The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II

Minkyu Kim[1], Nevan J. Krogan[2], Lidia Vasiljeva[1], Oliver J. Rando[3], Eduard Nedea[2], Jack F. Greenblatt[2] & Stephen Buratowski[1]

[1]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA*
[2]*Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario, Canada M5G 1L6*
[3]*Bauer Center for Genomics Research, Harvard University, 7 Divinity Ave., Cambridge, Massachusetts 02138, USA*

The carboxy-terminal domain (CTD) of the RNA polymerase II (RNApII) largest subunit consists of multiple heptapeptide repeats with the consensus sequence YSPTSPS. Different CTD phosphorylation patterns act as recognition sites for the binding of various messenger RNA processing factors, thereby coupling transcription and mRNA processing[1]. Polyadenylation factors are co-transcriptionally recruited by phosphorylation of CTD serine 2 (ref. 2) and these factors are also required for transcription termination[3,4]. RNApII transcribes past the poly(A) site, the RNA is cleaved by the polyadenylation machinery, and the RNA downstream of the cleavage site is degraded. Here we show that Rtt103 and the Rat1/Rai1 5′ → 3′ exonuclease are localized at 3′ ends of protein coding genes. In *rat1-1* or *rai1Δ* cells, RNA 3′ to polyadenylation sites is greatly stabilized and termination defects are seen at many genes. These findings support a model in which poly(A) site cleavage and subsequent degradation of the 3′-downstream RNA by Rat1 trigger transcription termination[5,6].

CTD peptides, consisting of four heptapeptide repeats with different phosphorylation states, were immobilized on beads and used to isolate novel CTD-binding proteins by affinity chromatography of yeast whole-cell extracts. A 60-kDa band that specifically bound to CTD phosphorylated at serine 2 (ser2-P) was identified as Rtt103/YDR289c by mass spectrometry (data not shown). Specific binding of haemagglutinin (HA)-tagged Rtt103 to the ser2-P CTD peptide was verified by immunoblotting (Fig. 1a). Rtt103 was originally isolated as a regulator of Ty1 transposition, because Ty1 RNA synthesis and transposition increase slightly in a *rtt103Δ* strain[7]. Sequence analysis of Rtt103 predicts that it has an RPR domain (also known as a CTD-interacting domain or CID) present in several proteins involved in regulation of nuclear pre-mRNA[8]. *Saccharomyces cerevisiae* has two other RPR/CID proteins, Pcf11 and Nrd1, both of which are involved in mRNA and small nucleolar RNA (snoRNA) 3′-end processing and both of which interact with phosphorylated CTD (ref. 9 and references therein). Rtt103 has homologues in higher eukaryotes, and a mouse homologue (BC021395/Q8VD54) was found in a phospho-CTD immunoprecipitation (S. McCracken and B. Blencowe, personal communication).

An Rtt103 deletion strain is viable. Rtt103 function was probed by synthetic genetic array (SGA) analysis[10]. The *rtt103Δ* strain was crossed to a set of viable deletion strains selected for their involvement in gene expression, and the resulting double-mutant spores were analysed (data not shown). Synthetic phenotypes were confirmed by tetrad dissections. Genetic interactions suggest a role for Rtt103 in 3′-end processing. Synthetic lethality of *rtt103Δ* was observed when combined with a deletion of *CTK1* and synthetic slow growth with a *CTK2* deletion. Ctk1 and Ctk2 are components of CTDK-1 (ref. 11), the kinase responsible for CTD ser2 phosphorylation and co-transcriptional recruitment of polyadenylation factors[2,12]. In addition, a double knockout of *REF2* and *RTT103* is non-viable. Ref2 is involved in snoRNA and mRNA 3′-end formation[13,14]. Synthetic slow growth phenotypes were seen with mutants of several transcription factors implicated in elongation or 3′ processing. These include Spt4, the Elongator complex (ELP1-6), members of the Paf complex, Bur2, Htz1 and its assembly complex SWR-C, and several ubiquitin/proteasome-related proteins. Finally, synthetic lethality was observed with a deletion of *RAI1*, which encodes a component of an RNA degradation complex (see below).

Tandem affinity purification (TAP) of Rtt103 was used to identify associated proteins. In addition to the expected subunits of RNApII, the Rat1 and Rai1 proteins co-purified with Rtt103 (Fig. 1b). Rat1 (also named Xrn2 and Hke1) is a nuclear 5′ → 3′ exoRNase involved in 5.8S RNA and snoRNA 5′ trimming[15,16]. Rai1 is a non-essential protein that co-purifies with Rat1 and enhances Rat1 activity *in vitro*[17]. Rat1/Rai1 has not previously been directly implicated in mRNA processing, but nuclear export of poly(A)$^+$ RNA is blocked in *rat1-1* temperature-sensitive cells at the non-permissive temperature[18]. This nuclear retention phenotype is also seen in cells mutated in some polyadenylation factors[19]. Interestingly, Rtt103 was also isolated in TAP purifications of cleavage factor Pcf11 when washing was performed at lower stringency, but not in purifications of Rna15 or other polyadenylation factors (Fig. 1c, ref. 14, and data not shown).

To determine whether Rtt103, Rat1 and Rai1 associate with transcription elongation complexes, chromatin immunoprecipitation (ChIP) was used to monitor their positions on the *ADH1*, *PMA1* and *PYK1* genes (Fig. 2a–c). Rtt103 crosslinked very strongly near 3′ ends of genes, a pattern similar to that of polyadenylation factors such as Rna14, Rna15 and Pcf11 (refs 2, 4). Consistent with its ser2-P CTD binding and interactions with polyadenylation

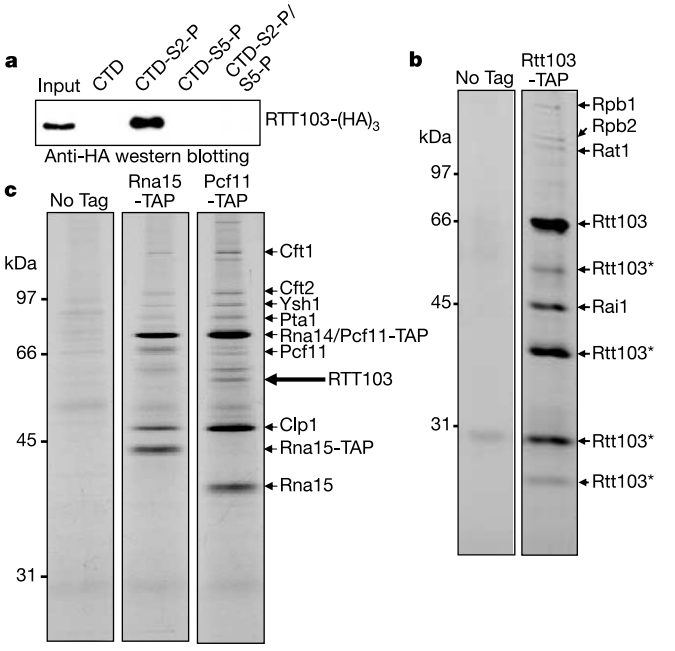

Figure 1 Isolation of a complex containing Rtt103, RNA polymerase II and the Rat1/Rai1 exonuclease. **a**, Rtt103 specifically interacts with CTD phosphorylated at serine 2. Yeast extracts from a triple haemagglutinin $(HA)_3$-tagged Rtt103 strain were incubated with beads carrying CTD peptides with the indicated phosphorylations. After washing, bound Rtt103 was monitored by immunoblot analysis with anti-HA antibody (12CA5). **b**, Purification of proteins associated with TAP-tagged Rtt103. Proteins were analysed by SDS–PAGE and silver staining, and identified by mass spectrometry. Asterisks indicate degradation products of Rtt103. **c**, Purification and mass spectroscopy analysis of TAP-tagged Rna15 and Pcf11 were performed as in **b**.

factors was our failure to observe recruitment of Rtt103 to *PMA1* or other genes in a *ctk1Δ* strain (Fig. 2d; data not shown). Crosslinking of Rat1 and Rai1 was also strongest at 3′ ends of genes, indicating a possible role in 3′-end processing. Rat1 and Rai1 also showed weaker crosslinking to promoter and coding regions. Interestingly, Rat1/Rai1 crosslinking was not dependent on Ctk1. Furthermore, Rat1 crosslinking was the same in *rtt103Δ* or *rai1Δ* strains (data not shown). Therefore, either Rat1 can intrinsically recognize cleaved mRNA 5′ ends or there are redundant mechanisms for targeting the nuclease. Redundancy could explain the synthetic lethality observed between *RAI1* and *RTT103* deletions.

What is the role of Rat1/Rai1 on mRNA genes? At 5′ ends and coding regions, Rat1 might act to degrade any uncapped mRNAs. This would be one of many 'quality control' surveillance mechanisms that prevent incomplete mRNAs from being transported and translated[20]. In support of this idea, Rat1/Rai1 contributes to increased nuclear mRNA degradation[21,22]. At 3′ ends of genes, RNApII transcribes past polyadenylation sites. Cleavage generates a new, uncapped 5′-RNA end on the downstream RNA. A coupled *in vitro* termination/polyadenylation system showed rapid degradation of this downstream RNA, which does not become part of the mature transcript[23]. Because Rat1 is a nuclear 5′ → 3′ exonuclease, we suspected it might be responsible for the degradation of the downstream RNA.

To test this possibility, total RNA was reverse transcribed with a primer complementary to the 3′-downstream region of the *ADH1* gene (Fig. 3a). This primer can produce two complementary DNAs: one from uncleaved pre-mRNA transcripts and one from the downstream RNA after cleavage. Polymerase chain reaction (PCR) was then performed with three primer pairs to generate products 1, 2 and 3, respectively (Fig. 3a). In wild-type cells, the amounts of the three PCR products were similar, indicating that most of the signal was from uncleaved pre-mRNA (Fig. 3b, left panel). In contrast, *rat1-1* cells at both the permissive and non-permissive temperatures had 10–20-fold higher levels of PCR product 3 (Fig. 3b, middle panel). Note that *rat1-1* is also partly defective in its snoRNA-processing functions at permissive temperatures[15]. Little increase was seen in products 1 and 2, suggesting that the increased product 3 is due to stabilization of the downstream RNA and not defects in mRNA 3′ cleavage. Rai1 enhances Rat1 activity *in vitro*[17] and a *rai1Δ* strain also showed stabilization of downstream RNA (Fig. 3b, right panel). A smaller but reproducible increase of product 3 was also seen with *rtt103Δ*, although some increase in product 2 was also observed. Similar results were obtained for the *PMA1* gene (data not shown). These results indicate that Rat1, aided by Rai1 and perhaps Rtt103, is the enzyme responsible for degrading the 3′-downstream RNA generated after poly(A) site cleavage.

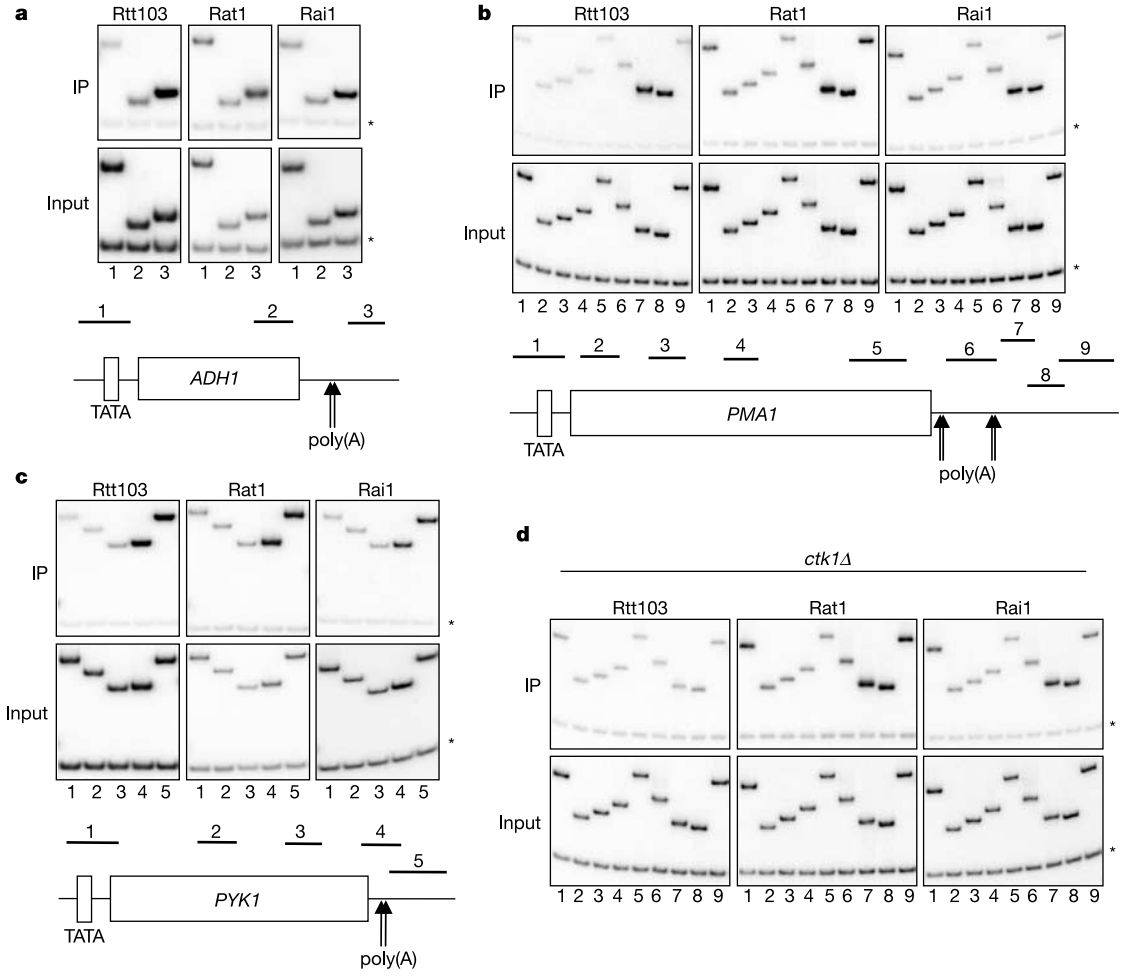

Figure 2 Rtt103 and Rat1/Rai1 localize at 3′ ends of genes. ChIP was performed with strains containing TAP-tagged Rtt103, Rat1 or Rai1. Sheared chromatin was precipitated with IgG-agarose and then amplified by PCR (upper panel) with primers as shown in the diagrams at the bottom. The upper band is the gene-specific band; the common lower band (asterisk) is an internal background control from a non-transcribed region on chromosome V. The middle panels show input controls used to normalize the PCR amplification efficiency of each primer pair. **a**–**c**, Analysis was performed on *ADH1* (**a**) *PMA1* (**b**) and *PYK1* (**c**). **d**, ChIP analysis on the *PMA1* gene in a *ctk1Δ* strain. IP, immunoprecipitation.

The *RAT1* gene was initially isolated in a screen for mutants defective in poly(A)$^+$ RNA transport out of the nucleus[18]. Many polyadenylation factor mutants exhibit a similar nuclear retention phenotype[19]. Extracts from a *rat1-1* mutant strain were assayed for transcript cleavage and polyadenylation *in vitro*, but no defects were observed (Fig. 3c and Supplementary Fig. 1a). To determine whether Rat1/Rai1 instead has a function in transcription termination, ChIP analysis of RNApII was performed in *rai1Δ* and *rat1-1* strains. In wild-type cells, Rpb3 (that is, RNApII) density is roughly even throughout the transcribed region of the *PMA1* gene, but decreases to near background about 200–400 base pairs downstream of the polyadenylation site (Fig. 4a, see primers 8 and 9). However, in *rat1-1* cells at both 23 °C and 37 °C, RNApII density did not decrease, indicating a possible role for Rat1 in termination. Cells lacking Rai1 were similar to *rat1-1* cells at 23 °C, but *rtt103Δ* cells had no obvious termination defect (Fig. 4b).

To see how general the 3′ effect was, immunoprecipitated chromatin was hybridized to a tiled oligonucleotide microarray covering *S. cerevisiae* chromosome III. Input and anti-Rpb3-immunoprecipitated DNAs were amplified, labelled with Cy3 or Cy5 and hybridized to the arrays. In *rat1-1* cells, most genes in which termination could be seen exhibited a 3′ extension of Rpb3 crosslinking relative to that observed in wild-type cells. Two such genes are shown in Fig. 4c. The effect was most obvious on strongly transcribed genes well separated from other strong transcription units. For example, the *PMP1* gene (transcription rate 38.7 mRNA h^{-1}; see ref. 24), which is upstream of *YCR024C* (0.1 mRNA h^{-1}), shows a steep peak of Rpb3 crosslinking in wild-type cells. In the *rai1Δ* strain, RNApII was seen much further downstream, declining at a more gradual rate. The *rat1-1* strain also showed extended 3′-RNApII crosslinking at 23 °C. At 37 °C, transcription in this strain was markedly extended several kilobases farther downstream (Fig. 4c). Thus, Rat1 probably has a function in transcription termination at multiple loci throughout the genome.

Two experiments were performed to distinguish whether the Rat1 exonuclease itself, or just the presence of the protein, was required for termination. First, immunoblotting of the Rat1-1 protein showed that it was stable at all temperatures (Supplementary Fig. 1b), showing that the termination defect was not simply due to an absence of Rat1. Next, a point mutant of Rat1 was created (D235A) in a critical residue that is absolutely conserved in this nuclease family[25]. The mutant produced normal levels of Rat1

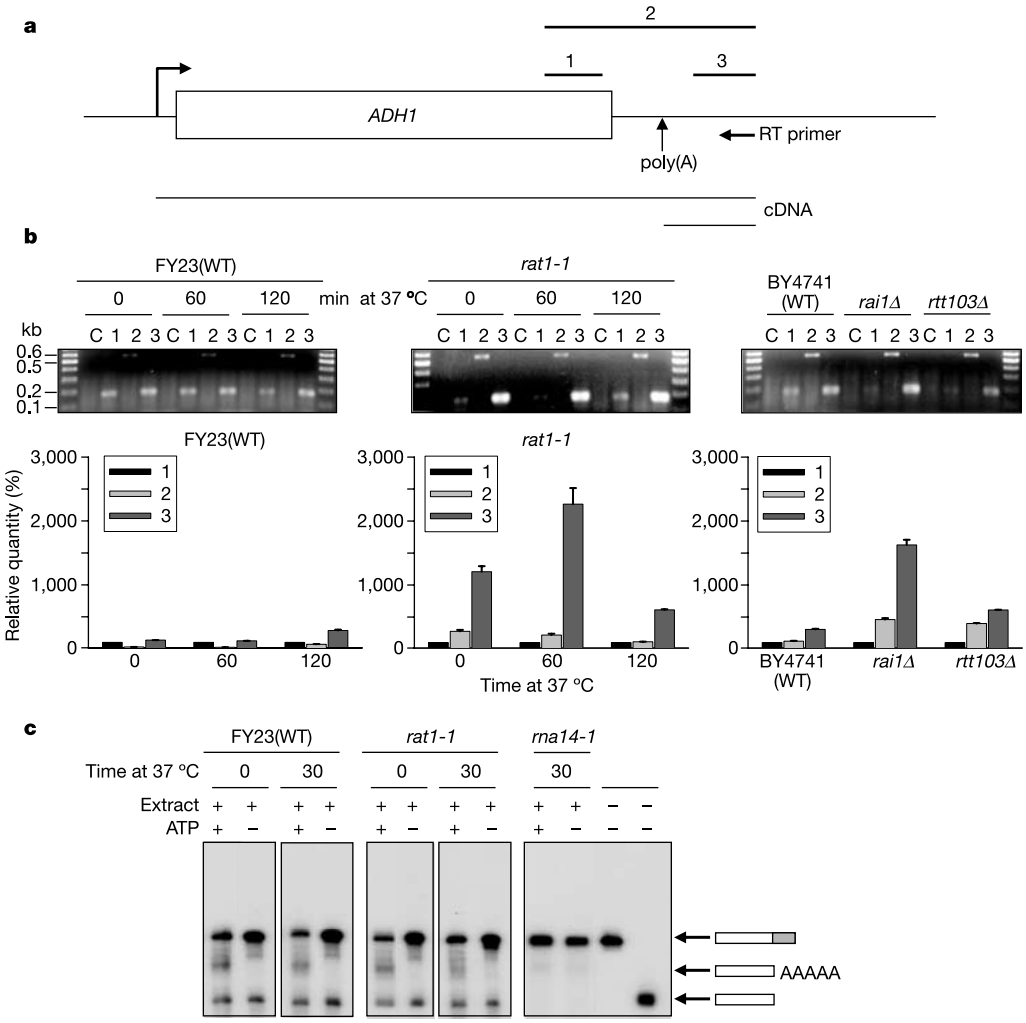

Figure 3 Rat1 degrades 3′-downstream RNA but is not required for mRNA cleavage or polyadenylation. **a**, Schematic diagram of *ADH1*. Lines below the gene indicate potential cDNAs. Bars above denote RT–PCR products shown in **b**. **b**, RT–PCR with RNA from wild-type (WT, left panel) and *rat1-1* cells (middle) shifted to 37 °C for the indicated duration. A similar analysis of *rai1Δ*, *rtt103Δ* and WT cells is shown at the right. C is a no-RT control. For quantification (below each gel), product 1 was set to 100% in each group. Results are means + s.d. for three repetitions. Black bars, product 1; light grey bars, product 2; dark grey bars, product 3. **c**, *In vitro* cleavage/polyadenylation assays were performed with extracts from the indicated strains. The last two lanes show standards for precursor RNA and cleavage product. The *rna14-1* extract was competent for polyadenylation of pre-cleaved product (Supplementary Fig. 1a).

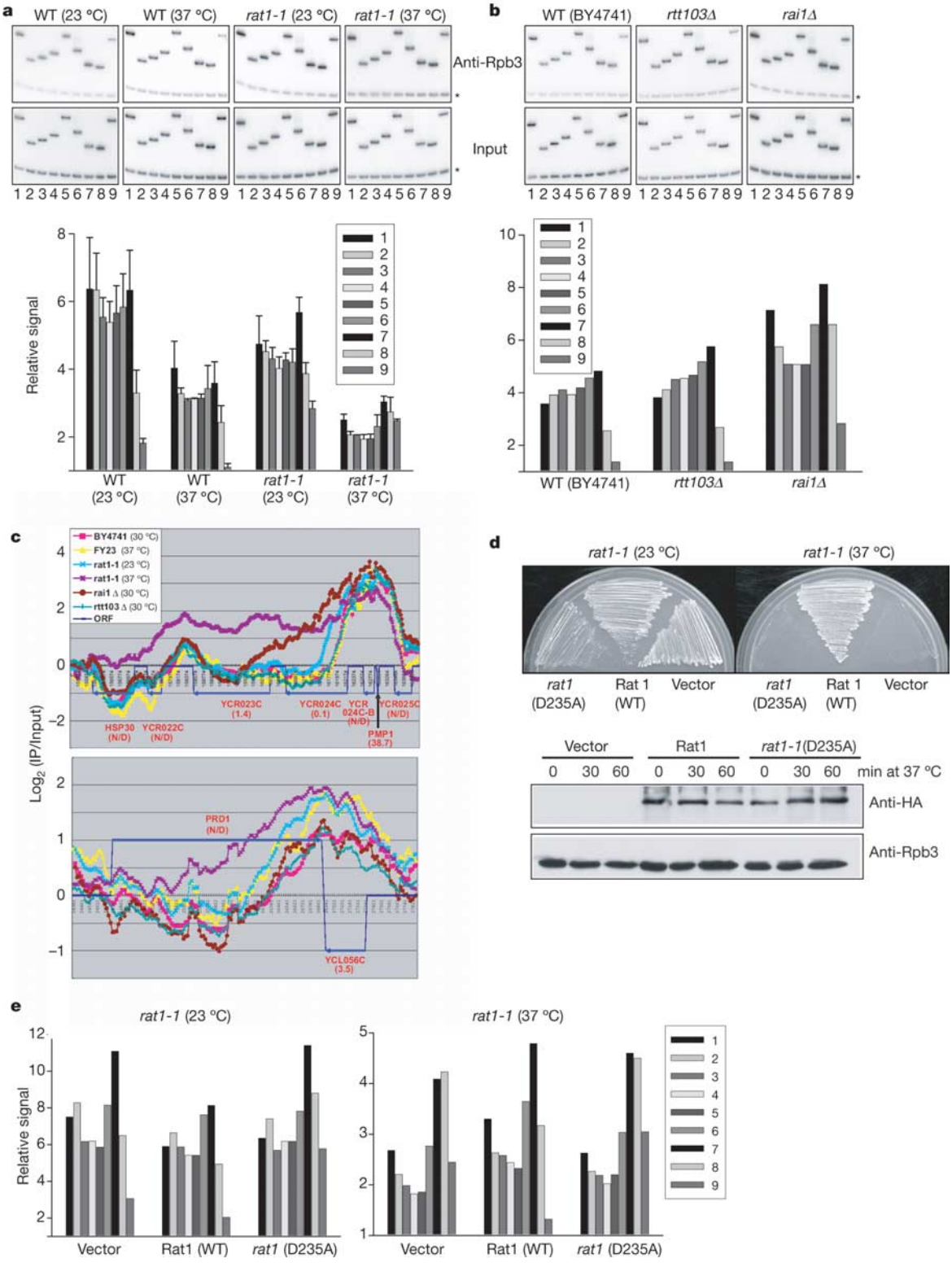

Figure 4 Rat1/Rai1 promotes transcription termination. **a**, ChIP analysis of Rpb3 on *PMA1* in wild-type (WT) and *rat1-1* strains grown at the indicated temperature. PCR products are shown as in Fig. 2b. Quantification is below each gel. The *y*-axis shows the ratio of *PMA1* signal relative to the internal negative control. A value of 1.0 therefore indicates no signal above background. Error bars show standard deviation from three repetitions. **b**, The same analysis with isogenic WT, *rai1Δ* and *rtt103Δ* strains. **c**, Chromatin immunoprecipitated with anti-Rpb3 was analysed using microarrays as described in Methods. Two representive genes are displayed here. The *y*-axis is $\log_2$ of the immunoprecipitation (IP) to Input ratio, and the *x*-axis represents position along the chromosome. Dark blue arrows denote open reading frames (ORFs). Note that, for the genes of interest, transcription is right to left. The numbers in parentheses show reported transcription rates (mRNA h^{-1}), with N/D being below the detection limit[24]. **d**, Rat1(D235A) does not support viability. Plasmids containing no insert, WT *RAT1* or the mutant D235A were transformed into a *rat1-1* strain and tested for growth at 23 and 37 °C. Below the plates are immunoblots showing expression of the tagged wild-type and D235A Rat1 proteins. **e**, Wild-type RAT1 gene restores termination, whereas the D235A mutant does not. Cells from **d** were used for ChIP experiments as in **a**.

protein but was inactive *in vitro* (Supplementary Fig. 2) and failed to complement the temperature sensitivity of the *rat1-1* strain (Fig. 4d). Whereas wild-type *RAT1* corrected the termination defect of the *rat1-1* strain, the D235A mutant did not (Fig. 4e). In fact, the point mutant caused slightly slower growth and exacerbated the read-through phenotype at the permissive temperature. The exonuclease activity of Rat1 is therefore required for proper termination.

How does Rat1 contribute to transcription termination? The most obvious mechanism is by degrading the transcript behind elongating RNApII. On reaching the polymerase, Rat1 might trigger transcript release by disrupting polymerase–RNA contacts. This mechanism echoes rho-dependent termination in bacteria. Although rho does not degrade RNA, it tracks along the RNA behind polymerase and triggers termination on reaching the elongation complex. Alternatively, degrading upstream RNA could promote 'backtracking' and arrest by RNApII (ref. 26). Without RNA behind the elongation complex, backtracking polymerases might release from the DNA template.

A cleavage-dependent tracking/termination model has been called the 'torpedo' model[5,6], and our results provide support for this idea. An alternative 'anti-terminator' model proposes that emergence of the polyadenylation site changes the properties of the elongation complex independently of cleavage to trigger termination, perhaps by dissociation of a positive elongation factor or recruitment of a termination factor. It is likely that termination involves aspects of both models. Association of the Paf and TREX/THO elongation factors with transcribing RNApII is reduced on passage through a poly(A) site[2,4], perhaps promoting termination. Polyadenylation factors are recruited to 3′ ends of genes dependent on CTD ser2 phosphorylation and the poly(A) site[2,4], leading to mRNA cleavage and degradation of the downstream RNA by Rat1. The role of Rat1 in termination seems to be conserved over evolution, because an RNAi knockdown of the mammalian Rat1 homologue (*XRN2*) also results in termination defects[27]. □

Methods

Yeast strains

Strains used are listed in Supplementary Table 1.

CTD peptide binding assay

CTD peptides were synthesized with four repeats of YSPTSPS and amino-terminal biotinylation. Unphosphorylated, ser2, ser5 and ser2/ser5 phosphorylated peptides were made. For the identification of ser2-P-specific binding proteins, 3 mg of streptavidin-coated magnetic beads (M-280; Dynal) was resuspended in 450 μl of PBS, mixed with 30 μg of ser2-P peptide and incubated at 4 °C for 90 min. After being washed twice with PBS, the beads were resuspended in 450 μl of PBS, mixed with 3 μg of unphosphorylated CTD peptide and incubated again for 90 min. Beads with the other peptides were prepared similarly. Beads were washed twice with PBS and resuspended in 900 μl of binding buffer (10 mM potassium phosphate pH 7.7, 100 mM potassium acetate, 20 mM magnesium acetate, 5 mM EGTA, 10% glycerol, 0.1% Nonidet P40, 0.05% Triton X-100, 1 mM dithiothreitol, 1 mM NaN_3, 1 mM NaF, 0.4 mM $NaVO_3$) plus protease inhibitors. Yeast whole cell extract (5 mg) was added and the mixture was incubated at 4 °C overnight. After washing of the beads four times with 450 μl of binding buffer, bound proteins were eluted with 40 μl of binding buffer in the presence of 1 M potassium acetate. The eluate was dialysed twice for 1 h against 200 ml of binding buffer at 4 °C, with Millipore type VS 0.025 μm filter and subjected to SDS–polyacrylamide-gel electrophoresis (SDS–PAGE). Protein bands were compared between unphosphorylated and ser2-P-affinity purification, and specific bands were analysed by mass spectrometry. For immunoblotting, protein-bound beads were boiled and loaded on SDS–PAGE gels directly. For Fig. 1a, a similar reaction was performed at a smaller scale with yeast extract from strain YSB815, in which Rtt103 is tagged with HA. Bound proteins were analysed by SDS–PAGE and anti-HA (monoclonal 12CA5) immunoblotting.

Rat1 mutants

The *rat1-1* mutant strain was from C. Cole (Dartmouth University, New Hampshire), and Rat1 plasmids were from A. Johnson (University of Texas, Austin). Rat1 and Rat1-1 were TAP-tagged as described previously[14,28]. The D235A mutant was constructed by PCR mutagenesis and cloned into a vector in which the HA-tagged protein was expressed from the *RAT1* promoter. A similar wild-type construct was made in parallel.

TAP purification

TAP purifications at 100 mM NaCl were performed as described[14,28].

SGA analysis

SGA analysis was performed as described previously[10].

ChIP

ChIP procedures and quantification were performed as described[4]. For the experiment in Fig. 4, both wild-type and *rat1-1* cells were incubated at 23–25 °C until OD_{595} reached 0.8. Equal amounts of medium prewarmed to 51 °C were added and cells were incubated at 37 °C for a further 25–30 min. Formaldehyde crosslinking and chromatin preparation were as described.

Polyadenylation and cleavage assays

RNA substrates containing the Gal7 3′ end were synthesized *in vitro* and assayed as described previously[29].

Reverse transcription (RT)–PCR analysis

Total RNA was isolated by extraction with hot phenol from isogenic FY23(WT) and *rat1-1* cells at 0, 60 and 120 min after temperature shift to 37 °C, or from isogenic BY4741(WT), *Δrtt103* and *Δrai1* cells at 30 °C. About 3 μg of total RNA was reverse transcribed. Priming was with 10 pmol of an anti-sense oligonucleotide spanning the 3′-downstream region of the *ADH1* gene; 5′-CCCAACTGAAGGCTAGGCTGTGG-3′. One-tenth of the reverse-transcribed cDNA products were then amplified by PCR. RT–PCR products were subjected to electrophoresis on 1.2% agarose gel and revealed by ethidium bromide staining. For each set of RT–PCR reactions, the relative intensities of PCR products 2 and 3 were calculated by setting that of PCR product 1 arbitrarily to 100%.

Microarray experiments

Formaldehyde crosslinking and ChIP were performed as above. Immunoprecipitated DNAs were amplified as described in www.microarrays.org/pdfs/Round_A_B_C.pdf with some modifications. Control and immunoprecipitated DNAs from 30 ml of cells were randomly tagged with Sequenase and primer A (5′-GTTTCCCAGTCACGATCNNNNNNNNN-3′) and then further amplified by PCR with primer B (5′-GTTTCCCAGTCACGATC-3′). Finally, 2 μg of amplified DNA was labelled by Klenow fragment in the presence of Cy dye-coupled dCTP and random primers (BioPrime labelling kit; Invitrogen). Input and immunoprecipitated DNA were labelled with different dyes for competitive hybridization to spotted microarrays[30] consisting of 50-base oligonucleotides spanning *S. cerevisiae* chromosome III (O.J.R., unpublished), tiled every 20 base pairs. Hybridization was at 55 °C for 4 h in buffer containing 3.4 × SSC and 0.3% SDS. Arrays were washed sequentially with 1 × SSC/0.03% SDS, 0.2 × SSC and 0.05 × SSC. Microarrays were scanned and fluorescence intensities were quantified with an Axon Genepix 4000B scanner and software. Data in Fig. 4c are from a moving-window average of five probes. Hybridizations with BY4741 and the *rat1-1* were repeated several times, whereas the FY23, *rai1Δ* and *rtt103Δ* strains were tested once. A full analysis of this and other data will be presented elsewhere (M.K., S.B. and O.J.R., unpublished work).

Received 15 June; accepted 17 September 2004; doi:10.1038/nature03041.

1. Buratowski, S. The CTD code. *Nature Struct. Biol.* **10,** 679–680 (2003).
2. Ahn, S. H., Kim, M. & Buratowski, S. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3′ end processing. *Mol. Cell* **13,** 67–76 (2004).
3. Birse, C. E., Minvielle-Sebastia, L., Lee, B. A., Keller, W. & Proudfoot, N. J. Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280,** 298–301 (1998).
4. Kim, M., Ahn, S. H., Krogan, N. J., Greenblatt, J. F. & Buratowski, S. Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. *EMBO J.* **23,** 354–364 (2004).
5. Connelly, S. & Manley, J. L. A CCAAT box sequence in the adenovirus major late promoter functions as part of an RNA polymerase II termination signal. *Cell* **57,** 561–571 (1989).
6. Proudfoot, N. J. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14,** 105–110 (1989).
7. Scholes, D. T., Banerjee, M., Bowen, B. & Curcio, M. J. Multiple regulators of Ty1 transposition in *Saccharomyces cerevisiae* have conserved roles in genome maintenance. *Genetics* **159,** 1449–1465 (2001).
8. Doerks, T., Copley, R. R., Schultz, J., Ponting, C. P. & Bork, P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res.* **12,** 47–56 (2002).
9. Meinhart, A. & Cramer, P. Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. *Nature* **430,** 223–226 (2004).
10. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294,** 2364–2368 (2001).
11. Sterner, D. E., Lee, J. M., Hardin, S. E. & Greenleaf, A. L. The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. *Mol. Cell. Biol.* **15,** 5716–5724 (1995).
12. Skaar, D. A. & Greenleaf, A. L. The RNA polymerase II CTD kinase CTDK-I affects pre-mRNA 3′ cleavage/polyadenylation through the processing component Pti1p. *Mol. Cell* **10,** 1429–1439 (2002).
13. Dheur, S. *et al.* Pti1p and Ref2p found in association with the mRNA 3′ end formation complex direct snoRNA maturation. *EMBO J.* **22,** 2831–2840 (2003).
14. Nedea, E. *et al.* Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3′-ends. *J. Biol. Chem.* **278,** 33000–33010 (2003).
15. Petfalski, E., Dandekar, T., Henry, Y. & Tollervey, D. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* **18,** 1181–1189 (1998).
16. Qu, L. H. *et al.* Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.* **19,** 1144–1158 (1999).
17. Xue, Y. *et al. Saccharomyces cerevisiae* RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol. Cell. Biol.* **20,** 4006–4015 (2000).
18. Amberg, D. C., Goldstein, A. L. & Cole, C. N. Isolation and characterization of RAT1: an essential gene

of *Saccharomyces cerevisiae* required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes Dev.* **6,** 1173–1189 (1992).
19. Hammell, C. M. *et al.* Coupling of termination, 3′ processing, and mRNA export. *Mol. Cell. Biol.* **22,** 6441–6457 (2002).
20. Vasudevan, S. & Peltz, S. W. Nuclear mRNA surveillance. *Curr. Opin. Cell Biol.* **15,** 332–337 (2003).
21. Bousquet-Antonelli, C., Presutti, C. & Tollervey, D. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102,** 765–775 (2000).
22. Das, B., Butler, J. S. & Sherman, F. Degradation of normal mRNA in the nucleus of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **23,** 5502–5515 (2003).
23. Yonaha, M. & Proudfoot, N. J. Transcriptional termination and coupled polyadenylation *in vitro*. *EMBO J.* **19,** 3770–3777 (2000).
24. Holstege, F. C. *et al.* Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95,** 717–728 (1998).
25. Solinger, J. A., Pascolini, D. & Heyer, W. D. Active-site mutations in the Xrn1p exoribonuclease of *Saccharomyces cerevisiae* reveal a specific role in meiosis. *Mol. Cell. Biol.* **19,** 5930–5942 (1999).
26. Ujvari, A., Pal, M. & Luse, D. S. RNA polymerase II transcription complexes may become arrested if the nascent RNA is shortened to less than 50 nucleotides. *J. Biol. Chem.* **277,** 32527–32537 (2002).
27. West, S., Gromak, N. & Proudfoot, N. Human 5′ → 3′ exonuclease XRN2 promotes transcription termination from co-transcriptional cleavage sites. *Nature* doi:10.1038/nature03035 (this issue).
28. Krogan, N. J. *et al.* RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol. Cell. Biol.* **22,** 6979–6992 (2002).
29. Zhao, J., Kessler, M., Helmling, S., O'Connor, J. P. & Moore, C. Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol.* **19,** 7733–7740 (1999).
30. Lashkari, D. A. *et al.* Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl Acad. Sci. USA* **94,** 13057–13062 (1997).

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

Acknowledgements We thank C. Moore and her laboratory for advice on polyadenylation reactions; N. Proudfoot, S. McCracken and B. Blencowe for sharing unpublished information; and the Taplin Mass Spectroscopy facility at Harvard Medical School and D. Richards of Affinium Pharmaceuticals for mass spectroscopy. N.J.K. was supported by a Doctoral Fellowship from the Canadian Institutes of Health Research (CIHR). O.J.R. is supported by funding from the Bauer Center for Genomics Research. This research was supported by grants to S.B. from the US National Institutes of Health, and to J.F.G. from the Canadian Institutes of Health Research, the Ontario Genomics Institute, and the National Cancer Institute of Canada with funds from the Canadian Cancer Society. L.V. is a Fellow and S.B. is a Scholar of the Leukemia and Lymphoma Society.

Authors' contributions All authors contributed to the conception and design of the experiments. M.K. performed the CTD affinity purification of Rtt103, RT–PCR analysis, ChIP analyses and Rat1 mutagenesis. N.K. did the Rtt103-TAP purification and the SGA analysis. L.V. conducted the immunoblotting, exonuclease and polyadenylation assays. E.N. performed the Pcf11 and Rna15 TAP purifications. The microarray experiments were done by M.K. and O.J.R. The paper was written by M.K., J.G. and S.B. with input from the other authors.

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

Correspondence and requests for materials should be addressed to S.B. (steveb@hms.harvard.edu).

Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites

Steven West*, Natalia Gromak* & Nick J. Proudfoot

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

* These authors contributed equally to this work

Eukaryotic protein-encoding genes possess poly(A) signals that define the end of the messenger RNA and mediate downstream transcriptional termination by RNA polymerase II (Pol II)[1]. Termination could occur through an 'anti-termination' mechanism whereby elongation factors dissociate when the poly(A) signal is encountered, producing termination-competent Pol II[2,3]. An alternative 'torpedo' model postulated that poly(A) site cleavage provides an unprotected RNA 5′ end that is degraded by 5′ → 3′ exonuclease activities (torpedoes) and so induces dissociation of Pol II from the DNA template[1,4]. This model has been questioned because unprocessed transcripts read all the way to the site of transcriptional termination before upstream polyadenylation[5–7]. However, nascent transcripts located 1 kilobase downstream of the human β-globin gene poly(A) signal are associated with a co-transcriptional cleavage (CoTC) activity[8] that acts with the poly(A) signal to elicit efficient transcriptional termination. The CoTC sequence is an autocatalytic RNA structure that undergoes rapid self-cleavage[9]. Here we show that CoTC acts as a precursor to termination by presenting a free RNA 5′ end that is recognized by the human 5′ → 3′ exonuclease Xrn2. Degradation of the downstream cleavage product by Xrn2 results in transcriptional termination, as envisaged in the torpedo model.

To confirm the requirement of a CoTC sequence for transcriptional termination we compared transcription of the human β-globin gene (WT), driven by a Tat-inducible HIV-1 promoter, with that of constructs either lacking CoTC (ΔCoTC) or bearing a mutant poly(A) signal (Δp(A)). The plasmids were transfected into HeLa cells and their nascent transcripts analysed by nuclear run-on (NRO) analysis (Fig. 1a). In WT (top panel), termination occurs very efficiently within CoTC as indicated by low signals over probes A (downstream of the CoTC element) and U3 (upstream of the promoter). Termination efficiency is significantly reduced when CoTC is deleted, as indicated by the high signal over probes A and U3 (middle panel), and the mutant p(A) causes an even greater termination defect (bottom panel), as shown previously[8]. In the accompanying study[9] the minimal autocatalytic, or core, CoTC element is also shown to promote Pol II termination by using the same NRO assay. We conclude that efficient transcriptional termination of the human β-globin gene requires CoTC.

The possibility that autocatalytic CoTC provides an entry site for a 5′ → 3′ exonuclease was tested by the depletion of mRNAs encoding such activities using RNA interference technology[10]. The major nuclear 5′ → 3′ exonuclease in humans is Xrn2 (refs 11, 12), a 950-residue protein homologous to *Saccharomyces cerevisiae* Rat1p (ref. 13). Rat1p is known to be involved in 5.8S rRNA processing[14], snoRNA maturation[15] and nuclear–cytoplasmic RNA transport[13]. Analyses by reverse transcriptase-mediated polymerase chain reaction (RT–PCR) and western blotting of Xrn2 mRNA and protein isolated from mock or Xrn2-specific short interfering RNA (siRNA)-treated HeLa cells are presented in Fig. 1b. An 80% decrease in the amount of Xrn2, but not actin, mRNA was observed. Similarly, western blotting with Xrn2 antibody[12] revealed that both forms of Xrn2 were significantly decreased (about fourfold), in comparison with transcription factor IIH (TFIIH). The smaller 70-kDa Xrn2 protein might result from the recognition of an internal methionine initiation codon[16].

To test the effect of Xrn2 depletion on transcriptional termination we performed NRO on Xrn2 knock-down HeLa cells that were transiently transfected with the WT plasmid (Fig. 1c, top graph). A reproducible twofold to threefold increase in transcripts over probes A and U3 in comparison with mock-treated cells indicates a role for Xrn2 in transcriptional termination (Fig. 1c, middle graph). We believe that the observed effect might be an under-representation: some cells will remain termination-competent because Xrn2 knock-down is not 100% efficient. The same experiment was conducted with control (luciferase-specific) siRNA, resulting in an identical profile to that of the mock-treated cells (data not shown). Although the level of transcriptional termination with the WT plasmid was lower in the mock-treated than the untreated cells (compare Fig. 1c with Fig. 1a), we suspect that the multiple transfections necessary for successful RNA-mediated interference (RNAi) might be responsible for this. Taken together, these data show that depletion of Xrn2 results in a decrease in termination efficiency of the β-globin gene. To confirm these data independently, however, we performed chromatin immunoprecipitation (ChIP) analysis with a Pol II-