Autocatalytic RNA cleavage in the human β-globin pre-mRNA promotes transcription termination

Alexandre Teixeira[1], Abdessamad Tahiri-Alaoui[1], Steve West[1], Benjamin Thomas[1], Aroul Ramadass[2], Igor Martianov[1], Mick Dye[1], William James[1], Nick J. Proudfoot[1] & Alexandre Akoulitchev[1]

[1]*Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK*
[2]*The Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK*

New evidence indicates that termination of transcription is an important regulatory step, closely related to transcriptional interference[1] and even transcriptional initiation[2]. However, how this occurs is poorly understood. Recently, *in vivo* analysis of transcriptional termination for the human β-globin gene revealed a new phenomenon—co-transcriptional cleavage (CoTC)[3]. This primary cleavage event within β-globin pre-messenger RNA, downstream of the poly(A) site, is critical for efficient transcriptional termination by RNA polymerase II[3]. Here we show that the CoTC process in the human β-globin gene involves an RNA self-cleaving activity. We characterize the autocatalytic core of the CoTC ribozyme and show its functional role in efficient termination *in vivo*. The identified core CoTC is highly conserved in the 3′ flanking regions of other primate β-globin genes. Functionally, it resembles the 3′ processive, self-cleaving ribozymes described for the protein-encoding genes from the myxomycetes *Didymium iridis* and *Physarum polycephalum*, indicating evolutionary conservation of this molecular process. We predict that regulated autocatalytic cleavage elements within pre-mRNAs may be a general phenomenon and that functionally it may provide the entry point for exonucleases involved in mRNA maturation, turnover and, in particular, transcriptional termination.

To characterize the mechanism of co-transcriptional cleavage and its role in transcriptional termination, we first analysed the effect of the CoTC element *in vitro* in a reconstituted transcription assay with RNA polymerase II[4]. The CoTC element from the β-globin gene[3] was cloned in both orientations between two G-less cassettes (Fig. 1a) in a template designed to monitor transcription elongation, pSLG402 (ref. 5). pSLG402 itself was also used as a control for efficient transcriptional elongation, in accordance with its original design[5]. In the presence of the correctly oriented CoTC element, transcripts from the distal G-less cassette were hardly detectable after the treatment with T1 RNase (Fig. 1a, lane A). To characterize the effect further we also compared the full-length run-off transcripts without RNase treatment. The time course demonstrated that full-length transcripts of 1.2 kilobases gradually accumulated from the control template; however, the termination template failed to produce detectable full-length transcripts

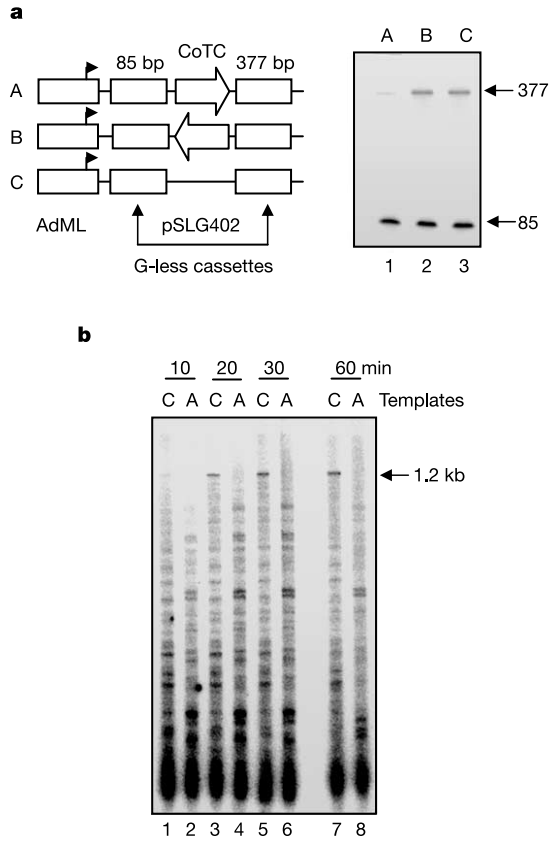

Figure 1 Effect of the CoTC element on reconstituted transcription *in vitro*. **a**, Reconstituted transcription from termination (A, lane 1) and control (B, C, lanes 2 and 3, respectively) templates after treatment with T1 RNase. Arrows mark the positions of the G-less transcripts from the promoter proximal (85 nt) and distal (377 nt) cassettes. **b**, Time course of reconstituted run-off transcription from the termination (A, even-numbered lanes) and control (C, odd-numbered lanes) templates. The arrow marks the position of the full-length transcript (1,200 nt). Abbreviation: kb, kilobases.

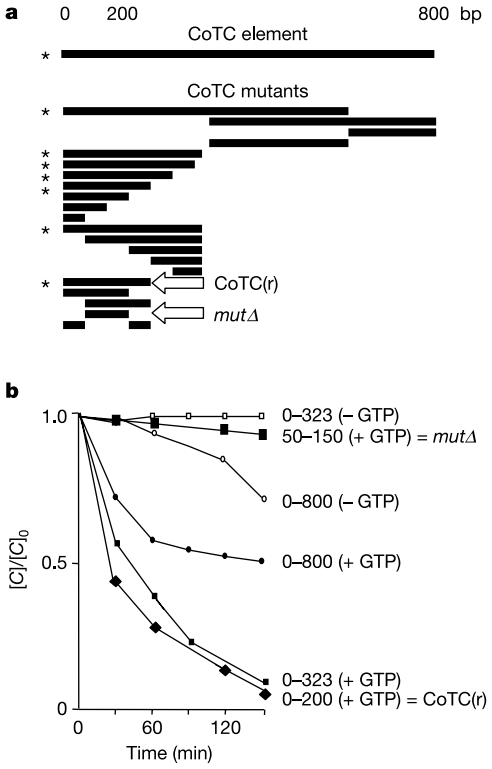

Figure 2 Mapping of the minimal catalytic core within the CoTC element. **a**, Deletion analysis of the synthetic CoTC transcripts. Deletion mutants of the synthetic CoTC RNA were tested for GTP-dependent degradation in a protein-free environment. Transcripts that displayed catalytic activity are marked with asterisks; arrows mark the minimal ribozyme core of the CoTC element (1–200), designated CoTC(r), and the deletion mutant (50–150), designated *mutΔ*. **b**, Kinetics of GTP-dependent degradation for the deletion mutants.

(Fig. 1b, compare A with C). Moreover, we observed additional partial-length products that had low stability after the first 30 min. We concluded that under conditions of reconstituted transcription, the presence of the CoTC element in the correct orientation prevented the production of stable transcripts from the distal portion of the template. This could be attributed to several causes, ranging from transcriptional pausing or termination to the short half-life of the CoTC transcript.

To analyse further the properties of transcripts containing the CoTC element, we decided to uncouple transcription from the subsequent RNA processing steps. We synthesized ^{32}P-labelled synthetic CoTC RNA by using T7 RNA polymerase as a potential substrate for specific RNA cleavage activities. Surprisingly, our analysis revealed that, unlike the control, the CoTC RNA had a short half-life with a distinct pattern of degradation/cleavage products in the presence of the nucleophilic cofactors Mg^{2+} and GTP[6] (Supplementary Fig. 1a). These products appeared under protein-free conditions, because the cleavage activity was maintained after gel purification in the presence of 8 M urea, after tryptic digestion, after extraction with phenol/chloroform, and in the presence of SDS[7] (Supplementary Fig. 4a). The degradation kinetics of the full-length synthetic CoTC RNA corresponds to a first-order catalytic reaction (Supplementary Fig. 1b) with $k = 0.015\,\mathrm{min}^{-1}$ and a half-life of 38 min, which is longer than for some classes of self-cleaving ribozymes[8]. However, the recently described autocleaving prokaryotic ribozyme in the *Bacillus subtilis glmS* mRNA[9] has a half-life of 4 h ($k = 0.003\,\mathrm{min}^{-1}$) under certain conditions. This is still 10^3-fold faster than spontaneous cleavage of an unconstrained RNA linkage[10] and tenfold faster than spontaneous nucleophilic cleavage of a constrained RNA linkage[11]. Our results indicate that CoTC RNA undergoes a single-molecule autocatalytic cleavage reaction, as described for certain classes of RNA ribozyme *in vivo* and *in vitro*[6,8]. The rates of autocatalytic cleavage can be subject to allosteric regulation. For example, *glmS* mRNA cleavage is activated almost 10^3-fold, with $k_{max} = 1\,\mathrm{min}^{-1}$, by its effector GlcN6P (ref. 9). If the complex kinetics of CoTC cleavage is taken into account (see below), it might well be subject to regulation and reach higher rates of cleavage *in vivo*. In fact, preliminary biochemical analysis confirms the existence of regulatory effectors and RNA chaperones for the CoTC element (A.A., unpublished data).

Next, we mapped the minimal region sufficient for autocatalytic cleavage of the synthetic CoTC RNA. We generated a series of 5′- and 3′-end deletion mutants spanning the 850 nucleotides (nt) of the CoTC element (Fig. 2a). In the cleavage analysis we monitored two parameters: the half-life of the transcripts and the dependence on GTP. A 200-nt region close to the 5′ end of the CoTC element displayed a cleavage half-life of 15 min in the presence of GTP cofactor, with further deletions from the 5′ and 3′ ends (*mutΔ*) rendering the transcripts stable under the same conditions (Fig. 2b). We conclude that the ribozyme core, CoTC(r), is located within these 200 nt.

To establish more comprehensive evidence for the autocatalytic activity of CoTC(r), we investigated its structural properties in more detail. First, we analysed single-stranded and double-stranded secondary structures within CoTC(r) by using the nucleases S1 and T1, which are specific for single-stranded RNA, as well as the double-strand-specific nuclease V1. As shown in Fig. 3a, digestion with V1 and S1 identifies a series of single-stranded and double-stranded regions (for a detailed analysis see Supplementary Fig. 2). We also analysed the secondary structure of CoTC(r) using a chemical modification approach (Supplementary Fig. 3). Single-stranded nucleosides were modified with dimethylsulphate (DMS) at A(N1), C(N3) and with 1-cyclohexyl-3-(2-morpholinoethyl)car-

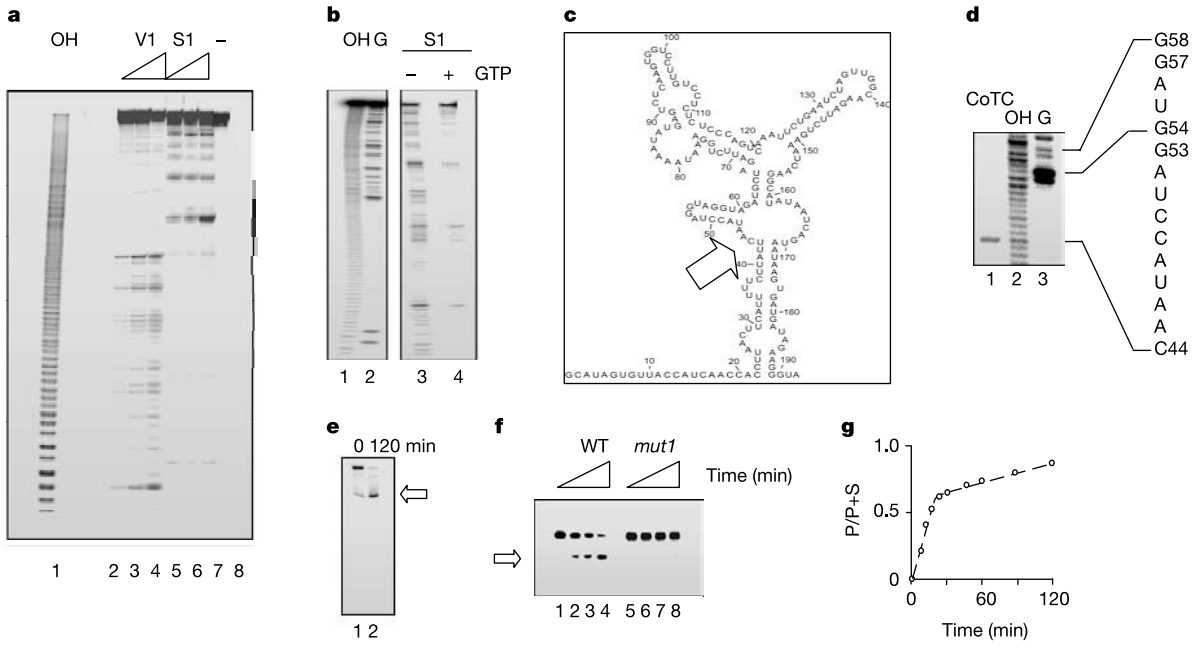

Figure 3 Analysis of the CoTC core secondary structure and mapping of the autocatalytic cleavage site. **a**, Cleavage profile for the 5′-^{32}P CoTC(r) RNA (lane 8) treated with endonucleases V1 (lanes 2–4) and S1 (lanes 5–7). **b**, S1 cleavage profile of the CoTC(r) RNA (lane 3) in the absence or presence of 1 mM GTP (lanes 3 and 4). **c**, Secondary structure of the catalytic CoTC core, as defined by the folding algorithm[13] and the data from enzymatic and chemical analyses. The arrow marks the position of the primary autocatalytic cleavage site at C44. **d**, The product of the autocatalytic cleavage isolated and resolved on a sequencing gel. OH identifies the alkaline hydrolysis ladder marking each nucleoside position on the gel; G identifies guanosine positions after partial digestion of the CoTC(r) with T1. **e**, The CoTC element (1–323) is cleaved (lane 2, arrow) under protein-free conditions in the presence of GTP. **f**, CoTC(r), the minimal catalytic core, is cleaved in a time course conducted under protein-free conditions (lanes 1–4 correspond to 0, 5, 15 and 120 min, respectively). Cleavage was performed with the 5′-^{32}P core RNA in the presence of 200 mM KCl. The arrow marks the primary cleavage site. *mut1* denotes a double mutant (C45C46) at the cleavage site. **g**, Profile of the biphasic product accumulation in the CoTC cleavage reaction. P, product; S, substrate.

bodiimide metho-*p*-toluene sulphonate (CMCT) at G(N1), U(N3). The positions of modified nucleosides were identified by primer extension with reverse transcriptase (Supplementary Fig. 3).

The data from the analyses of enzymatic and chemical modification were used to constrain simulations of secondary structure with the Mfold 3.1 program[12]. The resulting structure (Fig. 3c) is predicted to have a free energy change on folding of −32.87 kcal mol^{-1} and has clear structural and sequence characteristics. One notable feature is the presence of an AU-rich stem at the base of the structure, formed by the 5′ and 3′ ends of the sequence. Earlier deletion analysis with *mutΔ* identified this AU-rich stem as essential for GTP-dependent cleavage (Fig. 2). The highly structured organization of CoTC(r) compares favourably in complexity of folding with riboswitches described previously[13].

We also analysed the effect of GTP on the secondary structure of CoTC(r) (Fig. 3b). S1 digestion shows a loss of single-strand endoRNase cleavages in several positions in the presence of GTP (Fig. 3b), indicating significant changes in the secondary structure of CoTC(r). Classical studies of the group I ribozyme showed that nucleophilic attack by GTP was dependent on 2′- and 3′-hydroxy groups[14,15]. The diphosphate and monophosphate derivatives (GDP and GMP) still supported the cleavage reaction, whereas treatment with dideoxy-GTP (ddGTP) resulted in a loss of activity. For CoTC(r) we also observed a loss of cleavage with ddGTP (Supplementary Fig. 4a), whereas GTP, GDP and GMP were equally effective cofactors. We did not detect any covalent association of GTP with the cleavage products (data not shown). We observed a high affinity of GTP for the structured CoTC RNA. This was shown by pull-down experiments with the ^{32}P-labelled CoTC RNA in the presence of a tRNA competitor with agarose beads cross-linked to either GTP or UTP. Only with GTP was the labelled RNA retained on the agarose beads (Supplementary Fig. 4c). Further analysis of the cleavage mechanism will delineate the exact role played by GTP.

Degradation profiles of CoTC(r) indicate the presence of a primary cleavage site (Fig. 3c–e), which is also resistant to SDS treatment[7] (Supplementary Fig. 4a). We identified the position of the site by isolating the cleavage product and resolving it on the sequencing gel against defined RNA markers (Fig. 3c, d). We then generated several double-point mutants around the cleavage site, as well as at the stem of the CoTC structure. In particular, the cleavage site mutant *mut1* (C45C46) was severely impaired in its function (Fig. 3f). Several additional mutants within the stem (C39C40, C40C41) also lost autocatalytic activity (data not shown). These mutant results are consistent with a functional role of the stem, as shown in the earlier degradation assay of *mutΔ* CoTC RNA (Fig. 2). We also confirmed that the 3′ ends of the CoTC primary cleavage site could be labelled with pCp in a T4 RNA ligase reaction, whereas the 5′ ends could be labelled with polynucleotide kinase after treatment with phosphatase (data not shown). This labelling specificity suggests that the cleavage reaction generates 3′-hydroxy and 5′-phosphate ends.

An analysis of product accumulation during CoTC(r) cleavage reveals biphasic kinetics (Fig. 3g)[16]. Here, the early cleavage of a portion of the substrate is observed at a high k of about 1 min^{-1} in the initial linear range, followed by slower cleavage of the remaining substrate. The explanation for the biphasic kinetics might lie with the alternative conformation of E–S (ribozyme–substrate) forms and slow exchange with the active conformation. Indeed, correct folding before cleavage analysis proved important during experimental analysis of CoTC(r). In the context of highly regulated transcriptional termination *in vivo*, the biphasic behaviour of CoTC(r) might accommodate multilayered regulation by effectors and RNA chaperones, as mentioned above. We are currently investigating further details of this autocatalytic reaction because we hope to be able to classify and relate the CoTC(r) element to other known classes of catalytic RNA[6].

To confirm the biological relevance of the *in vitro* autocatalytic activity of CoTC(r) we determined whether the catalytic core of the CoTC element is implicated in the termination of RNA polymerase II-mediated transcription. The plasmid containing the β-globin gene was transfected into HeLa cells, followed by nuclear run-on (NRO) analysis of termination as described previously[3]. The β-globin gene contained one copy of the 200-nt core CoTC(r) or its mutant forms replacing the whole 3′ flanking region (region B5–B10)[3], as indicated. In the absence of CoTC(r), transcriptional termination is defective and NRO signals are detected far downstream of the cleavage site by probe U3 (Fig. 4a–c). After insertion of the minimal ribozyme core CoTC(r), more than 60% of polymerases successfully terminate (Fig. 4c). The cleavage-site mutant *mut1*, which is defective in catalytic cleavage (Fig. 3f), has a low termination efficiency (Fig. 4b, c). In addition, deletion mutant *mutΔ* (see Fig. 2b), as well as point mutants within the stem structure defective in cleavage (*mut 5* (C39C40) and *mut 6* (C40C41)), were similarly deficient in termination (data not shown). These experiments *in vivo* indicate strongly that the catalytic activity of the identified minimal ribozyme core of the CoTC element is essential in the transcriptional termination process.

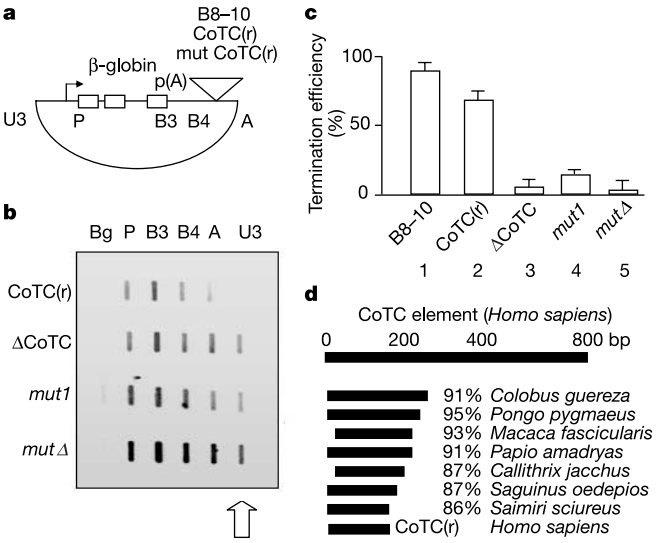

Figure 4 The catalytic ribozyme core of CoTC mediates termination *in vivo* and is conserved through evolution. **a**, Read-through *in vivo* assay: the 3′ end of the β-globin gene (B5–10) was removed and replaced with the partial sequence B8–B10, or the CoTC core constructs (CoTC(r)) and its mutants, as described in the text. The density of the RNA polymerase II that escaped termination was monitored by run-on assay. **b**, The upper panel marked CoTC(r) of the read-through run-on assay shows the distribution of the RNA polymerase II density on the plasmid where the 3′ end of the β-globin gene was replaced with CoTC(r). The lower panel marked ΔCoTC shows the result for the β-globin gene lacking the 3′ end. The arrow marks the U3 probe for the polymerase that escaped termination. *mut1* denotes the mutant at the ribozyme cleavage site (C45C46), deficient in autocatalytic cleavage, as shown in Fig. 3. *mutΔ* denotes the deletion mutant 50–150 of the CoTC (Fig. 2), also deficient in autocatalytic cleavage. **c**, Efficiency of termination as measured by the run-ons for the read-through assay. 100% corresponds to the level of termination observed with the U3 probe from the β-globin gene containing its full 3′ end, as described in ref. 3. The 3′ end was replaced with the β-globin gene 3′ end (B8–10) fragment containing the full CoTC element, the core ribozyme CoTC element, no core CoTC, cleavage mutant *mut1* or the deletion mutant *mutΔ*, respectively. Results are means + s.d. for four duplicate sets of independent assays. The results are statistically significant ($P < 0.05$). **d**, Conservation of the CoTC element in the non-coding 3′ end of the homologues of the β-globin gene among primates. The diagram shows regions of high homology and percentage of identity between human CoTC and the primate genes. *S. sciureus* features the catalytic core of the CoTC element, CoTC(r), as the only highly preserved part of the sequence.

We finally turned our attention to the possibility that autocatalytic elements such as core CoTC might have a more general function in transcriptional termination. A limited search in the close taxonomy group of primates, in which sequence similarity would be the highest, reveals that β-globin genes from the available primate sequences contain in their 3′ flanking regions sequences homologous to the full-length human CoTC element (Fig. 4d). The most highly conserved region corresponds exactly to the catalytic core CoTC(r). Because the 3′ flanking region sequences are normally highly divergent, this observed homology is especially significant. Indeed, in *Saimiri sciureus* the highly conserved homologue of the ribozyme core CoTC was identified 1,456 nt downstream of the conserved poly(A) site. The region immediately adjacent to the *S. sciureus* core CoTC homologue is not present in humans. The identified stem structure of the core CoTC also seems to be highly conserved: replacement of C31 by U in *Pongo pygmaeus*, and of A169 by G in *S. sciureus*, retains nucleoside pairing important for the structure. Recently, an independent study of the transcriptional termination model (A.R., unpublished data) has been conducted on the basis of a probabilistic generalized linear model and Bayes theorem[17]. The model identifies, through the primary sequence and secondary structure, an additional general termination marker downstream of the poly(A) site in several tested genes. When applied to the human β-globin gene, it successfully identifies the CoTC element as the termination marker of interest. Within the limitations of our current analysis these results encourage us to predict that the autocatalytic cleavage elements within β-globin mRNA might exist as a more general phenomenon. Intriguingly, there exists a biological precedent for a primary ribozyme cleavage downstream of poly(A) sites. In the slime moulds *D. iridis* and *P. polycephalum*[18,19] a complex twin-ribozyme class of group I introns produces a processive autocatalytic cleavage downstream of the poly(A) site of the intron-encoded mRNA. The site of the cleavage does not incorporate GTP, generates RNA products with 3′-hydroxy and 5′-phosphate ends, and is required for mRNA processing and poly(A) site recognition, in a similar manner to the CoTC element.

Earlier results identified CoTC as an essential part of transcriptional termination for the human β-globin gene *in vivo*[3]. We show here that this CoTC activity involves autocatalysis. Cleavage associated with recognition of the poly(A) signal does not promote termination in the absence of CoTC[3]. Previous studies found that poly(A) site cleavage was sufficient to induce termination only when a downstream pause site was present[20]. This suggests that poly(A) site cleavage, unlike CoTC, is kinetically slow. CoTC, but not for example a hammerhead ribozyme[21], generates 5′-phosphate and 3′-hydroxy ends at the site of cleavage. These are preferred substrates for Xrn2 and the 3′ → 5′ exoRNases within the exosome complex[22–24]. Indeed, the AU-rich stem of the CoTC element, essential for autocatalytic cleavage, might be related to the AU-rich element (ARE)[25] because both RNAs can efficiently and specifically recruit the exosome in a cell-free ATP-dependent RNA decay system (A.A., unpublished data). At the same time, the 3′ RNA cleavage product generated by CoTC activity is subjected to 5′ → 3′ exonuclease degradation by Xrn2, which in turn promotes transcriptional termination[26].

We suggest that regulated autocatalytic cleavage elements might have a more general function in RNA metabolism. Judiciously positioned, these elements might provide specific entry sites for distinct exoRNases, with direct implications for transcriptional termination, intron degradation and mRNA turnover[22]. □

Methods

Constructs

All DNA constructs were made by polymerase chain reaction (PCR) amplification of the relevant DNA by *Pfu* Turbo polymerase (Stratagene). Sequence coordinates of the CoTC element refer to the region 64,568–65,370 (GenBank accession no. U01317). For $p\beta 8_{191}$, the minimal 191 base pairs of the ribozyme were amplified by PCR with oligos B85′ and R_{200}. This was cloned blunt-ended into a vector prepared by PCR amplification of βΔ5–7[3] with oligos VR and VF (this vector will be described as βΔT). For $p\beta 38_{191}$ an insert was generated with oligos B85′ and R_{200}ClaI containing a *Cla*I restriction site at the 3′ end. This insert was cloned blunt-ended into a vector prepared by PCR amplification of $p\beta 8_{191}$ with oligos R_{200} and VF2. The resulting construct was subject to digestion with *Cla*I followed by insertion of another copy of the B85′/R_{200}ClaI-generated insert. pSKO1 was made by the insertion of a SKO5′/SKO3′-generated PCR product (coordinates 64,634–64,723) into βΔT. pSKO2 was made by inserting a SKO5′/SKO3′ PCR product into a vector prepared by PCR amplification of pSKO1 with oligos SKO3′ and VF2. The βΔ5–10 plasmid[3] and Tat plasmid[27] have been described previously. The CoTC element was cloned from the construct[5] instead of the original sequence between the G-less cassettes. Templates for the CoTC RNA and its mutants were generated by PCR with T7 promoter at the 5′ end of the upstream primer. Point mutants were made with the QuikChange site-directed mutagenesis kit (Stratagene).

Transcription *in vitro*

Reconstituted transcription was performed as described[4] in the presence of 5 μg of nuclear extracts. After transcription and optional treatment with RNase T1 (10 U for 15 min), transcripts were extracted with phenol/chloroform, precipitated, and resolved on 5% denaturing acrylamide gels. HeLa nuclear extract was purchased from the C4 Cell Culture Center (Belgium).

RNA production and purification

CoTC RNA was transcribed from the PCR templates with T7 polymerase in the presence of lower concentrations of Mg^{2+} and NTP to decrease autocleavage. Transcripts were gel-purified, phenol-extracted and precipitated. Tryptic digestion was performed before phenol extraction as indicated.

Cleavage

The reactions were conducted at 37 °C in the presence of Tris, Mg^{2+}, KCl or NH_2SO_4 and GTP as indicated. The reactions were stopped with EDTA and 80% formamide and the products of cleavage were resolved on 8 M urea gels.

S1, T1, V1 and chemical modification analyses

All reactions were conducted as described in ref. 28.

Transfection and run-ons

Subconfluent HeLa cells were transiently transfected with 20 μg of test plasmid and 3 μg of the Tat plasmid with the use of Lipofectamine 2000 (Invitrogen) in accordance with the manufacturer's guidelines. Probes β3 and β4 as well as P and U3 (ref. 29) have been described previously. The A singlestranded DNA NRO probe was prepared by the insertion of an *Aat*II/*Xmn*I digestion of βΔ5–7 into M13mp19 digested with *Hin*cII. NRO analysis was performed as described previously[30].

Bioinformatics

The folding algorithm[13] was modified to incorporate data from the enzymatic and chemical modification analyses. Sequence alignments were performed on the NIH server (www.ncbi.nlm.nih.gov).

Received 28 June; accepted 17 September 2004; doi:10.1038/nature03032.

1. Greger, I. H., Aranda, A. & Proudfoot, N. J. Balancing transcriptional interference and initiation on the *GAL7* promoter of *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **97,** 8415–8420 (2000).
2. Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C. & Hampsey, M. Ssu72 is an RNA polymerase II CTD phosphatase. *Mol. Cell* **14,** 387–394 (2004).
3. Dye, M. J. & Proudfoot, N. J. Multiple transcript cleavage precedes polymerase release in transcription by RNA polymerase II. *Cell* **105,** 669–681 (2001).
4. Kwek, K. Y. *et al.* U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nature Struct. Biol.* **9,** 800–805 (2002).
5. Lee, M. J. & Greenleaf, L. A. Modulation of RNA polymerase II elongation efficiency by C-terminal heptapeptide repeat domain kinase I. *J. Biol. Chem.* **272,** 10990–10993 (1997).
6. Lilley, D. M. J. The origins of RNA catalysis in ribozymes. *Trends Biochem. Sci.* **28,** 495–501 (2003).
7. Cech, T. R., Zaug, A. J. & Grabowski, P. J. *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **27,** 487–496 (1981).
8. Doherty, E. & Doudna, J. Ribozyme structures and mechanisms. *Annu. Rev. Biochem.* **69,** 597–615 (2000).
9. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428,** 281–286 (2004).
10. Li, Y. & Breaker, R. R. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2′-hydroxyl group. *J. Am. Chem. Soc.* **121,** 5364–5372 (1999).
11. Emilsson, G. M., Nakamura, S., Roth, A. & Breaker, R. R. Ribozyme speed limits. *RNA* **9,** 907–918 (2003).
12. Zuker, M. Mfold web server for nucleic acid folding and hybridisation prediction. *Nucleic Acids Res.* **31,** 3406–3415 (2003).
13. Winkler, W. C. & Breaker, R. R. Genetic control by metabolite-binding riboswitches. *Chembiochem* **4,** 1024–1032 (2003).
14. Cech, T. R. The chemistry of self-splicing RNA and RNA enzymes. *Science* **236,** 1532–1539 (1987).
15. Bass, B. L. & Cech, T. R. Specific interaction between the self-splicing RNA of *Tetrahynema* and its guanosine substrate: implications for biological catalysis by RNA. *Nature* **308,** 820–826 (1984).
16. Stage, K. T. & Uhlenbeck, O. C. Hammerhead ribozyme kinetics. *RNA* **4,** 875–889 (1998).

17. Down, T. A. & Hubbard, J. P. Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res.* **12,** 458–461 (2002).
18. Vader, A., Nielsen, H. & Johansen, S. *In vivo* expression of the nucleolar group I intron-encoded I-dirI homing endonuclease involves the removal of a spliceosomal intron. *EMBO J.* **18,** 1003–1013 (1999).
19. Rocheleau, G. A. & Woodson, S. A. Requirements for self-splicing of a group I intron from *Physarum polycephalum*. *Nucleic Acids Res.* **22,** 4315–4320 (1994).
20. Yonaha, M. & Proudfoot, N. J. Transcriptional termination and coupled polyadenylation *in vitro*. *EMBO J.* **19,** 3770–3777 (2000).
21. Samarsky, D. A. *et al.* A small nucleolar RNA: ribozyme hybrid cleaves a nucleolar RNA target *in vivo* with near-perfect efficiency. *Proc. Natl Acad. Sci. USA* **96,** 6609–6614 (1999).
22. Parker, R. & Song, H. The enzymes and control of eukaryotic mRNA turnover. *Nature Struct. Mol. Biol.* **11,** 121–127 (2004).
23. Bousquet-Antonelli, C., Presutti, C. & Tollervey, D. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102,** 765–775 (2000).
24. van Hoof, V. A. & Parker, R. Messenger RNA degradation: beginning at the end. *Curr. Biol.* **12,** 285–287 (2002).
25. Chen, C. *et al.* AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107,** 451–464 (2001).
26. West, S., Gromak, N. & Proudfoot, N. J. Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* doi:10.1038/nature03035 (this issue).
27. Adams, S. E. *et al.* Synthesis of a gene for the HIV transactivator protein Tat by a novel single-stranded approach using gap repair. *Nucleic Acids Res.* **15,** 4287–4287 (1988).
28. Tahiri-Alaoui, A. *et al.* High affinity nucleic acid aptamers for streptavidin incorporated into bi-specific capture ligands. *Nucleic Acids Res.* **30,** 1–9 (2002).
29. Dye, M. J. & Proudfoot, N. J. Terminal exon definition occurs co-transcriptionally and promotes termination of RNA polymerase II. *Mol. Cell* **3,** 371–378 (1999).
30. Ashe, H. L., Monks, J., Wijgerde, M., Fraser, P. & Proudfoot, N. J. Intergenic transcription and transinduction of the human β-globin locus. *Genes Dev.* **11,** 2494–2509 (1997).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank T. Nilsen, J. Manley, S. Valadkhan, C. Smith and P. C. Branco for their critical comments and constructive support. This work was supported by grants to W.J. from BBSRC and Edward P. Abraham Research Fund, to N.J.P. from the Wellcome Trust and to A.A. from the Wellcome Trust Career Development Programme, Medical Research Council, Cancer Research UK, Edward P. Abraham Research Fund and Exeter College (Oxford).

Authors' contributions A.T. is the lead author. A.T.-A. and S.W. are second authors, and contributed equally to this work. A.T.-A. and W.J. are responsible for the RNA secondary structure analysis. S.W. and N.J.P. are responsible for making the CoTC mutants and testing them by nuclear run-on analysis.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.A. (alexandre.akoulitchev@path.ox.ac.uk).