 sequence is indicated. (d) Schematic representation of the tame sequence in the proα1(I) collagen gene and sequences of the two 5- and 6-kb transcripts, as inferred from both the human (5) and mouse (23) genes. Note that in both species there is no intron in the 3' UTR.

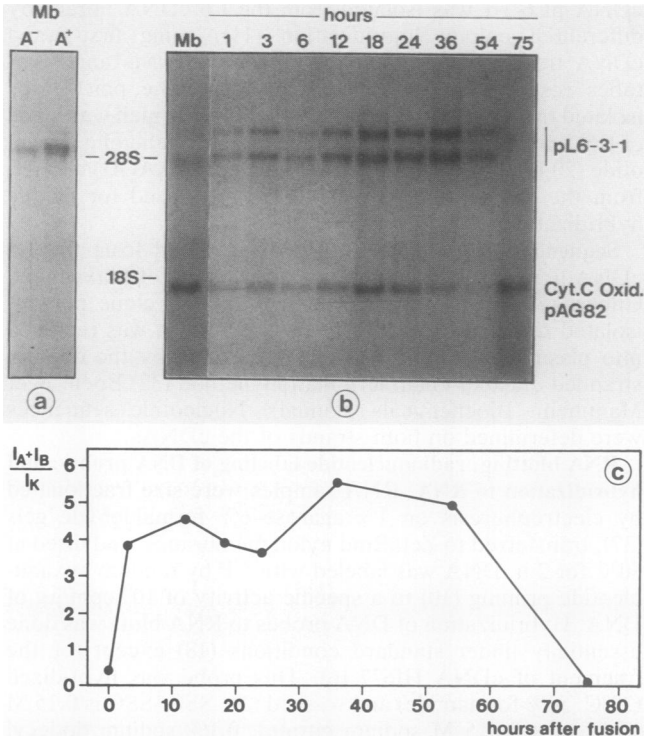

FIG. 2. Northern blot analysis with RNA from L6 cells. (a) Hybridization of DNA pL6-3-1 to 3 μg of poly(A)$^-$ and 3 μg of poly (A)$^+$ RNA from L6 myoblasts. The 6-kb RNA was primarily detected in the poly(A)$^+$ fraction; the 5-kb RNA was detected mainly in the poly(A)$^-$ fraction but was also found in the poly(A)$^+$ fraction. (b) Northern blot analysis with probe pL6-3-1 of poly(A$^+$)-enriched cytoplasmic RNA (5 μg per lane) from L6 cells at 0 (lane Mb [myoblasts]), 1, 3, 6, 12, 18, 24, 36, 54, and 75 h after induction of myotube formation. Two bands, of 5 and approximately 6 kb, were found up to 54 h after induction of myotube formation. After hybridization to pL6-3-1, the blot was rehybridized with probe pAG82, which contains cDNA sequences of the mouse cytochrome *c* oxidase (25, 33), and exposed to Kodak XAR-5 film at −70°C for 3 days. (c) Quantitation of the proα1(I) collagen mRNAs relative to the cytochrome *c* oxidase control by densitometry (model 620 video densitometer; Bio-Rad Laboratories). I_A and I_B are the densities of the collagen mRNAs; I_K is the density of the cytochrome *c* oxidase mRNA. The relative amounts of the collagen mRNAs were calculated by integrating the peaks of (I_A + I_B) and I_K.

cytoplasmic RNA of L6 cells (Fig. 1a). These two probes detected the same 5- and 6-kb RNAs. We conclude from these Northern blot results and from the nucleotide sequence analysis that cDNAs pL6-3-1 and pmt-14 were derived from proα1(I) collagen mRNAs.

For the gel retardation and DNase I protection experiments described below, we used the 3'-UTR segment that was most highly conserved among rat, human, and murine

proα1(I) collagen RNAs. This was a segment of 135 nucleotides with 96% homology in the rat (represented in cDNAs pL6-3-1 [Fig. 1b] and pmt-14 [Fig. 1c]), human, and mouse (Fig. 1c) sequences. Henceforth, for reasons of convenience, we call this conserved 3′-UTR segment the tame sequence. Its location in the collagen gene and in the transcripts is schematically shown in Fig. 1d.

Down-regulation of the the two collagen RNAs during L6 myotube formation. Poly(A)$^+$-enriched, cytoplasmic RNA from L6 cells was prepared at 1, 3, 6, 12, 18, 24, 36, 54, and 75 h after induction of myotube formation and from L6 myoblasts. Northern blots were performed and probed with ^{32}P-labeled cDNA pL6-3-1 (Fig. 2b). Two collagen RNAs of 5 and about 6 kb were detected up to 54 h after induction of myotube formation and had decreased at approximately 75 h. Subsequent hybridization of this Northern blot with cDNA pAG82 containing coding sequences of the mouse cytochrome *c* oxidase subunit 1 (25, 33) showed that L6 cells, although they were kept in serum-free medium, contained this mRNA, which is essential for normal cell function. This finding indicated that the decrease of the two collagen RNAs identified with cDNA pL6-3-1 was not related to nutritional deficiencies of myotube-forming L6 cells. Hybridization to pAG82 also revealed the integrity of the RNA samples tested.

Furthermore, quantitation of the collagen mRNAs relative to the cytochrome *c* oxidase control by densitometry (Fig. 2c) showed that the level of the proα1(I) collagen mRNAs was rather constant between 0 and 54 h of myotube development and then appeared to decrease sharply between 54 and 75 h after induction of myotube formation. This drastic decrease was reproducible in L6 cells that had ≥90% of nuclei in myotubes. L6 cells that were less differentiated had higher levels of collagen mRNAs (unpublished observations). Most likely, this variability in collagen expression was the reason why cDNA pL6-3-1 was isolated from myotube-forming L6 cells. Percentage of fusion was checked in sister plates at 96 h after induction of myotube formation, when development of L6 myotubes in culture was complete. Because of the observations described above, RNA and nuclear extracts were subsequently used from L6 myotube cultures that had ≥90% of nuclei in myotubes. The data shown in Fig. 2b and 1a are also in agreement with previous analyses (12) showing that collagen proteins are present in L6 myoblasts and become undetectable during L6 myotube formation.

Comparison of the hybridization patterns of poly(A)$^-$- and poly(A)$^+$-enriched cytoplasmic L6 RNAs to pL6-3-1 (Fig. 2a) showed that a substantial if not major fraction of the 5-kb RNA was present in the poly(A)$^-$ fraction as well as the poly(A)$^+$ fraction. The 6-kb RNA, however, was primarily found in the polyadenylated form (Fig. 2a).

Recognition of the tame sequence by regulated nuclear proteins from L6 cells. Nuclear extracts were prepared from L6 myoblasts. A 0.5-ng sample of ^{32}P-end-labeled pL6-3-1 insert DNA, representing the tame sequence, was added to 4 μg of nuclear protein; complexes were allowed to form at 25°C for 30 min and then subjected to electrophoresis on 4% polyacrylamide gels. These gel retardation experiments (Fig. 3, lanes 2 and 3) showed that nuclear factors from L6 myoblasts bound to the tame sequence. Binding of these myoblast factors could be inhibited by specific competitor DNA containing the tame sequence (Fig. 3, lanes 4 to 6) but not by unrelated DNA fragments such as pUC18 DNA (Fig. 3, lane 19). pUR290 plasmid polylinker itself, which flanks

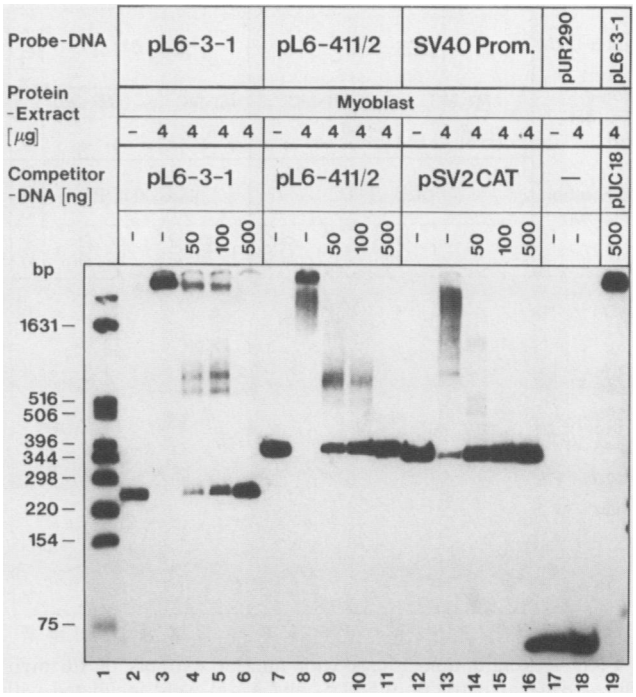

FIG. 3. Retardation assays with nuclear extracts from L6 myoblasts, using the 250-bp insert fragment of cDNA pL6-3-1 (tame plus linker sequences of pUR290; lanes 2 to 6 and 19), the 350-bp insert fragment of cDNA pL6-411/2, containing the TGGCA-binding site (lanes 7 to 11), the 345-bp *Hin*dIII-*Pst*I simian virus 40 promoter fragment of plasmid pSV2CAT (lanes 12 to 16), and the *Eco*RI-*Hin*dIII linker fragment of plasmid pUR290 (lanes 17 and 18). Binding of nuclear L6 proteins to the pL6-3-1 fragment was inhibited by specific competitor DNA (lanes 4 to 6) but not by unrelated, linearized pUC18 DNA (lane 19). Molar excesses of competitor DNA over probe DNA (0.5 ng; 1,000 cpm) were 5-, 10-, and 50-fold for pL6-3-1 (lanes 4 to 6), 10-, 20-, and 100-fold for pL6-411/2 (lanes 9 to 11), and 8-, 16-, and 80-fold for pSV2CAT (lanes 14 to 16). The linker fragment of pUR290 was not retarded (lane 18). Size marker was pBR322 DNA digested with *Hin*fI (lane 1). The gel was exposed to Kodak XAR-5 film for 12 h.

the isolated pL6-3-1 cDNA insert, was not recognized by nuclear L6 proteins (Fig. 3, lanes 17 and 18).

Positive-control experiments were performed with the TGGCA-binding site (also called nuclear factor 1-binding site; 17) which is present on cDNA pL6-411 subclone 2 (pL6-411/2; 16) and does not compete with pL6-3-1 (data not shown). Nuclear extracts also worked with these recognition sequences (Fig. 3, lanes 7 to 11). Previous studies have shown that L6 muscle cells contain high amounts of TG GCA-binding protein (T. Herget et al., unpublished observations). A second positive-control experiment was done with the *Eco*RI-*Pst*I fragment from plasmid pSV2-CAT, containing the simian virus 40 promoter-enhancer element (14; Fig. 3, lanes 12 to 16). This fragment was also shifted in the gel retardation experiments.

Treatment of the DNA-protein complexes with either proteinase K or RNase A showed that protein factors were responsible for the reduced electrophoretic mobility of the tame-positive DNA fragment (data not shown).

Tame-DNA-binding activity and proα1(I) collagen RNAs are undetectable in L6 myotubes but can be coordinately reinduced. Nuclear proteins were also prepared from myotubes at 80 h (Fig. 4) and at 100 h (data not shown) of

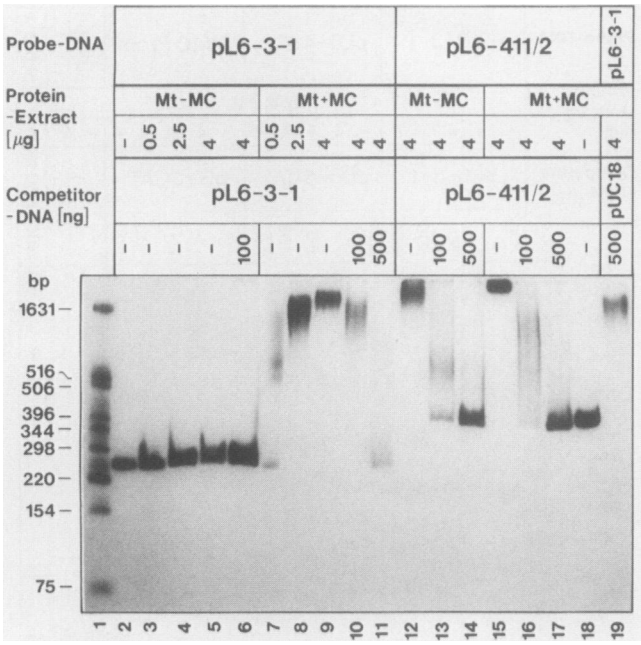

FIG. 4. Retardation assays with nuclear extracts of L6 myotubes. Nuclear extracts (0.5, 2.5, and 4 μg) were incubated with DNA fragments from pL6-3-1 and pL6-411/2 (see legend to Fig. 3). Nuclear extracts were prepared from L6 myotubes at 80 h after induction of myotube formation in serum-free medium without medium change (Mt−MC) or with medium change without serum at 80 h and preparation of extracts at 100 h (Mt+MC) of myotube development. Both nuclear protein extracts retarded insert DNA from pL6-411/2 (lanes 12 to 17). Insert DNA pL6-3-1 bound only to proteins of the Mt+MC extract (lane 7 to 11) and not to proteins of the Mt-MC extract (lane 2 to 6).

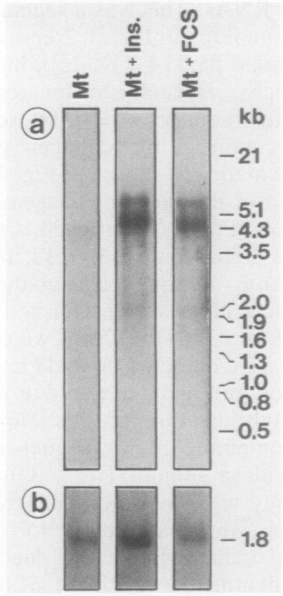

FIG. 5. Northern blot analysis of total cytoplasmic RNA (10 μg per lane) from L6 myotubes at 100 h after induction of myotube formation in serum-free medium without medium change (Mt; equivalent to Mt−MC in Fig. 4) and with change to medium containing insulin (Mt+ins.; equivalent to Mt+MC in Fig. 4) or fetal calf serum (Mt+FCS) at 80 h of development. (a) Hybridization of the blot to ^{32}P-labeled cDNA pL6-3-1; (b) rehybridization of the same blot to cDNA pAG82 of the mouse cytochrome *c* oxidase subunit 1.

development in serum-free medium. The protein in these extracts had no detectable affinity to the tame sequence when assayed in gel retardation assays. However, when the medium of the myotube cultures was changed to fresh medium without serum, nuclear proteins recognizing the tame sequence were again detectable (Fig. 4, lanes 7 to 11). Control experiments with the nuclear myotube extracts demonstrated that proteins from all of these preparations retained the fragment with the TGGCA-binding site (Fig. 4, lanes 12 to 18).

In parallel, some dishes of the L6 cell preparations described above were used for isolation of cytoplasmic RNA in order to correlate the accumulation of the proα1(I) collagen RNAs with the activity of the tame-DNA-recognizing proteins. The RNA samples were analyzed in Northern blots with ^{32}P-labeled cDNA pL6-3-1 as the probe. Reinduction of the tame-recognizing activity was accompanied by reaccumulation of the proα1(I) collagen RNAs containing the tame sequence (Fig. 5a). Control hybridization with plasmid pAG82 showed that the cytochrome *c* oxidase mRNA was detected in all three RNA samples equally well (Fig. 5b). This and the result of Fig. 1 demonstrated that the reduction of collagen RNAs and tame-binding activity was not due to nutritional deficiencies of the L6 cells. Thus, the coordinate accumulation and decrease of collagen RNAs and tame-binding activity strongly indicated a functional relationship between the two activities (see Discussion).

Determination of sequences recognized by the tame-binding protein(s). DNase I protection experiments were performed to locate the binding sites of the nuclear L6 proteins within the tame sequence. The results (Fig. 6) indicated the protection of four sequence motifs, A1, A2, B1, and B2, within the tame element from DNase I digestion. The protected nucleotides are marked in the pL6-3-1 sequence (Fig. 1b). There are two footprints on each strand. Remarkably, the corresponding regions on the complementary strands are not protected. The protection below B1 (Fib. 6b) was not reproducible in different assays and thus was regarded as nonspecific.

We compared the protected regions with published binding sites. Homology was found in two cases. (i) In the protected motifs A1 and A2, six of seven and five of seven nucleotides, respectively, were identical to the M-CAT heptamer 5′-CATTCCT-3′, which is required for skeletal muscle-specific activity of the cardiac troponin T gene promoter (19). (ii) The sequence 5′-TCTCCCA-3′ (motif B1 in Fig. 1 and 6) was identified as a partial complementary sequence of a TFIIIA recognition site present in the internal control region of the *Xenopus* 5S rRNA gene (31). However, it is not clear whether the homology of these recognition sites reflects any functional similarity of the binding proteins.

Tame-DNA-binding protein(s) also recognizes the sequence on single-stranded DNA. Since we found protection of only one of the tame-DNA strands in the DNase I footprint experiments (Fig. 6), we tested whether the tame-binding protein would also recognize single-stranded DNA. For this purpose, we synthesized oligonucleotide A (5′-CATTCATC TCTCCATTCAACCTTAC), which was derived from the windows A1 and A2 protected on the sense strand in the footprint (Fib. 1b and Table 1). This oligonucleotide was used in gel retardation assays with nuclear proteins from L6 myoblasts. In these experiments (Fig. 7), assay conditions

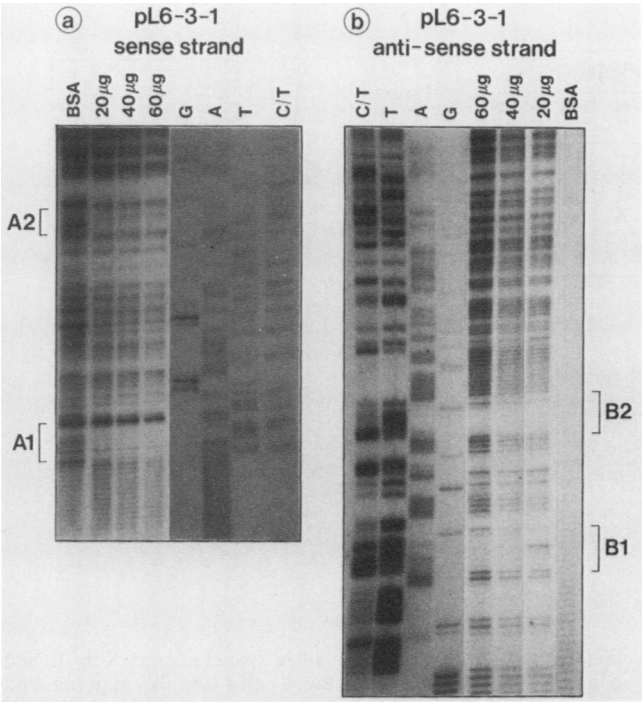

FIG. 6. Footprint analyses. Nuclear extracts of L6 myoblasts (20, 40, and 60 μg) or 20 μg of bovine serum albumin were incubated with 10 ng (10^5 cpm) of pL6-3-1 insert fragment labeled at the sense (a) or antisense (b) strand. The protein-DNA complexes were treated with 50 ng of DNase I at 25°C for 3 min. The protected nucleotides were localized in comparison with the sequence ladder obtained from sequence determination of either strand by the method of Maxam and Gilbert (21). The DNA fragment shown in panel a corresponds to the RNA; that in panel b is complementary to the RNA. The protected nucleotides are shown in Fig. 1.

were different from those for the experiments shown in Fig. 3 and 4, since protein binding to oligonucleotides was good in acetate buffer but very weak in TBE, which seems to be better for resolution of complexed fragments. Therefore, the competition experiments, in which oligonucleotides and plasmid fragment were used together, were done under conditions optimized for protein binding to oligonucleotides.

The single-stranded oligonucleotide A was retained by nuclear factor(s) from L6 myoblasts (Fig. 7a). In contrast, oligonucleotide A′, which was complementary to A, did not bind to nuclear L6 protein(s) (Fig. 7a). Competition of binding to oligonucleotide A could be achieved both with oligonucleotide A and with the double-stranded tame-DNA fragment. In agreement with these results, oligonucleotide A, but not A′, could compete with the double-stranded tame-DNA fragment for the binding proteins (Fig. 7b). Thus, the tame-recognizing factor(s) was able to find its target on double-stranded as well as on single-stranded DNA.

On the basis of the results shown in Fig. 7a and b, we determined whether windows B1 and B2 were also specifically recognized on single-stranded DNA. Nuclear extracts from L6 myoblasts were incubated with ^{32}P-labeled oligonucleotides B1, B2, B1′, and B2′, the latter two being complementary to B1 and B2. Oligonucleotide B1 was weakly and oligonucleotide B2 was strongly shifted by protein(s) from L6 myoblasts (Fig. 7c, lanes 8 and 10). Sequences complementary to B1 and B2 were not recognized by this protein(s) (Fig. 7c, lanes 9 and 11). This result supports the idea that footprints B1 and B2 were specific. In contrast to the results for oligonucleotide A, containing windows A1 and A2 (Fig. 7c, lane 2), retardation assays with oligonucleotides B1 and B2 (Fig. 7c, lanes 8 and 10) did not give rise to two distinct bands.

To obtain further evidence that both windows A1 and A2 could be recognized by tame-binding protein(s), we mutated these sites simultaneously [oligonucleotide A(m1, m2)] or at either site separately (oligonucleotides A(m1) and A(m2); for details, see Table 1]. Mutation of either site, A1 or A2, still allowed retardation of the oligonucleotides (Fig. 7c, lanes 6 and 7). However, the upper of the two bands seen with oligonucleotide A (Fig. 7c, lane 2) disappeared in both cases. Mutation of both windows completely abolished binding (Fig. 7c, lane 5). These results support the idea that each of the windows, A1 and A2, can be independently occupied by tame-binding protein(s).

Finally, it was shown that retardation of oligonucleotide A was very much reduced with nuclear extract from L6 cells at 80 h of myotube development (Fig. 7c, lane 3). This observation is in agreement with the idea that the same tame-binding protein(s) binds on double-stranded as well as on single-stranded DNA.

DISCUSSION

In this report, we show that a conserved 3′-UTR segment of the proα1(I) collagen gene, called tame DNA, is recog-

TABLE 1. Design of synthetic tame oligonucleotides derived from footprints in DNase I protection analyses

Oligonucleotide	Window	Sequence of window (5′→3′)[a]	Sequence of oligonucleotides (5′→3′)[b]
A	A1	CATTCATCTCTC	CATTCATCTCTCCATTCAACCTTAC
	A2	CATTCAACCTTAC	
A(m1,m2)	A1 and A2 mutated		GAAAGAGGTCTCGAAAGAGGCTTAC
A(m1)	A1 mutated		GAAAGAGGTCTCCATTCAACCTTAC
A(m2)	A2 mutated		CATTCATCTCTCGAAAGAGGCTTAC
A′	Complementary to A		GTAAGGTTGAATGGAGAGATGAATG
B1	B1	TCTCCCA	TCTCCCAATTTTTGGC[c]
B1′	Complementary to B1		GCCAAAAATTGGGAGA
B2	B2	GTTTTCCA	GTTTTCCAGTTTTCCAGTTTTCCA[d]
B2′	Complementary to B2		TGGAAAACTGGAAAACTGGAAAAC

[a] See Fig. 1b.
[b] Mutated oligonucleotides are underlined.
[c] Window B1 plus downstream sequence.
[d] Three times window B2.

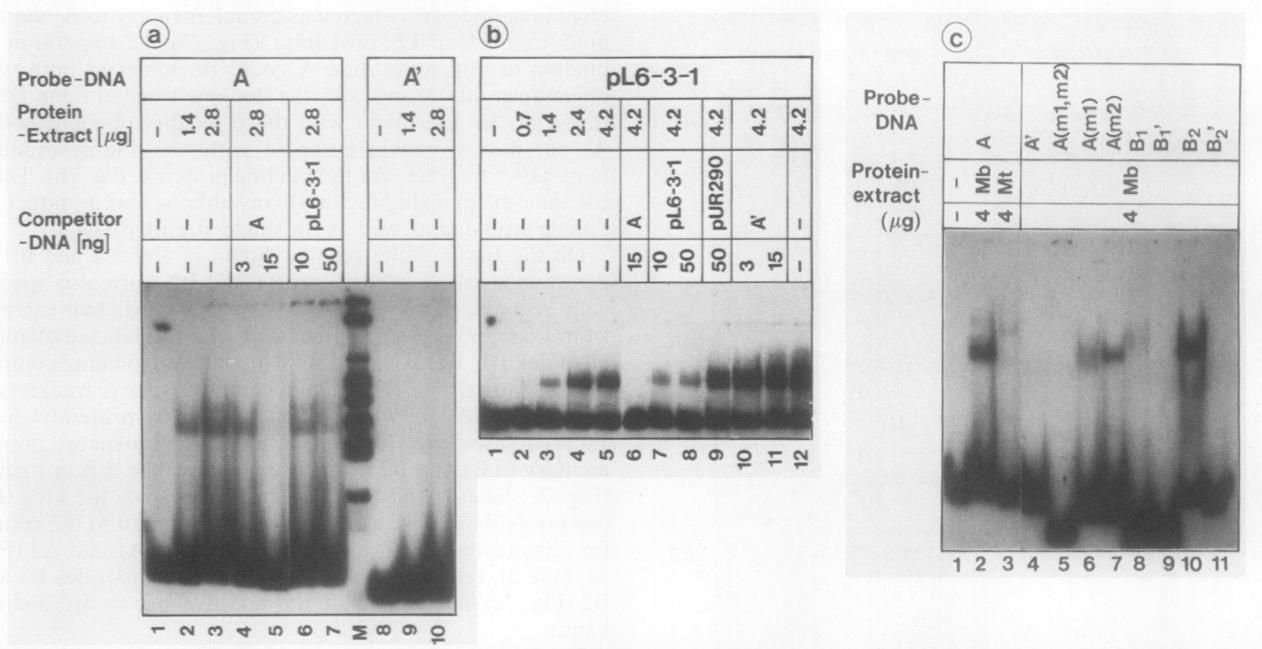

FIG. 7. Gel retardation assays with ^{32}P-labeled oligonucleotides A (see below; panel a, lanes 1 to 7) and A′ (panel a, lanes 8 to 10) and pL6-3-1 insert fragment (b). Buffer conditions were different from those used in the experiments shown in Fig. 3 and 4 (see also Materials and Methods and Results). Nuclear extracts were prepared from L6 myoblasts. Competitor DNA was added to the reaction mixtures as indicated. Lane M in panel a is a ^{32}P-labeled DNA size marker (Bethesda Research Laboratories ladder). (c) Confirmation of the DNase I footprints (see Fig. 6) by gel retardation assay with ^{32}P-labeled oligonucleotide probes. Synthetic oligonucleotide sequences were derived from the protected windows A1 and A2 (A; lane 1 to 3), B1 (lane 8), and B2 (lane 10). A′ (lane 4), B1′ (lane 9), and B2′ (lane 11) are oligonucleotides complementary to A, B1, and B2, respectively. Furthermore, oligonucleotide A was mutated in different ways. In A(m1,m2) (lane 5), both A1 and A2 were mutated; in A(m1) (lane 6), window A1 was mutated; and in A(m2) (lane 7), window A2 was mutated. Details of the synthetic oligonucleotides are given in Table 1. Nuclear extracts were prepared from L6 myoblasts (Mb; lanes 2 and 4 to 11) and from L6 myotubes (Mt; lane 2) at 80 h after induction of differentiation. DNA-protein complexes were formed in the presence of 1,500 ng of poly(dI-dC) DNA. Exposure to Kodak XAR-5 film was at −70°C for 40 h.

nized by regulated nuclear protein(s) from myogenic L6 cells. Both the tame-DNA-recognizing nuclear proteins and the proα1(I) collagen RNAs were readily detectable in L6 myoblasts but not in L6 myotubes. This observation suggests that the collagen RNAs and the binding proteins are coregulated. The finding that the tame-binding activity was recoverable from the nucleus may indicate that it is involved in regulation of proα1(I) collagen expression by direct interaction with the tame sequence at the DNA level. The two bands seen in the gel retardation experiments (Fig. 3 and 7) could indicate that two identical protein molecules interact with two target sites on one DNA molecule. The fact that in vitro the protein(s) showed protection of one strand in double-stranded tame DNA (Fig. 6) and shifted the appropriate single-stranded oligonucleotides (Fig. 7) raises the possibility that the tame-recognizing factor(s) bind to RNA as well as to DNA. Such a property has been already reported for transcription factor TFIIIa (13, 36). It will be important to identify the nature of the tame-binding target in vivo.

Mechanistically, it is difficult to imagine a protein binding to double-stranded DNA and protecting only one of the strands and not its direct complementary sequence against DNase I attack (Fig. 6). Perhaps the tame-binding protein(s) has the intrinsic property of recognizing its target in double-stranded DNA and separating the two strands in order to bind to a single-stranded molecule. No special DNA topology would be required for this activity, since linear DNA was used in the DNase I footprint analyses.

It has been reported that two proα1(I) collagen RNAs of about 5 and 6 kb exist in human (5), mouse (23), and rat (Fig. 1 and 2) cells. As shown in human (5) and mouse (23) cells, these two RNAs differ in the lengths of their 3′ UTRs, a result explained by the differential use of two polyadenylation signals that are about 1 kbp apart (23). It is interesting to note that the protected regions A1 and A2 on the sense strand of the tame sequence (Fig. 1 and 6) are located close to the polyadenylation signal of the smaller transcript, which was determined in both mouse (23) and human (5) cells. Thus, the tame-binding protein(s) is possibly involved in the decision as to whether an individual transcript is going to be 5 or 6 kb long by directing the polyadenylation machinery to either the proximal or distal site. The physiological importance of this decision may be reflected by the fact that the differential use of the collagen polyadenylation sites is maintained during evolution.

Furthermore, the differential use of polyadenylation sites has been also reported for other gene products, such as the L-PK transcripts (20). In this case, it seems that the ratio of the different L-PK mRNAs which are specified by the length of the 3′ UTR is even regulated in a developmentally and tissue-specific manner (20). Nevertheless, the functional significance of these multiple L-PK transcripts remains unclear, since they have similar half-lives and appear to be translatable with comparable efficiencies (20).

It was also demonstrated that a complete medium change at 80 h of L6 myotube development led to reappearance of the tame-binding protein(s) as well as of the collagen RNAs. Nutritional problems were unlikely be the reason for this phenomenon. It seems conceivable that L6 cells secrete and

accumulate a factor in the supernatant that is able to repress expression of the proα1(I) collagen RNAs and tame-recognizing activity, for instance by binding to the L6 cell surface (autocrine regulation). Removal of such a factor by a change of medium would then result in reexpression of the collagen RNAs and tame-binding proteins(s) (Fig. 4 and 5).

Perhaps the most important result to emerge from this work is the coordinate accumulation and disappearance of the collagen RNAs and the tame-DNA-binding activity. Clearly, the binding protein(s) must be inactivated either by degradation or by modification. At present, we are attempting to purify sufficient amounts of the protein to raise antibodies with which these questions can be addressed.

ACKNOWLEDGMENTS

We thank A. Nowock, R. Jack, and B. von Wilcken-Bergmann for helpful discussions and critical reading of the manuscript, B. Müller-Hill for encouragement, J. Merlie for providing his lambda gt11 cDNA library from primary rat myotubes, and M.-L. Chu for cDNA clone Hf677.

This research was supported by grants from the Deutsche Forschungsgemeinschaft (Sta 187/5-1,2) and the Bundesministerium für Forschung und Technologie. Fellowships of the Boehringer Ingelheim Fonds to T.H., of the DECHEMA to K.Z., and of the Deutsche Forschungsgemeinschaft (Sta 187/6-1) to A.S.-P. are gratefully acknowledged.

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Bernard, M. P., M.-L. Chu, J. C. Myers, F. Ramirez, E. F. Eikenberry, and D. J. Prockop.** 1983. Nucleotide sequences of complementary deoxyribonucleic acids for the proα1 chain of human type I procollagen. Statistical evaluation of structures that are conserved during evolution. Biochemistry **22**:5213–5223.
3. **Buonanno, A., J. Mudd, V. Shah, and J. P. Merlie.** 1986. A universal oligonucleotide probe for acetylcholine receptor genes. J. Biol. Chem. **261**:16451–16458.
4. **Carlin, B. E., J. C. Lawrence, Jr., J. M. Lindstrom, and J. P. Merlie.** 1986. Inhibition of acetylcholine receptor assembly by activity in primary cultures of embryonic rat muscle cells. J. Biol. Chem. **261**:5180–5186.
5. **Chu, M.-L., W. de Wet, M. Bernard, and F. Ramirez.** 1985. Fine structural analysis of the human proα1(I) collagen gene. J. Biol. Chem. **260**:2315–2320.
6. **Chu, M.-L., J. C. Myers, M. P. Bernard, J.-F. Ding, and F. Ramirez.** 1982. Cloning and characterization of five overlapping cDNAs specific for the human proα1(I) collagen chain. Nucleic Acids Res. **10**:5925–5934.
7. **Cleveland, D. W.** 1988. Autoregulated instability of tubulin mRNAs: a novel eukaryotic regulatory mechanism. Trends Biochem. Sci. **13**:339–343.
8. **De Ponti-Zilli, L., A. Seiler-Tuyns, and B. M. Paterson.** 1988. A 40-base-pair sequence in the 3′ end of the β′ actin gene regulates β-actin mRNA transcription during myogenesis. Proc. Natl. Acad. Sci. USA **85**:1389–1393.
9. **Favoloro, J., R. Freisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
10. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
11. **Garfinkel, L. I., M. Periasamy, and B. Nadal-Ginard.** 1982. Cloning and characterization of cDNA sequences corresponding to myosin light chains 1, 2 and 3, troponin-C, troponin-T, α-tropomyosin, and α-actin. J. Biol. Chem. **257**:11078–11086.
12. **Garrels, J. I.** 1979. Changes in protein synthesis during myogenesis in a clonal cell line. Dev. Biol. **73**:134–152.
13. **Ginsberg, A. M., B. O. King, and R. G. Roeder.** 1984. Xenopus 5S gene transcription factor, TFIIIA: characterization of a cDNA clone and measurement of RNA levels throughout development. Cell **39**:479–489.
14. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
15. **Gunning, P., T. Mohun, S.-Y. Ng, P. Ponte, and L. Kedes.** 1984. Evolution of the human sarcomeric actin genes: evidence for unit of selection within the 3′ untranslated regions of the mRNAs. J. Mol. Evol. **20**:202–214.
16. **Herget, T., M. Reich, K. Stüber, and A. Starzinski-Powitz.** 1986. Regulated expression of repetitive sequences including the identifier sequence during myotube formation in culture. EMBO J. **5**:659–664.
17. **Leegwater, P. A. J., P. C. van der Vliet, R. A. W. Rupp, J. Nowock, and A. E. Sippel.** 1986. Functional homology between the sequence-specific DNA-binding proteins nuclear factor 1 from HeLa cells and the TGGCA protein from chicken liver. EMBO J. **5**:381–386.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mar, J. H., and C. P. Ordahl.** 1988. A conserved CATTCCT motif is required for skeletal muscle-specific activity of the cardiac troponin T gene promoter. Proc. Natl. Acad. Sci. USA **85**:6404–6408.
20. **Marie, J., M.-P. Simon, Y.-C. Lone, M. Cognet, and A. Kahn.** 1987. Tissue-specific heterogeneity of the 3′ untranslated region of L-type pyruvate kinase mRNAs. Eur. J. Biochem. **158**:33–41.
21. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–525.
22. **Miksicek, R., U. Borgmeyer, and J. Nowock.** 1987. Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus. EMBO J. **6**:1355–1360.
23. **Mooslehner, K., and K. Harbers.** 1988. Two mRNAs of mouse proα1(I) collagen gene differ in the size of the 3′-untranslated region. Nucleic Acids Res. **16**:773.
24. **Müllner, E. W., and L. C. Kühn.** 1988. A stem loop in the 3′ untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell **53**:815–825.
25. **Murphy, D., P. M. Brickell, D. S. Latchman, K. Willison, and P. W. J. Rigby.** 1983. Transcripts regulated during normal embryonic development and oncogenic transformation share a repetitive element. Cell **35**:865–871.
26. **Nowock, J., and A. E. Sippel.** 1982. Specific protein-DNA interaction at four sites flanking the chicken lysozyme gene. Cell **30**:607–615.
27. **Owen, D., and L. C. Kuehn.** 1987. Noncoding 3′ sequences of the transferring receptor gene are required for mRNA regulation by iron. EMBO J. **6**:1287–1293.
28. **Pinset, C., and R. G. Whalen.** 1985. Induction of myogenic differentiation in serum-free medium does not require DNA synthesis. Dev. Biol. **108**:284–289.
29. **Ponte, P., P. Gunning, H. Blau, and L. Kedes.** 1982. Human actin genes are single copy for α-skeletal and α-cardiac actin but multicopy for β- and γ-cytoskeletal genes: 3′ untranslated regions are isotype specific but are conserved in evolution. Mol. Cell. Biol. **3**:1783–1791.
30. **Ponte, P., S. Ng, P. Engel, P. Gunning, and L. Kedes.** 1984. Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucleic Acids Res. **12**:1687–1696.
31. **Sakonju, S., and D. D. Brown.** 1982. Contact points between a positive transcription factor and the Xenopus 5S RNA gene. Cell **31**:395–405.
32. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. J. Biol. Chem. **74**:5463–5467.

33. **Scott, M. R. D., K.-H. Westphal, and P. W. J. Rigby.** 1983. Activation of mouse genes in transformed cells. Cell **34**:557–567.
34. **Shaw, G., and R. Kamen.** 1986. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**:659–667.
35. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
36. **Smith, D. R., I. J. Jackson, and D. Brown.** 1984. Domains of the positive transcription factor specific for the Xenopus 5S RNA gene. Cell **37**:645–652.
37. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
38. **Treisman, R.** 1985. Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5′ element and c-fos 3′ sequences. Cell **42**:889–902.
39. **Verma, I. M., J. Deschamps, C. Van Beveren, and P. Sassone-Corsi.** 1986. Human c-fos. Cold Spring Harbor Symp. Quant. Biol. **51**:949–958.