3218–3225 Nucleic Acids Research, 1994, Vol. 22, No. 15

Production and cleavage of *Drosophila hsp70* transcripts extending beyond the polyadenylation site

Shelley L.Berger* and Matthew Meselson
Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138, USA

Received December 17, 1993; Revised May 12, 1994; Accepted June 7, 1994 GenBank accession no. U03466

ABSTRACT

Transcription downstream of the polyadenylation site was studied in the *Drosophila hsp70* gene, whose high level of transcription in response to temperature elevation facilitates detection of rare and possibly short-lived transcripts. Transcription downstream of the polyadenylation site was detected both in cultured cells and in intact animals. Even shortly after temperature elevation the extended non-polyadenylated RNAs were rare relative to mature message, and their level continued to increase following temperature elevation even after the amount of mature message stopped increasing. The extended transcripts therefore are unlikely to be message precursors. Although continuous transcripts were detected extending as far as 2 kb downstream of the normal polyadenylation site, the predominant extended transcript was 0.45 kb long, apparently produced by cleavage of longer transcripts. Its amount relative to mature message increased with the duration and severity of heat-shock. As is the case in nonpolyadenylated histone mRNA, there is a potential stem-loop structure just upstream of the cleavage site. These data and other lines of evidence suggest that this extended transcript results from an alternative mode of stable 3′-end formation.

INTRODUCTION

Studies of 3′ end formation of pol II transcripts have shown that the consensus signal AATAAA, approximately 30 bp upstream of the polyadenylation site, and a downstream sequence rich in the dinucleotide GT are both required for the normal formation and polyadenylation of 3′ ends (for reviews, see references 1, 2). Mutations of either of these sequences give rise to non-polyadenylated transcripts that extend heterogeneously well beyond the normal polyadenylation site (3, 4). A striking example of this is seen when either sequence is deleted from a hybrid SV40 gene cloned in a circular plasmid, causing transcription to continue repeatedly around the circle (3).

Even when the polyadenylation signal and the downstream element are intact, however, transcription of sequences far beyond the polyadenylation site has been consistently noted in run-on studies in isolated nuclei (1). In the mouse β-globin, mouse α-amylase and other genes these regions tend to be far downstream (0.3 to 4.0 kb) of the ultimate processed 3′ ends, and primary ends are heterogeneous and occur over a large region (for review, see 5). Whether such extended transcription occurs in intact cells and, if so, whether it occurs normally or instead results only from occasional read-through events is still not well-established.

Two mechanisms have been proposed for transcriptional termination (6). In the first, 3′ end formation is dependent on the polyadenylation process itself, as suggested by the effect of mutations in polyadenylation sites described above (3). Cleavage of the elongating RNA at the site of polyadenylation would expose the uncapped 5′ end to exonucleolytic degradation, ultimately causing release of the polymerase complex (3). A related effect is possible modification of the transcription complex after transiting the polyadenylation site, causing destabilization of the complex and termination (6, 7, 8). In either case termination may be facilitated by pause sites found downstream of the polyadenylation site (8).

A different proposal is that termination is independent of the polyadenylation process. For example, a sequence found downstream of the mouse gastrin gene causes termination in the absence of polyadenylation (9).

We have examined newly synthesized transcripts of an polII gene in intact cells, focusing on the pattern of transcription in the 3′ region. In order to obtain detectable levels of possible short-lived RNA species, we chose a gene that is transcribed at a particularly high rate, the *Drosophila hsp70* heat shock gene.

Transcripts containing sequences downstream of the polyadenylation site of the *hsp70* gene are clearly evident within 2 minutes after temperature elevation. However, even at early times the amount of downstream RNA is very small compared to the amount of RNA transcribed from the adjacent region upstream of the polyadenylation site. Unless the downstream transcripts are extremely short-lived, persisting for less than about 10 seconds, this indicates that in normal intact cells *hsp70* transcription only occasionally extends much beyond the normal 3′ end.

The rare downstream transcripts are not polyadenylated and extend continuously as far as 2 kb past the normal polyadenylation site. The most prominent of them is a discrete species 450 nt longer than mature transcript that probably arises by cleavage

*To whom correspondence should be addressed at: Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

of transcripts extending farther downstream. Its accumulation and that of even longer transcripts is especially enhanced relative to normal message, after prolonged or severe heat shock. Under these conditions, there is apparently a failure of normal 3′ processing, allowing accumulation of transcripts that extend beyond the normal cleavage site.

Although most transcripts are polyadenylated at the expected position, the stable non-polyadenylated transcript extending 450 nt beyond the normal polyadenylation site must result from an alternative mechanism of 3′-end formation. This alternative processing pathway could serve to stabilize mRNA 3′ ends when normal 3′ end formation fails.

MATERIALS AND METHODS

Materials

Cultured *Drosophila* cells derived from Oregon R flys [Schneider line 2 (S2); 10], were grown in M3 medium supplemented with 10% fetal calf serum at 25°C (11). Fly stocks were wild-type Oregon R. All probes were derived from pPW223 containing Oregon R genomic DNA cloned from the heat shock locus at chromosome sub-division 87A (12,13). A *SalI*−*Bam*HI (2.5 kbp) fragment of pPW223 was subcloned from pPW223 into pSP72 (Promega Biotec). This fragment was restriction mapped and various fragments, designated mat.1, mat.2, term.1 and term.2 were subcloned into pSP72 or pBS$^+$ (Stratigene), as depicted in Figure 1.

RNA isolation

S2 cells (approximately 1×10^7) or Oregon R flies (approximately 50 adults) were heat shocked at indicated temperatures and times and were rapidly cooled in a −10°C salt water bath. The cells were centrifuged at 4°C and lysed by addition of 2 ml Holmes−Bonner solution [7M urea, 2% SDS, 0.35M NaCl, 1mM EDTA, 10 mM Tris−HCl, pH 7.4 (14)]. For the experiment depicted in Figure 9, cells were centrifuged and resuspended in 1 ml of medium before heat shock and were lysed immediately afterwards by the addition of Holmes−Bonner solution. Adult flies were disrupted with the aid of a Dounce homogenizer and extracted 2−3 times with 1:1 phenol:chloroform (PC). Nucleic acids were precipitated with 2.5 volumes of ethanol and dissolved in TES (10 mM Tris−HCl, pH 7.4, 1 mM EDTA, 0.2% SDS). RNA was precipitated by addition of sodium acetate (pH 5.3) to 3 M. After centrifugation, the RNA pellet was washed with 3 M sodium acetate and 70% ethanol and stored at −20°C. Polyadenylated RNA was prepared by oligo-dT chromatography (15).

Northern blot analysis

RNA was denatured, ethidium bromide was added, and the samples were electrophoresed on 0.8% formaldehyde gels, essentially as described (15), using RNA size markers (BRL; 0.15−9.5 kb ladder). Gels were blotted overnight to nylon membranes (ICN transblot) and crosslinked to the membrane by UV irradiation (600 μW/cm^2 for 5 min). Single-stranded continuously labeled RNA probes were prepared by *in vitro* transcription of plasmid DNAs using Sp6 or T7 polymerase (16). Hybridizations were carried out in 50% formamide, 10% PEG-6000, 7% SDS, 0.125 M sodium phosphate (pH 6.5), 0.25 M NaCl and 1 mM EDTA at the following temperatures, which were optimized for each probe: 37°C for mat.1 and 55°C for term.1 and term.2. The blots were washed 3 times each in 2×SSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate), 5 mM sodium phosphate (pH 6.5), 0.02% sodium pyrophosphate, and 1% SDS followed by 3 washes in the same solution except 0.1×SSC. Wash temperatures were 50°C for mat.1 or 60°C for term.1 and term.2. Hybridized probes were removed from the membrane by a 2 hour wash at 80°C in 75% formamide before rehybridization with other probes. Autoradiographic exposures (Kodak XAR film) varied from 1 to 40 hours at −70°C using intensifying screens.

RNA probe protection analysis

For RNA probe protection, probes were synthesized on linearized DNA templates and purified by gel electrophoresis. 2.5−25 μg of total cellular RNA (except where otherwise indicated) was hybridized to 6−300,000 cpm antisense RNA probe in 30 μl of 40 mM PIPES, pH 6.4, 1 mM EDTA, 0.4 M NaCl) at 37°C overnight. Probes were always in excess during hybridization, as demonstrated by doubling the amount of cellular RNA in control experiments. The hybrids were treated with RNAase A (40 μg/ml) and RNAase T1 (600 u/ml; BRL) in 0.35 ml buffer (10 mM Tris−HCl, pH 7.5, 5 mM EDTA, 300 mM NaCl) for 30 min at 25°C. Digestion was stopped by addition of 10 μl 20% SDS and 10 μl 10 mg/ml proteinase K. Proteinase K digestion (37°C for 15 min) was followed by phenol/chloroform extraction, and ethanol precipitation. The nucleic acids were resuspended in 2 μl of 80% formamide, 0.1% each bromophenol blue and xylene cyanol, and 5 mM EDTA) and separated by electrophoresis on a 6% polyacrylamide−urea gel. The gel was dried and exposed to autoradiographic film for various times at −70°C using intensifying screens. End-labeled *Hpa*II fragments of pBR322 were used as size markers.

Construction of Bal31 deletions

The 2.5kb *SalI*−*Bam*HI subclone of pPW223 was digested to completion with *SalI* and digested with Bal31 nuclease (Boehringer Mannheim) removing aliquots every five minutes for 25 minutes. Reactions were stopped by adding EGTA to 10 mM. The deleted ends were blunted using T4 DNA polymerase and were ligated to *Pst*I linkers. Miniprep DNAs were digested with *Pst*I and *Bgl*II to estimate the extent of each deletion. Promising deletion DNAs were digested with *Bgl*II, electrophoresed on low melting point agarose (Seakem) and the large fragment was excised from the gel and self-ligated. The exact endpoints of the deletions were determined by Maxam−Gilbert sequencing (17).

Sequencing of term.1

*Pst*I to *Bgl*II fragments of pSP72-term.1 and the Bal31 deletion derivatives were cloned into the *E.coli* phage vectors mp18 and mp19 (18). The dideoxynucleotide chain termination method (18) was used to sequence each insert. In one direction deletion derivatives permitted accurate sequencing through the insert. In the other direction a complementary oligomer (+360 to +373 nt relative to the *Aha*III site) was synthesized to complete the sequence. The free energy of formation of the stems of the structures shown in Fig. 9 were calculated according to Tinoco *et al.* (19).

The GenBank accession number is U03466.

S1 nuclease protection analysis

pSP72-term.1 was digested with *Fok*I and the 3′ end was labeled using Klenow polymerase. The plasmid was then digested with

*Bgl*II and the fragments were electrophoresed on a 4% polyacrylimide gel. The 393 *Fok*I−*Bgl*II fragment was excised from the gel and purified. Hybridization and S1 nuclease digestion were done as described (17) and the products were electrophoresed on an 6% polyacrylimide gel.

RESULTS

Transcripts found shortly after induction

In order to detect possible intermediates involved in RNA 3′ end-formation, we analyzed transcripts of the *Drosophila hsp70* gene shortly after heat-shock. The detection of such processing intermediates, if they exist, should be favored by this heat-shock genes's rapid inducibility and very high transcription rate.

Figure 1 shows the 3′ region of the *hsp70* gene and the anti-sense RNA probes used for Northern and probe protection analyses. The downstream end of the 243 nt anti-sense RNA probe mat.1 lies at an *Aha*III site 15 nt beyond the polyadenylation signal, therefore detecting only mature mRNA. The 472 nt anti-sense RNA probe mat.2 extends to a *Sac*I site 244 nt farther downstream and is useful for detecting transcripts that extend across the polyadenylation site. The anti-sense RNA probes term.1 (*Aha*III−*Bgl*II) and term.2 (*Bgl*II−*Eco*RI) only detect transcripts downstream of the polyadenylation site.

Cultured cells propagated at 23°C were kept at 37°C for various times, rapidly cooled in an ice-salt bath, centrifuged at 4°C and lysed. RNA was prepared by phenol extraction, salting out and ethanol precipitation. A Northern blot showing the successive hybridization of probes mat.1, term.1 and term.2 to total cellular RNA isolated at 0, 5, 10, 15, 20 and 30 min after temperature elevation is shown in Fig. 2. Panel A shows that 2.8 kb mature polyadenylated *hsp70* message, is seen by 5 min (lane 2) and then accumulates rapidly, reaching an approximately constant level at 20 min (lane 5). Hybridization to mat.1 gives no clear evidence of transcripts longer than mature message at any of the times examined.

In order to achieve more sensitive detection of longer transcripts without interference from mature mRNA, the membrane in Figure 2 panel A was washed, probed with term.1, and exposed for a prolonged period. Panel B reveals a discrete 3 kb RNA evident within 10 min after temperature elevation (lane 3) that accumulates until 30 min (lane 6), the latest time shown. Even larger molecules are evident at late times (lane 6). Prominent among these is an RNA of approximately 4.5 kb, revealed better by hybridization to term.2 (panel C, lane 6). The same species of *hsp70* RNA and the same kinetics of accumulation following temperature elevation were observed in RNA extracted from adult *Drosophila* (not shown).

The 3′ ends of mature transcripts and of the transcripts giving rise to the band at 3 kb were mapped by RNA probe protection with mat.2 and term.1, as depicted in Fig. 3. Protection of probe mat.2 (panel A) reveals three major fragments of 200−230 nt (indicated by circles), the size range expected for protection by the 3′ region of mature *hsp*70 message. The presence of multiple species may be due to alternative polyadenylation sites, since there are several regions rich in AT nucleotides at the predicted position upstream of the cleavage site (see Figure 6). Mature message accumulates until about 20 min (lane 5) and then remains approximately constant in amount until at least 60 min (Figure 3, lane 6, and data not shown).

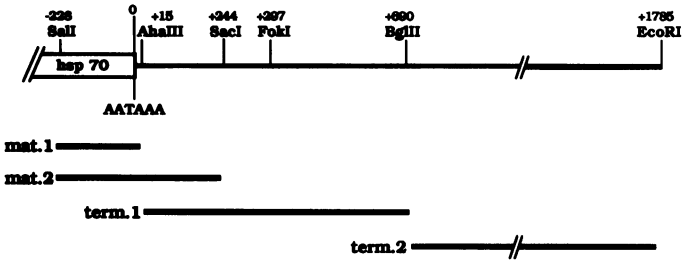

Figure 1. Map of the distal region of the *hsp70* gene at chromosome sub-division 87A. A portion of the *hsp70* structural gene is shown. A consensus polyadenylation signal sequence AATAAA is found at the 3′ end of the gene. The location and distance in nucleotides from the distal end of the polyadenylation signal of several restriction endonuclease cleavage sites are shown. Below the map are the names and locations of probes used in the study. The map is not drawn to scale.

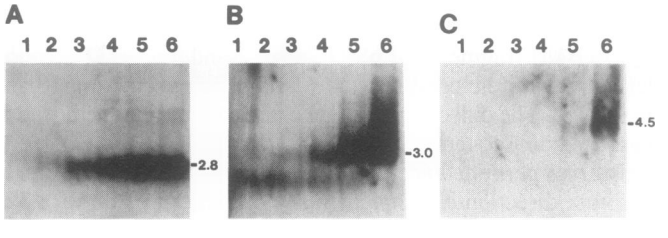

Figure 2. Northern analysis of *hsp70* RNA accumulation during temperature elevation. S2 cells were heat shocked at 37°C for: lane 1, 0 min; lane 2, 5 min; lane 3, 10 min; lane 4, 15 min; lane 5, 20 min; lane 6, 30 min. Total cellular RNA was prepared from each time point and 2.5 μg were used in each lane. Panels are autoradiograms of the Northern blot hybridized to the (**A**) mat.1, (**B**) term.1 and (**C**) term.2 antisense RNA probes. The sizes of the major bands were calculated based on RNA markers present on the same gel. Autoradiograms were exposed for (A) 2 hours, (B) and (C) 20 hours.

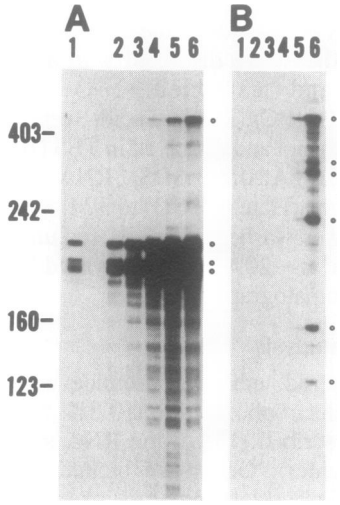

Figure 3. RNA probe protection analysis of *hsp70* RNA during temperature elevation. S2 cells were heat shocked at 37°C for: lane 1, 0 min; lane 2, 5 min; lane 3, 10 min; lane 4, 15 min; lane 5, 20 min; lane 6, 30 min. Total cellular RNA was prepared from each time point. 2.5 μg (panel **A**) or 25 μg (panel **B**) were analyzed using radiolabeled antisense RNA probes. The probes used were: panel A, mat.2; panel B, term.1. Radiolabeled *Hpa*II-digested pBR322 markers were present on the same gel. Autoradiograms were exposed for (A) 2 hours and (B) 20 hours.

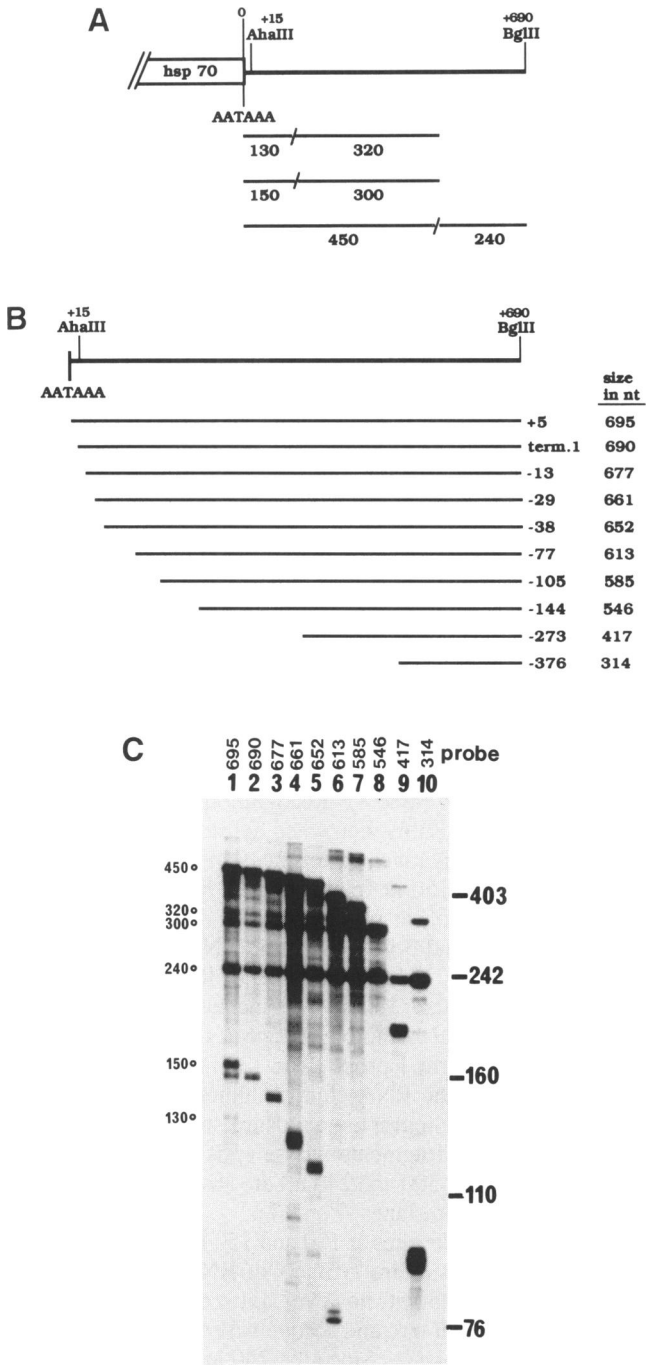

Figure 4. Positioning of the 5′ ends of *hsp70* RNA downstream of the polyadenylation site using Bal31 deletions of the term.1 probe. Panel **A:** The deduced locations of the cleavage sites downstream of *hsp70*. The sites were mapped using the series of deletion probes of term.1 shown in Figure 4B. As explained in the text, the fragments 130 and 150 nt in length are believed to be artifacts of probe protection assay, while the interruption at 450 nt downstream of the polyadenylation site is believed to be an authentic 3′ end. Panel **B:** The location (in nt) of Bal 31 deletion endpoints relative to the term.1 probe. The size (in nt) of each probe is also indicated. The sequence of the 5′ end of each probe was determined. Panel **C:** Determination of the 5′ end position of the cleavage sites downstream of *hsp70*. RNA was isolated from S2 cells which had been heat-shocked for 20 min at 41°C. RNA probe protection analysis was done, using 20 μg of RNA hybridized to various radiolabeled antisense RNA probes, prepared by Bal31 nuclease digestion. Above each lane are listed the size (in nt) of each deletion probe, as shown in Figure 4B. On the right side are shown the sizes of expected protected fragments of the probe term.1 (690 nt). Radiolabeled *Hpa*II-digested pBR322 markers were present on the same gel and are indicated on the left side. The autoradiogram was exposed for 20 hours.

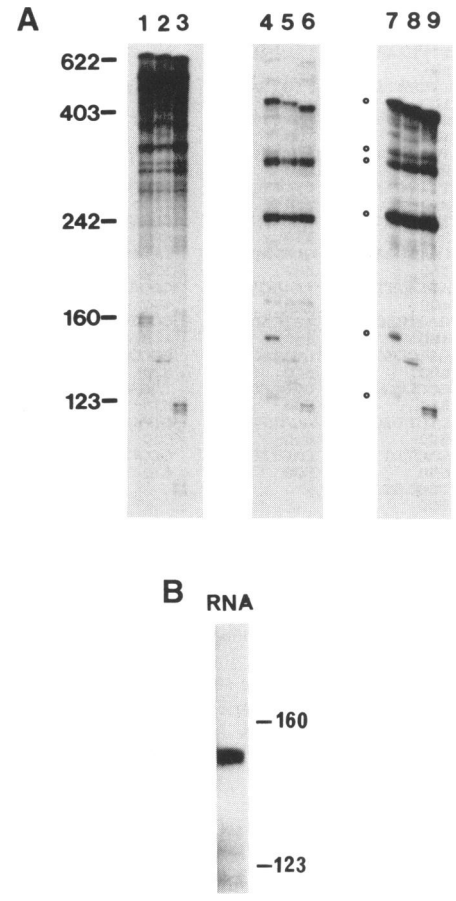

Figure 5. A. Comparison of RNAase digestion products of radiolabeled term.1 probe hybridized to synthetic term.1 RNA, adult *Drosophila* RNA, or S2 cellular RNA. Radiolabeled antisense RNA probes were hybridized to: lanes 1, 2, 3, synthetic RNA complementary to term.1 probe; lanes 4, 5, 6, RNA isolated from OregonR flys which had been heat-shocked for 30 min; lanes 7, 8, 9, RNA isolated from S2 cells which had been heat-shocked for 30 min. The probes were: lanes 1, 4, 7, term.1 probe; lanes 2, 5, 8, Bal31 deletion probe −13; lanes 3, 6, 9, Bal31 deletion probe −29. The autoradiogram was exposed for: lanes 1, 2 and 3, 2 hours; lanes 4, 5, and 6, 30 hours; lanes 7, 8 and 9, 20 hours. Radiolabeled *Hpa*II-digested pBR322 markers were present on the same gel. Fragments of interest are indicated by circles next to lane 7: 450, 320, 300, 240, 150, and 130 nt. **B.** S1 nuclease protection analysis of the cleavage site in *hsp70* RNA located 450 bases downstream of the polyadenylation site. The 3′ end of the *Fok*I site (Figure 1) was radiolabeled. The 393 bp *FokI*−*Bgl*II fragment was purified and hybridized to RNA isolated from S2 cells previously heat-shocked for 30 minutes at 37°C. The 150 base S1 nuclease resistant fragment corresponds in size to the cleavage site detected using RNAase protection which was mapped 450 bases downstream of the polyadenylation site. Radiolabeled *Hpa*II-digested pBR322 markers were present on the same gel and are indicated on the left side. The autoradiogram was exposed for 10 hours.

Induced RNAs protecting the entire mat.2 probe are also seen in Panel A (indicated by circle at 470 nt in lanes 4, 5 and 6), reflecting the presence of continuous transcripts that extend at least to the end of the probe, 250 nt beyond the polyadenylation site of mature message. The amount of these long transcripts is small compared to the amount of mature message but continues to increase until at least 60 min (lane 6, and not shown), long after the amount of normal polyadenylated message levels off at 20 min (lane 5), in agreement with the results of the Northern analysis (Fig. 2).

Protection with the 690 nt probe term.1 yields several prominent fragments shorter than the probe and only a relatively

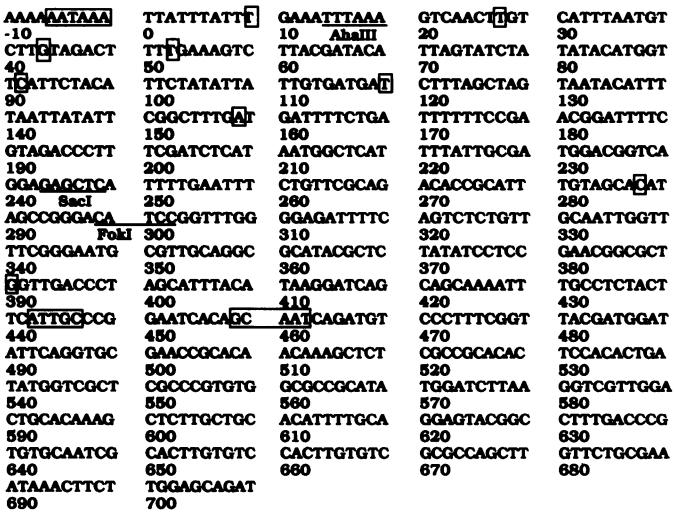

Figure 6. Sequence of the 3′ end region of *hsp70*. The sequence was determined from the *SalI* site (not shown) to the *BglII* site in both directions by the dideoxynucleotide chain termination method. The numbering is relative to the polyadenylation signal AATAAA, with the first nt following this sequence assigned the number zero. Upstream sequences are numbered negatively and downstream sequences are numbered positively. Restriction endonuclease sites used in this study are underlined. The location of the 5′ endpoints of the Bal31-generated deletions are boxed. The AATAAA sequence and hypothetical stem structure at nt 450 are boxed.

small amount of full length material (Fig. 3, panel B). The most prominent protected fragments are of sizes 450, 320, 300, 240, 150 and 130 nt (indicated by circles). Note that the 300 and 150 nt fragments, and also the 320 and 130 nt fragments add to give 450 nt, while the 240 and the 450 fragments add to give the full-length 690. Thus, these fragments could be produced by cleavage of long transcripts at sites 130, 150, and 450 nt downstream from the polyadenylation signal, as depicted in Figure 4A.

Mapping of transcripts downstream of the polyadenylation site

To test the explanation that the fragments protected by probe term.1 result from cleavage, we determined the positions of the 6 fragments within the term.1 probe. To do this we conducted a series of protections of 20 min heat shock RNA with derivatives of term.1 lengthened or shortened at its upstream end (Figure 4B). All such changes in the probe will correspondingly change the length of any protected fragment that extends to the upstream end of term.1. Fragments having 5′ ends within term.1 will be changed only later in the series of deletion probes. Nine cloned variants of term.1, with lengths ranging from 695 (a 5 nt addition) to 214 nt (a 376 nt deletion), were used for this analysis.

It is seen in Figure 4C that the size of one member of each of the three pairs of protected fragments deduced above (130, 150 and 450 nt, i.e. the two smallest and the largest fragments indicated by circles) was correspondingly altered by every probe in the series, indicating that their 5′ ends lie at the upstream end of term.1, as shown in Figure 4A. On the other hand, the remaining three fragments (320, 300 and 240 nt, i.e. the three middle fragments indicated by the circles) are affected only later in the series (300 and 320 nt are shortened beginning in lane 8) or remain intact through the series (240 nt). We conclude that the 6 downstream fragments are indeed produced by cleavage of downstream transcripts at sites 130, 150 and 450 nt beyond the polyadenylation site, and have the map positions shown in Figure 4A.

Comparison of cleavage sites in synthetic and authentic *Drosophila* RNA

The observed protected fragments could have resulted from cleavage *in vivo* or as an artifact of probe-protection analysis due to RNAse sensitivity at preferred sites of the RNA duplex. In order to detect possible artifacts of this sort, controls were performed with synthetic RNA complementary to the term.1 probe. Full-length synthetic RNA was hybridized with radiolabeled term.1 and also with two of the deletion derivatives (the −13 and −29 nt probes in Figure 4B), and incubated with RNAse. These digests (Figure 5A, lanes 1, 2 and 3) were compared to digests of RNA obtained from heat-shocked adult flys (lanes 4, 5 and 6) or from S2 cells (lanes 7, 8 and 9).

Gel-electrophoretic analysis of the resulting digests shows the expected full-length protected fragment of term.1 in the synthetic RNA (Figure 5A, lane 1, 690 nt). In addition, there is a prominent series of fragments resulting from cleavage of term.1 at specific sites in the RNA duplex. If any of the three sites we previously detected at 130, 150 or 450 nt result as artifacts from such preferential cleavage, corresponding fragments should appear in the control with synthetic RNA.

As before, the 150 nt fragment is seen in both S2 cell RNA (lane 7) and fly RNA (lane 4). Recognition of this fragment is facilitated by the bandshifts seen when hybridization is done with the two deletion derivatives of term.1 (lanes 5 and 6 or lanes 8 and 9). The same 150 nt fragment does indeed appear in the synthetic RNA control, as may be seen by comparing synthetic RNA in lanes 1, 2 and 3 with authentic RNA in lanes 4, 5 and 6 or lanes 7, 8 and 9. Note that the size of this fragment in lane 1 is somewhat larger than 150 nt. This increased size results from 10 bases of additional homology between the synthetic term.1 antiprobe with the term.1 probe. The 130 nt fragment was also present in the synthetic RNA control, although it is difficult to discern in the autoradiographic exposure shown in Figure 5A. In contrast, all four fragments whose ends are adjacent to the 450 nt site (450, 320, 300 and 240 nt) are absent from the control (compare lanes 1 with lanes 4 and 7).

We conclude that the sites at 130 and 150 nt downstream from the polyadenylation site are artifacts of RNA probe-protection analysis. The site at 450 nt, however, must result from cleavage of longer transcripts *in vivo* and accounts for the 3 kb *hsp70* RNA seen in Northern analyses. Since the 240 and 450 nt fragments are much more abundant than the full-length 690 nt full-length fragment (in lanes 4 and 7), we further conclude that relatively few long transcripts escape cleavage at the site at 450 nt.

Sequence of the 3′ region of the *hsp70* gene and other characteristics of longer transcripts

In order to examine the 3′ end of the *hsp70* gene for features that might be responsible for cleavage 450 nt downstream of the polyadenylation signal, we determined the sequence of the 820 bp *SalI*−*BglII* interval of the 3′ region (Figure 6). The location of the RNA cleavage site at 450 nt, initially determined by RNA probe protection analysis against DNA standards, was determined to be at 450 ± 10 nt by DNA probe protection and S1 nuclease analysis with a 393 bp *FokI*−*BglII* fragment as probe (Figure 5B).

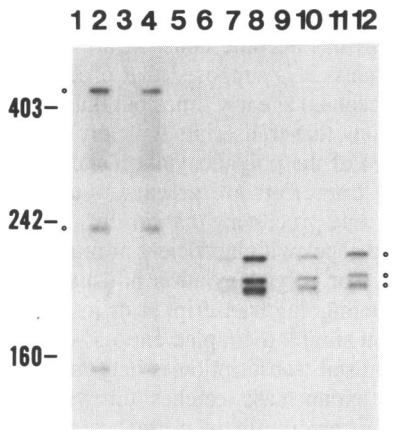

Figure 7. RNA probe protection analysis of *hsp70* RNA following oligo-dT fractionation. Total cellular RNA isolated from heat-shocked (60 min) or untreated S2 cells was fractionated by oligo-dT chromatography. RNA was hybridized to (lanes 1–6) term.1 or (lanes 7–12) mat.1 probes followed by protection analysis. 20 μg of RNA was used in lanes 1–6 and 2 μg of RNA was used in lanes 7–12. RNA samples were: lanes 1 and 7, total cellular RNA; lanes 2 and 8, heat-shocked total cellular RNA; lanes 3 and 9, oligo-dT flow-through RNA; lanes 4 and 10, heat-shocked oligo-dT flow-through RNA; lanes 5 and 11, oligo-dT bound RNA; lanes 6 and 12, heat-shocked oligo-dT bound RNA. Radiolabeled *Hpa*II-digested pBR322 markers were present on the same gel. The autoradiogram was exposed for 20 hours.

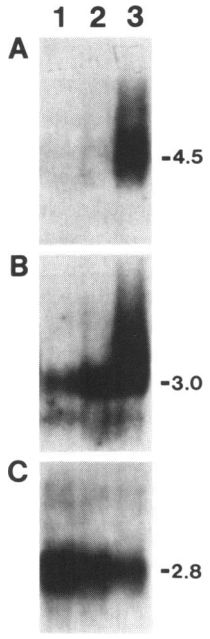

Figure 8. Northern analysis of *hsp70* RNA after temperature elevation at 37°, 39°C and 41°C. S2 cells were heat shocked for 20 min at: lane 1, 37°C; lane 2, 39°C; lane 3, 41°C. Total cellular RNA was prepared from each sample and 2.5 μg were electrophoresed in each lane. The radiolabeled probe used in each panel were: panel **A**, term.2, panel **B**, term.1, panel **C**, mat.1. Autoradiograms were exposed for (panels A and B) 20 hours, (panel C) 2 hours.

The polyadenylation status of the 3′ end at 450 nt was investigated by RNA probe protection analysis of RNA fractionated by chromatography on oligo-dT cellulose (Figure 7). As expected, the polyadenylated fraction (lane 12) protected

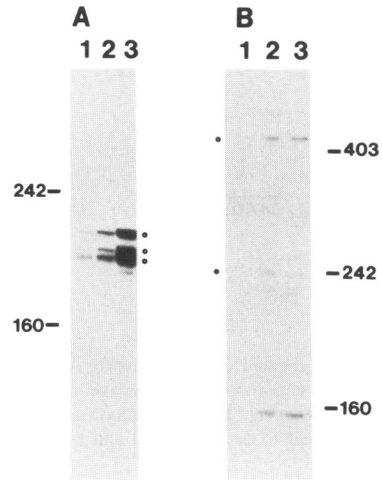

Figure 9. RNA probe-protection analysis of *hsp70* RNA produced early after temperature elevation. S2 cells were centrifuged and concentrated prior to temperature elevation. Lysis was performed immediately following heat shock. Total cellular RNA was prepared and 2 μg (panel **A**) or 20 μg (panel **B**) RNA was hybridized to mat.1 (panel A) or term.1 (panel B) probes. Cells were heat-shocked at 37°C for: lane 1, 0 min; lane 2, 2 min; lane 3, 4 min. The autoradiogram was exposed for (panel A) 2 hours and (panel B) 20 hours.

the three fragments of the mat.1 probe attributed to the 3′ region of mature message more strongly than did the non-polyadenylated fraction (lane 10). In contrast, the transcripts that protected the 450 nt and 240 nt fragments of term.1 were completely confined to the non-polyadenylated fraction (lane 4) rather than the polyadenylated fraction (lane 6). Total RNA is shown in lanes 1 and 8.

Both here (Figure 7) and in protection analyses shown earlier (Figures 3, 4 and 5) it is seen that the amounts of the 240 and 450 nt protected fragments are comparable. This indicates that transcription extends beyond the site at 450, since the sum of the lengths of these two fragments equals that of the full-length probe.

Heat elevation above 37°C causes increased accumulation of these long transcripts

It was seen in Figures 2 and 3 that transcription beyond the polyadenylation site, including that extending to the site at 450 nt, continues after the amount of mature message reaches a plateau. A more striking increase in the amount of the same long transcripts and a parallel decrease in the amount of mature message is seen with temperature elevation above 37°C, as shown in Figure 8. Successive Northern hybridization was done with probes mat.1 (panel C), term.1 (panel B) and term.2 (panel A) to RNA isolated from S2 cells previously heat-shocked at 37°C (lane 1), 39°C (lane 2) or 41°C (lane 3) for 20 min. It is clear that the amount of RNA extending to the site at 450 nt (panel B) increases at the higher temperatures, while the amount of mature mRNA actually declines (panel C). This was also shown by RNA probe-protection analysis of the same RNA samples (not shown). These effects suggest that prolonged heat shock or particularly high temperature causes failure of the normal mechanism of 3′ end formation without affecting the process giving rise to transcripts cleaved at 450 nt.

Yost and Lindquist (20) have shown that intron splicing is also inhibited at high heat shock temperatures, and can be prevented

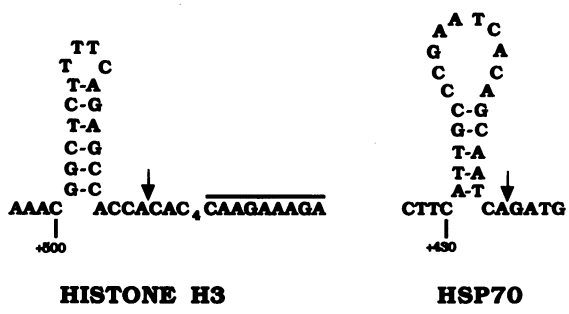

Figure 10. Comparison of the stem-loop structure formed at the 3′ end of histone H3 RNA with a possible stem-loop at the cleavage site 450 nt downstream of the *hsp70* polyadenylation signal. The denaturation free energy of the histone H3 structure was calculated to be 9 kcal/mol and the *hsp70* structure was 5 kcal/mol.

by a previous lower temperature heat shock. In contrast, using Northern and RNA probe-protection analyses, we found that such heat pretreatment does not restore the normal level of polyadenylation after severe heat shock (not shown).

Accumulation of the longer transcripts at very early times

Having established that transcripts extending well past the normal polyadenylation site continue to accumulate at late times after temperature elevation, when the amount of mature message becomes essentially constant, we sought to examine the relative accumulation of long transcripts at very early times. For this purpose, lysis solution was added directly to concentrated cell cultures at various times after temperature elevation, without cooling or centrifugation. This precaution was taken in order to avoid possible RNA processing that might occur during the harvesting steps employed in other experiments.

Protection of the probes term.1 and mat.1 with whole cell RNA isolated at 0, 2 and 4 min after temperature elevation is shown in Figure 9. Full-length protection of mat.1 (panel A), and protection of the 450 nt and 250 nt proximal and distal segments of term.1 (panel B) were clearly evident by 2 minutes (lane 2). Taking into account the different amounts of RNA used in the two protection series (10-fold more for term.1 than mat.1), it is seen that there is about 100 times less protection of term.1 than of mat.1 in the 2 and 4 minute samples. Thus, if the long RNA molecules protecting term.1 are precursors to mature message, their lifetime must be only a few seconds. Rather than precursors, then, we believe that long RNAs result from failure of the normal termination reaction and shunting of processing along an alternate path.

DISCUSSION

We have sought to determine the pattern of transcription and processing in the 3′ region of the *hsp70* heat shock gene. Shortly after temperature elevation, by far the most abundant induced *hsp70* RNA observed is the mature polyadenylated message. In addition, at the earliest time at which induced transcripts are reliably observed, 2 minutes after temperature elevation, we detect approximately 1/100 as many longer transcripts as of mature message. Thereafter, long transcripts continue to increase but the amount of mature message levels off by about 20 minutes.

The low level of long transcripts relative to the amount of mature message when the latter is increasing most rapidly and the continued increase of long transcripts after the amount of mature message has become constant are not characteristic of a precursor, unless it is supposed that processing is very rapid (less than 10 seconds) at early times but fails later. We therefore question the view that transcripts typically extend continuously far downstream of the polyadenylation site in normal cells and that such long transcripts are precursors to mRNA.

A view of 3′ end processing that fits our observations and also those including genes with inefficient or mutant polyadenylation sites (3,21) is that polyadenylation is usually accompanied by cleavage of the ongoing transcript at or just downstream of the polyadenylation site by a coupled endo-5′-exonuclease (1,3,8). The extent of distal transcription is determined by the location at which the exonuclease catches up with the polymerase. According to this picture, in the normal case, as with *hsp70* soon after a moderate heat shock, cleavage occurs at or near the polyadenylation site and rapid 5′ exonucleolytic action limits the production of long transcripts to a low level. But prolonged or severe heat shock, or mutations in the polyadenylation signal (3), cause failure of normal cleavage and/or normal 5′ exonucleolytic digestion, allowing the accumulation of long transcripts.

Two studies (4, 22) reached similar conclusions for termination and polyadenylation of yeast CYC1 RNAs. CYC1 termination occurred within 100 nt of the normal polyadenylation site and mutation of the polyadenylation signal caused the accumulation of extended transcripts (23). Since we did not detect 3′ ends immediately downstream of the polyadenylation site, we surmise that termination of *hsp70* RNA occurs close to the site.

Under conditions of prolonged or intense heat shock, we suppose that normal cleavage of *hsp70* RNA often fails to occur, allowing transcription to extend far downstream. The failure to cleave could be due to inhibitory effects of heat shock itself on the machinery of processing, or the high level of transcripts could overwhelm the processing capacity. In such cases, we presume that cleavage by some different mechanism, not coupled to the 5′ exonucleolytic activity, occurs at the 450 nt site farther downstream. Such lack of 5′ nucleolysis could explain the stability of the distal fragments. Alternatively, the accumulation of longer transcripts could be due to differential stability of RNA after heat shock, resulting in a failure to cleave the longer transcripts.

We do not know whether the formation of stable non-polyadenylated transcripts extending to 450 nt is functional or fortuitous. Non-polyadenylated 3′ ends are generally unstable. The non-polyadenylated transcripts of histone mRNA are stabilized by specific 3′ loop structures (24, 25). The stability of the 3′ end of the non-polyadenylated 450 nt transcript of *hsp70* may result from formation of the stem-loop depicted in Figure 10, for which the calculated denaturation free energy is about 5 kcal. It may therefore be that the formation of stable non-polyadenylated transcripts extending to 450 serves to provide an alternative species of stable messenger when normal cleavage upstream fails, as in the case of extreme heat shock.

Switching between alternative modes of generating 3′ ends of messenger RNA may be a general method of providing stability under changing cellular conditions. It is interesting to note that nonpolyadenylated histone H1 RNA, which fluctuates in amount during the avian cell cycle, is replaced in some nondividing cell types by transcription of a different H1 gene, which produces polyadenylated RNA (26). This polyadenylated H1 RNA contains within its 3′ end the conserved sequences responsible for directing cleavage into the nonpolyadenylated form of the RNA. Thus alternative modes of 3′ end processing may operate in either

direction: switching from dominant polyadenylated to nonpolyadenylated RNA as observed here for *hsp70*, or from dominant nonpolyadenylated to polyadenylated RNA, as demonstrated for histone H1.

ACKNOWLEDGEMENTS

We wish to thank Dale Dorsett, Linda Hyman, Claire Moore, Gordon Moore, Mark Mortin and Richard Southgate for helpful discussions and critical reading of the manuscript. The work was supported by postdoctoral fellowship grants from the National Institutes of Health and the Arthritis Foundation to S.L.B.

REFERENCES

1. Proudfoot, N. 1989. Trends Biochem. Sci. **14**, 105–110.
2. Kerppola, T. and Kane, C. 1991. FASEB J. **5**, 2833–2842.
3. Connelly, S. and Manley, J. 1988. Genes Devel. **2**, 440–452.
4. Russo, P. and Sherman, F. 1989. Proc. Natl. Acad. Sci. USA **86**, 8348–8352.
5. Proudfoot, N. and Whitelaw, E. 1988. in *Frontiers in Molecular Biology*. pp. 97–129. IRL Press
6. Richardson, J. 1993. Crit. rev. in Biochem. and Mol. Biol. **28**, 1–30.
7. Hyman, L. and Moore, C. 1993. Mol. Cell. Biol. **13**, 5159–5167.
8. Eggermont, J. and Proudfoot, N. 1993. EMBO J. **12**, 2539–2548.
9. Ashfield, R., Enriquez-Harris, P. and Proudfoot, N. 1991. EMBO J. **10**, 4197–4207.
10. Kerppola, T. and Kane, C. 1990. Biochemistry, **29**, 269–278.
11. Sato, K., Ito, R., Baek, K., and Agarwal, K. 1986. Mol. Cell. Biol. **6**, 1032–1043.
12. Schneider, I. 1972. J. Embryol. Exp. Morph. **27**, 353–365.
13. Shields, G. and Sang, J. 1977. Dros. Inform. Serv. **52**, 161.
14. Livak, K., Freund, R., Schweber, M., Wensink, P., and Meselson, M. 1978. Proc. Natl. Acad. Sci. USA. **75**, 5613–5617.
15. Wensink, P., Finnegan, D., Donelson, J., and Hogness, D. 1974. Cell **3**, 315–325.
16. Holmes, D., and Bonner, J. 1973. Biochemistry **12**, 2330–2338.
17. Maniatis, T., Fritsch, E., and Sambrook, J. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
18. Melton, D., Krieg, M., Rebagliati, M., Maniatis, T., Zinn, K., and Green, M. 1984. Nucl. Acids Res. **12**, 7035–7056.
19. Maxam, A., and Gilbert, W. 1980. Meth. Enzymol. **65**, 499–560.
20. Sanger, F., Nicklen, S., and Coulson, A. 1977. Proc. Natl. Acad. Sci. USA **74**, 5463–5488.
21. Tinoco *et al.* 1973. Nature New Biol. **246**, 40–41.
22. Yost, H. and Lindquist, S. 1986. Cell **45**, 185–193.
23. Lanoix, J. and Acheson, N. 1988. EMBO J. **7**, 2515–2522.
24. Osborne, B. and Guarente, L. 1988. Genes Dev. **2**, 766–772.
25. Zaret, K., and Sherman, F. 1982. Cell **29**, 563–573.
26. Mowry, K., Oh, R., and Steitz, J. 1989. Mol. Cell. Biol. **9**, 3105–3108.
27. Vasserot, A., Schaufele, F., and Birnstiel, M. 1989. Proc. Natl. Acad. Sci. USA **86**, 4345–4349.
28. Kirsh, A., Groudine, M., and Challoner, P. 1989. Genes Devel. **3**, 2172–2179.