MOLECULAR AND CELLULAR BIOLOGY, May 2005, p. 4023–4033 Vol. 25, No. 10
0270-7306/05/$08.00+0 doi:10.1128/MCB.25.10.4023–4033.2005

Human hHpr1/p84/Thoc1 Regulates Transcriptional Elongation and Physically Links RNA Polymerase II and RNA Processing Factors

Yanping Li, Xiaoling Wang, Xiaojing Zhang, and David W. Goodrich*

Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, New York 14263

Received 10 September 2004/Returned for modification 20 October 2004/Accepted 23 February 2005

Cotranscriptional loading of RNA processing factors onto nascent RNA facilitates efficient gene expression. Mechanisms responsible for coupling transcription and RNA processing are not well defined, but the *Saccharomyces cerevisiae* TREX complex provides an example. TREX is composed of the subcomplex THO that associates with RNA polymerase II and is required for normal transcriptional elongation. THO associates with proteins involved in RNA splicing and export to form the larger TREX complex. Hence, assembly of TREX physically couples transcriptional elongation with RNA processing factors. Whether metazoan species with long, intron-containing genes utilize a similar mechanism has not been established. Here we show that human hHpr1/p84/Thoc1 associates with elongating RNA polymerase II and the RNA splicing and export factor UAP56 in intact cells. Depletion of hHpr1/p84/Thoc1 causes transcriptional elongation defects and associated cellular phenotypes similar to those observed in THO-deficient yeast. We conclude that hHpr1/p84/Thoc1 regulates transcriptional elongation and may participate in a protein complex functionally analogous to yeast TREX, physically linking elongating RNA polymerase II with RNA processing factors.

Efficient expression of mRNA is facilitated by cotranscriptional loading of RNA processing and export factors onto nascent RNA (22). Molecular mechanisms utilized to couple transcribing RNA polymerase II complexes (RNA PolII) to these factors are incompletely defined, but the *Saccharomyces cerevisiae* TREX complex provides one example. TREX is composed of the THO subcomplex, containing four proteins implicated in transcriptional elongation, Hpr1p, Tho2p, Mft1p, and Thp2p (6). THO associates with two additional proteins implicated in nuclear RNA export, Sub2p and Yra1, to form the larger TREX complex (30). Sub2p is the yeast orthologue of human UAP56 that has also been implicated in RNA splicing (18). THO component Hpr1p genetically and physically interacts with RNA PolII (4, 6, 10, 30) and is essential for recruitment of Sub2p to genes regulated by THO (34). Loss of Hpr1p impairs both transcriptional elongation and nuclear RNA export for a subset of yeast genes (5, 7, 10, 13, 24, 30). Hence, Hpr1p is a critical component of THO and TREX. These observations have engendered the hypothesis that TREX assembly couples the processes of transcription elongation and RNA processing, thereby ensuring efficient, coregulated expression of a subset of yeast genes.

It is currently unknown whether metazoan species with long, intron-containing genes utilize an analogous mechanism to couple transcriptional elongation and RNA processing. Human structural homologues are apparent in the protein databases only for some yeast TREX proteins, including Tho2 (yeast THO component Tho2p), UAP56 (yeast RNA export factor Sub2p), and Aly (yeast RNA export factor Yra1p) (18, 28, 31, 33). In vitro pulldown assays, however, have identified additional human proteins that associate with a glutathione *S*-transferase (GST)-UAP56 fusion protein (30). These proteins may represent additional components of human THO or TREX complexes. It is unknown whether any of these putative human THO/TREX proteins regulate transcriptional elongation, whether they associate with each other in intact cells, or whether they bind elongating RNA PolII.

One of the proteins interacting with GST-UAP56 has previously been discovered based on its ability to associate with the retinoblastoma tumor suppressor protein (Rb1) and originally named p84 (9). The gene and its encoded protein are also referred to as hHpr1 and p84N5. The official symbol as annotated in the human genome database is Thoc1. Here we will refer to the gene as Thoc1 and its encoded protein as pThoc1. While pThoc1 does not show overall, statistically significant, primary amino acid sequence similarity to known yeast TREX proteins in BLASTP alignments, its primary sequence and predicted molecular mass most closely match those of yeast THO component Hpr1p. In this study, we explore the hypothesis that pThoc1 is a functional orthologue of yeast Hpr1p. Consistent with this hypothesis pThoc1 binds UAP56, human Tho2, and elongating forms of RNA PolII in intact cells. Depletion of pThoc1 compromises the expression of some genes due to a molecular defect in transcriptional elongation similar to that observed in Hpr1p-deficient yeast. Other cellular phenotypes previously associated with loss of Hpr1p, including slow growth, increased sensitivity to DNA damage, and defects in nucleotide excision repair (NER), are also observed in pThoc1-depleted human cells. We conclude that pThoc1 is required for the normal transcriptional elongation of some human genes and physically links transcribing RNA PolII to the RNA splicing and export factor UAP56. We suggest that pThoc1 is a human functional orthologue of yeast Hpr1p and participates in a human protein complex analogous to yeast TREX, coupling the processes of transcriptional elongation and RNA processing.

* Corresponding author. Mailing address: Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263. Phone: (716) 845-4506. Fax: (716) 845-8857. E-mail: david.goodrich@roswellpark.org.

MATERIALS AND METHODS

Cell culture, plasmids, and adenovirus. Cell lines were obtained from the American Type Culture Collection (Manassas, VA). HeLa cells with stably integrated copies of the UHG16-3 plasmid were provided by Manfred Gossen (Max Delbruck Center for Molecular Medicine). UHG16-3 expresses the bacterial *lacZ* gene under control of a tetracycline-regulatable promoter (Tet-*lacZ*). Cells were cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 0.5% penicillin and streptomycin and grown at 37°C in a humidified atmosphere of 5% CO_2.

Cloning of the pThoc1 full-length cDNA was described previously (8, 9). The UAP56 and Tho2 cDNAs were cloned by PCR amplification from total HeLa cell cDNA generated using the SMART cDNA synthesis procedure (Clontech, Palo Alto, CA). The cDNAs were subcloned into pDNR-CMV (Clontech, Palo Alto, CA) or pCEP4 (Invitrogen, Carlsbad, CA), and the coding DNA sequences were verified against published sequences. The *pThoc1 expression plasmids were constructed by mutating the Thoc1 cDNA to introduce silent mutations in codons 472 and 473 of the pThoc1 cDNA (wild-type GAA CAG to GA**G** CA**A**) by the QuikChange method (Stratagene, La Jolla, CA). For expression of green fluorescent protein (GFP) fusion proteins, the cDNAs were transferred to pLP-EGFP-C1 (Clontech, Palo Alto, CA) by in vitro *cre*-mediated recombination. Plasmids were introduced into cultured cells using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations.

Adenoviral stocks were obtained from the University of Texas M. D. Anderson Cancer Center Vector Core Facility and used as recommended. Expression of the cytomegalovirus (CMV)-*lacZ* adenoviral gene is driven by the early promoter of CMV. Expression of the GV16-*lacZ* gene is driven by a promoter responsive to a GAL4-VP16 fusion protein expressed by the Ad/PKG-GV16 adenovirus. Infections were typically performed at a multiplicity of infection of 25 to 50 IU per cell. For the UV reactivation assay, Ad/CMV-GFP was UV irradiated (0.1 to 0.3 J/m^2) before infection.

siRNAs. Small interfering RNA (siRNA) oligonucleotides were introduced into cultured cells by transfection with Lipofectamine 2000 as recommended by the manufacturer (Invitrogen, Carlsbad, CA). The concentration of siRNA duplexes used was 10 to 50 nM. The siRNA duplexes directed against pThoc1 were N52 (5′GCCAUUGAACAGGCAGACCdTdT and 5′GGUCUGCCUGUUCA AUGGCdTdT) and N54 (5′CACAUCCUGUUGCAGUAUCdTdT and 5′GA UACUGCAACAGGAUGUGdTdT). The siRNA duplexes directed against UAP56 were UAP562 (5′GUCACACUCGGGAGUUGGCdTdT and 5′GCCA ACUCCCGAGUGUGACdTdT) and UAP564 (5′AGAGAUCCGUCCAGUC UGCdTdT and 5′GCAGACUGGACGGAUCUCUdTdT). The Tho2 siRNA duplexes were Tho21 (5′ACACUGGAAUCAUUAGGGCdTdT and 5′GCCC UAAUGAUUCCAGUGUdTdT) and Tho22 (5′GCCAGUCGAUACGGAAG GUdTdT and 5′ACCUUCCGUAUCGACUGGCdTdT). The control N52 M siRNA targets the same region as N52 but contains mismatches to the wild-type sequence (5′GCCAUU**AAG**CAGG**AC**GACCdTdT and 5″GGUC**GU**CCUG**CU** **U**AAUGGCdTdT). The nonspecific control siRNA duplex (SC) contained scrambled sequence from the firefly luciferase gene (Dharmacon, Lafayette, CO).

Growth and gene expression assays. Cells were extracted, and β-galactosidase activity was measured by *o*-nitrophenyl-β-D-galactopyranoside (ONPG) assay according to the manufacturer's instructions (Promega, Madison, WI). Optical density of samples was measured at 405 nm. β-Galactosidase activity in intact cells was visualized using 1 mg/ml X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) according to standard procedures. Growth of cells was monitored by direct counting of viable, trypan blue-excluding cells or by XTT assay performed as recommended (Roche Applied Science, Indianapolis, IN) with similar results. After incubation with the XTT labeling reagent, optical density of the samples was measured at 490 nm.

For analysis of RNA, total RNA was extracted from cells using the RNeasy method (QIAGEN, Chatsworth, CA). For reverse transcription (RT), 0.8 μg of total RNA and 50 ng of random primer were used in a 20-μl reaction volume with Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA), followed by PCR amplification with 1 μl of the synthesized cDNA and *Taq* polymerase (Fermentas, Hanover, MD). The PCR products were resolved by agarose gel electrophoresis and stained with ethidium bromide. The primers used to amplify the respective cDNAs were as follows: pThoc1, 5′CCCTTGACCAAGCTTTC AGA and 5′TTCCCAGCAGCATAGAAGGT; UAP56, 5′ATGGCAGAGAA CGATGTG and 5′TAACTGGCTCCAGCTGTT; Tho2, 5′ATGTTTGTTTCA GACACAGT and 5′GTTCCTGGTAAAACTTGAAT; ACC1, 5′TGGCCAGA TTCAAGCCATGT and 5′ACCAAGCTGCGGATTTGCTT; calreticulin, 5′A CATGCACGGAGACTCAGAAT and 5′TGTTGTCTGGCCGCACAAT; topoisomerase I, 5′AGGTTCCTTCTCCTCCTCCA and 5′ CCTTTTCATTGCCTG CTCTC; β-actin, 5′GTGGCATCCACGAAACTACC and 5′AGTACTTGCGC TCAGGAGGA.

RNA levels were quantitated by slot or Northern blotting. Total cellular RNA (1 μg to 10 μg) isolated as described above was blotted onto Hybond N^+ membrane (Amersham Biosciences Corp, Piscataway, NJ). The blots were hybridized to random-primed, ^{32}P-labeled hybridization probes for *lacZ* (600 bp spanning the middle of the coding sequence) or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in Rapid-hyb buffer (Amersham Biosciences Corp, Piscataway, NJ) at 65°C for 2 h. Blots were washed according to the manufacturer's recommendations, and the radioactive signals were quantitated and visualized using a PhosphorImager and ImageQuant software (Molecular Dynamics, Sunnyvale, CA).

For DNA-RNA hybrid analysis, siRNA-transfected HeLa cells were infected with the CMV-*lacZ* adenovirus at a multiplicity of infection of 50 and nucleic acid was subsequently extracted by the RNeasy method (QIAGEN, Chatsworth, CA). Four micrograms of total nucleic acid was treated with 3 U of RNase-free DNase I (Promega) for 2 h at 37°C, followed by phenol-chloroform purification and ethanol precipitation. Another 4-μg aliquot was treated with 5 U of RNase H (Fermentas) or RNase H plus RNase A for 2 h at 37°C in 40 mM Tris (pH 8.0)–8 mM $MgCl_2$–50 mM NaCl–3 mM dithiothreitol–2.5 mM spermidine. The samples were purified as described above and then digested with 3 U of RNase-free DNase I. The treated samples were analyzed by RT-PCR as described above or by slot blot analysis. Random prime labeled full-length *lacZ* or GAPDH cDNA was used as a hybridization probe.

Immunostaining, Western blotting, and immunoprecipitation. For immunostaining, 2 days following transfection cells were washed in phosphate-buffered saline (PBS) and fixed in 4% formaldehyde–10% acetic acid in PBS. Cells were permeabilized with 0.1% Triton X-100 and 0.1% sodium citrate in PBS, washed, and incubated for 1 h at 37°C with mouse anti-pThoc1 antibody diluted 1:1,000 (GeneTex, San Antonio, TX). After washing, cells were incubated for 1 h at 37°C with Texas red anti-mouse immunoglobulin G (IgG) antibody diluted 1:200. Stained cells were analyzed by confocal fluorescent microscopy.

For Western blotting and immunoprecipitation, cells were harvested and washed with ice-cold PBS. Cell pellets were extracted in buffer containing 50 mM Tris (pH 7.4), 250 mM NaCl, 5 mM EDTA, and 0.1% NP-40 and supplemented with a cocktail of protease and phosphatase inhibitors (Sigma, St. Louis, MO). Extracts were clarified by centrifugation, and protein concentration were measured by Bradford assay. For Western blotting, 20 μg of total protein for each sample was resolved by 7.5% sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. The proteins were transferred to polyvinylidene difluoride membrane, and the blot was blocked with 5% nonfat dry milk in 20 mM Tris (pH 7.5)–150 mM NaCl–0.1% Tween 20 (TBST) for 1 h at room temperature. The blot was incubated with primary antibody diluted in fresh TBST-5%milk, and the bound antibody was detected using horseradish peroxidase-conjugated secondary antibody and chemiluminescence (Amersham Biosciences Corp, Piscataway, NJ).

For immunoprecipitation, cell extracts (600 μg to 1 mg of total cell protein) were incubated overnight with rabbit polyclonal antibody directed against pThoc1 or GFP (Clontech, Palo Alto, CA). A GST fusion protein containing the carboxy-terminal half of pThoc1 was used to raise the rabbit anti-pThoc1 antibody. Immune complexes were collected by incubation with protein A/G plus agarose beads (Santa Cruz Biotechnology, Santa Cruz, CA) and eluted by boiling in 2× SDS Laemmli loading buffer. Proteins were resolved by 7.5% SDS-polyacrylamide gel electrophoresis and analyzed by Western blotting using 1:1,000-diluted mouse monoclonal GFP antibody (Clontech, Palo Alto, CA), 1:5,000-diluted mouse monoclonal pThoc1 antibody, or H5 monoclonal antibody (Covance, Berkeley, CA).

ChIP. Formaldehyde cross-linking and chromatin immunoprecipitation (ChIP) were carried out as recommended by the manufacturer (Upstate, Charlottesville, VA). Cells were fixed in Dulbecco modified Eagle medium with 1% formaldehyde. Sonicated, cross-linked chromatin was adjusted to 1× ChIP dilution buffer before immunoprecipitation. Affinity-purified rabbit pThoc1 antibody (see above) or RNA PolII antibody (N-20; Santa Cruz Biotechnology, Santa Cruz, CA) was used to immunoprecipitate protein-chromatin complexes. Preimmune rabbit IgG was used as a negative control. Immunoprecipitated samples were washed in wash buffer I~IV, followed by cross-link reversal and purification of the precipitated chromatin.

Immunoprecipitated DNA was detected by PCR using 2 μl of immunoprecipitated material, 0.5 μM each PCR primer, 2 mM $MgCl_2$, 0.2 mM deoxynucleoside triphosphates, 1× *Taq* buffer, and 1 U of *Taq* DNA polymerase (Fermentas, Hanover, MD). The primers used to detect the *lacZ* gene were as follows: P, 5′TCGGTACCCGGGTCGAGTA (227 bases 5′ of the start codon) and 5′TCT TTCGATCCCGGGGAT (28 bases 5′ of the start codon); M, 5′ATGGTCAG

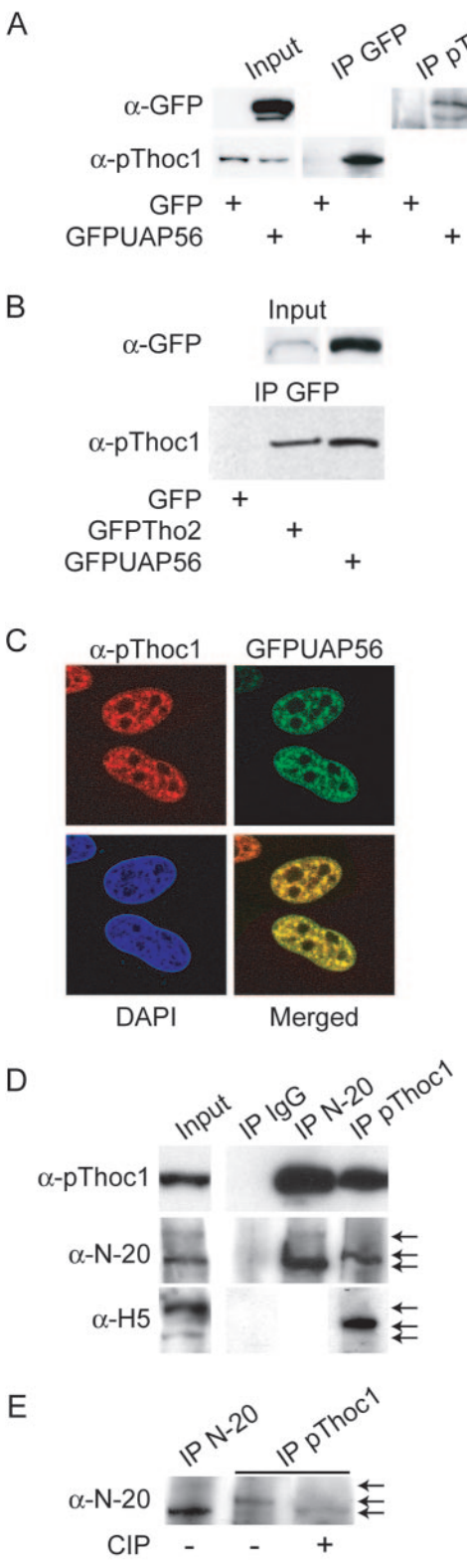

FIG. 1. pThoc1 stably associates with putative TREX proteins and elongating RNA PolII in intact cells. (A) HeLa cells were transfected with the indicated expression plasmids, and total cell extracts were prepared 2 days later. Aliquots of the extracts representing 5% of the amounts used in the immunoprecipitations were analyzed directly by Western blotting (Input) using monoclonal antibodies specific for GFP (α-GFP) or pThoc1 (α-pThoc1). Immunoprecipitates were prepared using polyclonal antibodies specific for GFP (IP GFP) or pThoc1 (IP pThoc1), and the immunoprecipitates were analyzed by Western blotting. Note the reciprocal coimmunoprecipitation of pThoc1 and GFPUAP56. (B) Cell extracts expressing GFP, GFPUAP56, or GFPTho2 were analyzed by immunoprecipitation and Western blotting as in panel A. Note that pThoc1 coimmunoprecipitates with both GFPTho2 and GFPUAP56. (C) HeLa cells transfected with GFPUAP56 expression plasmid were fixed and stained for pThoc1 (α-pThoc1) or DNA (4′,6′-diamidino-2-phenylindole [DAPI]). Staining and GFP signal were visualized by confocal fluorescence microscopy. The merged image combines the GFPUAP56 signal and the signal from pThoc1 staining. (D) HeLa total cell extracts were analyzed by immunoprecipitation and Western blotting as in panel A using antibodies specific for RNA PolII (α-N-20), serine 2 CTD-phosphorylated RNA PolII (α-H5), pThoc1 (α-pThoc1), or preimmune IgG. The arrows indicate three differentially migrating forms of RNA PolII resolved by Western blotting. The upper band corresponds to hyperphosphorylated RNA PolIIo, the lower band corresponds to hypophosphorylated RNA PolIIa, and the serine 2-phosphorylated band of intermediate mobility preferentially associates with pThoc1. (E) Total HeLa cell extracts were analyzed by immunoprecipitation and Western blotting as in panel D, except that aliquots of the pThoc1 immunoprecipitate were treated with calf intestinal phosphatase (CIP) as indicated before Western blot analysis.

GTCATGGATG (1,159 bases 3′ of the start codon) and 5′TCTCCAGGTAGC GAAAGC (1,640 bases 3′ of the start codon); 3′, 5′ATGGCGATTACCGT TGAT (2,658 bases 3′ of the start codon) and 5′TTTTTGACACCAGACCAA (3,140 bases 3′ of the start codon). The primers used to detect the human calreticulin gene were as follows: P, 5′GCGCTCCCGCGGTTTCTTTA (434 bases 5′ of the start codon) and 5′TGCCCATGGCCGTCTTTT (170 bases 5′ of the start codon); M, 5′TGAGTACAAGGTGAGTTTGG (1,963 bases 3′ of the start codon) and 5′CAGCACGCCAAAGTTATC (2,170 bases 3′ of the start codon); 3′, 5′TAGGAACACAGGTGGAAA (4,775 bases 3′ of the start codon) and 5′TTGTTTCTCTGCTGCCTG (5,031 bases 3′ of the start codon). Input DNA was diluted 200-fold, and 0.5 μl was analyzed by PCR as described above. Relative PCR signals were quantitated using ImageQuant software (Molecular Dynamics, Sunnyvale, CA).

RESULTS

pThoc1 stably interacts with elongating RNA PolII and putative human THO and TREX proteins in intact cells. Data from yeast suggest that the THO component Hpr1p is essential for recruiting the RNA export factor Sub2p to transcribed genes regulated by THO (34). Here we test whether human pThoc1 forms stable interactions with both UAP56 (yeast Sub2p) and elongating RNA PolII in intact cells. Since antibodies against UAP56 are not generally available, the protein is expressed as a GFP fusion. GFP immunoprecipitates prepared from GFP-UAP56-expressing cells, but not GFP-expressing cells, contain readily detectable pThoc1 (Fig. 1A). Likewise, pThoc1 immunoprecipitates prepared from GFP-UAP56-expressing cells contain detectable levels of GFP-UAP56, indicating that pThoc1 and GFP-UAP56 can reciprocally coimmunoprecipitate in intact cells. pThoc1 has previously been immunolocalized to the nucleus with significant concentration in nuclear speckles rich in RNA processing factors (9). We have tested whether endogenous pThoc1 colocalizes with GFP-UAP56. GFP-UAP56's subnuclear localization is very similar to the immunofluorescent staining pattern of pThoc1 (Fig. 1C). Indeed, upon merging the GFP-UAP56 and immunostained pThoc1 signals it is apparent that their subnuclear localization patterns overlap significantly, particularly within nuclear speckles. To assess whether pThoc1 binds putative human THO proteins, we have performed coimmunoprecipitation assays with GFP-tagged

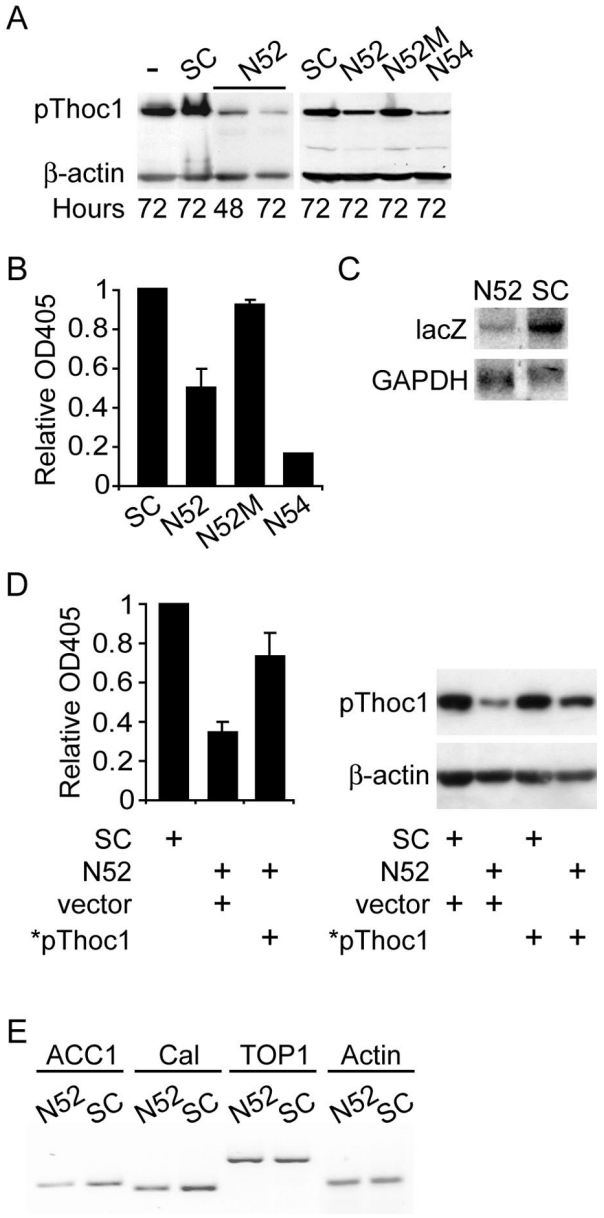

FIG. 2. Depletion of pThoc1 reduces RNA expression from a subset of genes, including bacterial *lacZ*. (A) siRNA duplexes targeted against two different regions of pThoc1 (N52 and N54), SC, N52 M, or no siRNA (−) were transfected into HeLa cells. At the indicated times after transfection, total cell extracts were prepared and analyzed by Western blotting for pThoc1 or β-actin as indicated. (B) HeLa cells treated with the indicated siRNA were infected with a recombinant adenovirus containing a CMV-*lacZ* gene. β-Galactosidase activity within total protein extracts from these cells was measured by ONPG assay. The optical density at 405 nm (OD405) was normalized to the value in the SC-treated control cells. The data represent the mean and standard deviation of three or more experiments. Note the correlation between pThoc1 levels and β-galactosidase activity. (C) Total RNA extracted from HeLa cells transfected with the indicated siRNAs was analyzed for *lacZ* or GAPDH mRNA levels by Northern blotting. Note the decrease in *lacZ* mRNA, but not GAPDH mRNA, upon depletion of pThoc1. (D) HeLa cells were transfected with the indicated siRNAs and expression plasmids and infected with the CMV-*lacZ* adenovirus, and total protein extracts were made as in panel B. Vector indicates the empty control vector. *pThoc1 is a plasmid designed to express pThoc1 mRNA containing silent mutations in the region targeted by the N52 siRNA. Extracts were tested for β-galactosidase activity by ONPG assay (left side). The data show the mean and standard deviation of two experiments. Expression of *pThoc1 in the presence of N52 was assayed by Western blot assay of HeLa cell extracts transfected with the indicated plasmids and siRNAs (right side). β-Actin was used as a loading control for the Western blot assay. (E) RNA expression of four endogenous human genes in HeLa cells transfected with the indicated siRNA was analyzed by RT-PCR. The genes analyzed include those for ACC1, calreticulin (Cal), topoisomerase I (TOP1), and β-actin (Actin). The results are representative of multiple experiments.

human Tho2 (homologue of yeast THO component Tho2p). pThoc1 protein is readily detected in GFP immunoprecipitates prepared from GFP-hTho2-expressing cells, but not from GFP-expressing cells (Fig. 1B).

If pThoc1 functions like yeast Hpr1p, then pThoc1 should stably associate with elongating forms of RNA PolII. We have tested this possibility using coimmunoprecipitation assays. Immunoprecipitates prepared with the N-20 RNA PolII antibody, but not preimmune IgG, contain readily detectable pThoc1 (Fig. 1D, top). Comparison of this signal with the signal from pThoc1 immunoprecipitates indicates that most of the soluble pThoc1 that can be captured by immunoprecipitation is stably associated with RNA PolII. Immunoprecipitates prepared with pThoc1 antibody also contain detectable levels of RNA PolII, demonstrating reciprocal coimmunoprecipitation of these endogenous proteins (Fig. 1D, middle). The N-20 antibody recognizes two prominent species of RNA PolII that differ in phosphorylation status within the carboxy-terminal domain (CTD) (3). The fastest-migrating species is hypophosphorylated. The slowest-migrating species is heavily phosphorylated, including phosphorylation on serine 2 of the CTD that is typically associated with elongating transcription complexes (16). The predominant RNA PolII species associated with pThoc1 migrates with a mobility intermediate between these two forms. To test whether pThoc1-associated RNA PolII is phosphorylated on serine 2, we have probed a similar blotting assay with the serine 2 phospho-specific antibody H5. The RNA PolII species associated with pThoc1 is strongly immunoreactive with the H5 antibody (Fig. 1D, bottom). Consistent with this finding, treatment of pThoc1 immunoprecipitates with phosphatase prior to Western blot analysis alters the mobility of pThoc1-associated RNA PolII (Fig. 1E).

Inhibition of RNA expression in pThoc1-depleted cells. The expression of a subset of yeast genes is dependent on Hpr1p (7). In particular, expression of the bacterial *lacZ* gene in yeast is highly dependent on Hpr1p, possibly due to its high G+C content (7). We have used the bacterial *lacZ* gene driven by the early promoter of the CMV and contained within a replication-defective recombinant adenovirus to measure the effects of pThoc1 depletion on gene expression in human cells. siRNA targeting pThoc1 (N52 or N54), but not SC or a point-mutated version of N52 (N52 M), effectively reduces pThoc1 levels but has no detectable effect on the levels of β-actin (Fig. 2A). Depletion of pThoc1 by N52 or N54 decreases β-galactosidase activity expressed from the *lacZ* adenovirus by two- to fivefold relative to cells treated with SC or N52 M (Fig. 2B). To ascertain whether the defect in β-galactosidase activity extends to RNA expression, *lacZ* RNA has been assayed by Northern blotting in pThoc1-depleted or control cells. The level of full-

length *lacZ* mRNA decreases upon pThoc1 depletion (Fig. 2C). GAPDH RNA levels are not affected by loss of pThoc1. Consistent with the loss in β-galactosidase activity, quantitation of *lacZ* RNA by slot blotting indicates that depletion of pThoc1 causes a twofold reduction in total *lacZ* RNA relative to control cells (see Fig. 5D; data not shown). To demonstrate the specificity of the observed effects for pThoc1 depletion, we have tested whether β-galactosidase activity can be rescued by expression of pThoc1 mRNA containing silent mutations in the region targeted by the N52 siRNA (*pThoc1) and thus resistant to siRNA-mediated gene silencing. Cotransfection of the *pThoc1 plasmid with N52 siRNA causes a significant increase in both β-galactosidase and pThoc1 levels relative to N52-treated cells cotransfected with the empty vector control (Fig. 2D).

Human genes regulated by the putative human THO or TREX complexes have not been identified. Preliminary gene expression profiling suggests that 7% of genes show reduced expression in pThoc1-depleted 293 cells (Y. Li, unpublished observation). Reduced RNA expression from the human ACC1 and calreticulin genes upon pThoc1 depletion has been verified by RT-PCR (Fig. 2E). In contrast, the levels of β-actin and topoisomerase I RNA are unaffected. Hence, the expression of some endogenous human genes is also affected by pThoc1 depletion. Interestingly, ACC1 expression in yeast is dependent on Hpr1p (27).

Depletion of pThoc1 causes a transcriptional elongation defect. Reduced *lacZ* RNA expression may be due to defects in the initiation of transcription. To address this possibility, we have tested whether the effects of pThoc1 depletion on β-galactosidase activity were dependent on the particular promoter used to drive expression. We have used three different promoters to measure the effects of pThoc1 depletion on *lacZ* gene expression. In addition to the CMV promoter, *lacZ* genes utilizing a GAL4VP16 fusion protein-dependent synthetic promoter (GV16) and a tetracycline-inducible synthetic promoter (Tet) have been tested. The CMV-*lacZ* and GV16-*lacZ* genes are expressed from within a replication-defective recombinant adenovirus. The Tet-*lacZ* gene is stably integrated into the HeLa cell genome. Depletion of pThoc1 by N52 treatment decreases β-galactosidase activity expressed from each of the three *lacZ* genes tested by two- to sixfold relative to control siRNA-treated cells (Fig. 3A). Hence, the effects of pThoc1 depletion on *lacZ* expression are promoter independent. The decrease in *lacZ* expression is specific since GFP expression driven by an identical CMV promoter is relatively unaffected by pThoc1 depletion, perhaps owing to its lower G+C content (GFP at 39% versus *lacZ* at 56%) (Fig. 3B).

Since the effects of pThoc1 depletion were not dependent on the nature of the promoter, pThoc1 might affect steps subsequent to initiation of transcription. If pThoc1 functions during transcription elongation, then it should associate with promoter-distal regions of target genes, perhaps through its association with elongating RNA PolII. We used ChIP with pThoc1-specific antibody to test this possibility. Like ChIP analysis of yeast Hpr1p (15), pThoc1 antibody immunoprecipitated both promoter-proximal and promoter-distal regions of a chromosomally integrated *lacZ* gene (Fig. 4B). Since the endogenous calreticulin gene was sensitive to pThoc1 depletion, we tested whether pThoc1 associates with this endogenous human gene.

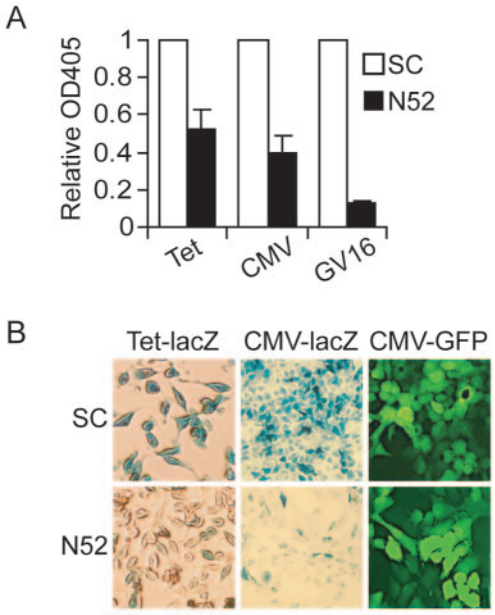

FIG. 3. Reduction in *lacZ* expression upon pThoc1 depletion is promoter independent. (A) HeLa cells containing a chromosomally integrated Tet-*lacZ* gene (Tet) or wild-type HeLa cells were transfected with the indicated siRNA. Wild-type transfected HeLa cells were subsequently infected with adenovirus containing the CMV-*lacZ* gene (CMV) or the GV16-*lacZ* gene (GV16). Total cell extracts were then assayed for β-galactosidase activity by ONPG assay. The data indicate the mean optical density at 405 nm (OD405), normalized to the SC-transfected control, for more than three experiments. The error bars represent the standard deviation from the mean. (B) Expression of the integrated Tet-*lacZ* or adenoviral CMV-*lacZ* gene was visualized under phase-contrast microscopy by 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) staining after treatment with control (SC) or pThoc1 (N52) siRNA. Expression of GFP from the adenoviral CMV-GFP gene was directly visualized under fluorescence microscopy. Note the decrease in *lacZ* expression, but not GFP expression, upon depletion of pThoc1.

PCR primer pairs spanning promoter-proximal and promoter-distal regions of the calreticulin gene were able to amplify detectable bands of appropriate size from immunoprecipitates prepared with the pThoc1 antibody, but not with preimmune IgG (Fig. 4C). To test whether pThoc1 was recruited specifically to actively transcribing genes, we compared the ChIP signal from a transcriptionally silent *lacZ* gene (doxycycline treated) and a transcriptionally active *lacZ* gene (vehicle treated). pThoc1 immunoprecipitates from vehicle-treated cells (transcriptionally active) consistently generated a greater *lacZ* ChIP signal than immunoprecipitates from doxycycline-treated cells (transcriptionally inactive) (Fig. 4D). The ability to detect pThoc1 associated with promoter-distal regions of actively transcribing target genes was consistent with a hypothesized role for pThoc1 in transcriptional elongation.

If pThoc1 depletion affects *lacZ* expression by decreasing the efficiency of transcriptional elongation, then loss of pThoc1 should decrease the density of RNA PolII associated with promoter-distal regions of the *lacZ* gene. We tested this by ChIP using the N20 RNA PolII antibody. Promoter-proximal and promoter-distal regions of an actively transcribing *lacZ* gene could be readily detected from N20 immunoprecipitates but not from preimmune IgG immunoprecipitates (Fig. 4E).

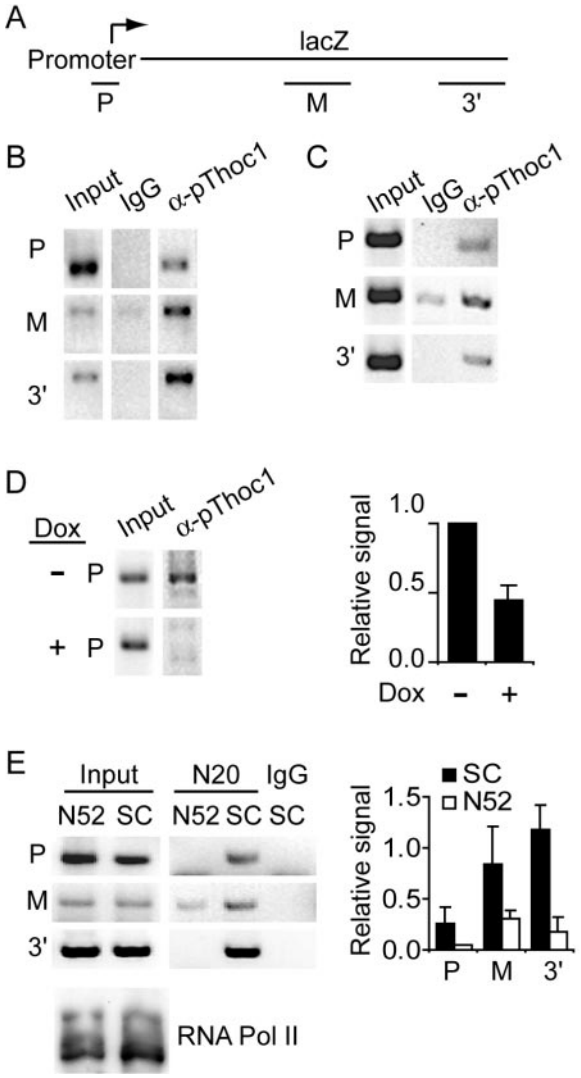

FIG. 4. pThoc1 associates with promoter-distal regions of genes and affects the density of transcribing RNA PolII along those genes. (A) Schematic representation of the chromosomally integrated Tet-*lacZ* gene indicating the regions of the gene (P, M, and 3′) analyzed by ChIP in relation to the promoter and coding sequence. (B) ChIP analysis of the integrated Tet-*lacZ* gene using antibody specific for pThoc1 (α-pThoc1) or a preimmune IgG control antibody (IgG). Input represents PCR amplification of the starting extract prior to immunoprecipitation. The labels for the gene regions assayed (P, M, and 3′) correspond to the positions indicated in panel A. The results are representative of multiple experiments. (C) The endogenous human calreticulin gene was analyzed by ChIP as in panel B using the indicated antibodies or input control. The regions of the calreticulin gene analyzed correspond to the promoter region (P), a region in the middle of the gene (M), and a region at the end of the gene (3′). The PCR primers used to amplify these regions are listed in Materials and Methods. (D) The integrated *lacZ* gene was analyzed by ChIP using the pThoc1 antibody and the P primer pair as in panel B, in either the absence (−) or the presence (+) of doxycycline (Dox). The right side shows the mean and standard deviation of the PCR signals for three experiments. The data were corrected by normalization to the input controls and are expressed relative to the untreated sample. (E) The Tet-*lacZ* gene was analyzed by ChIP as in panel B using an RNA PolII antibody (N20) or a nonimmune IgG control, in cells treated with either SC or N52 siRNA. The right side shows the mean and standard deviation of the PCR signals for each primer pair for three experiments. The data were normalized to the input control for each sample. At the bottom is an analysis of total RNA PolII levels in SC- or N52-treated cells by Western blotting using the N-20 antibody.

These ChIP signals were consistently reduced when immunoprecipitating RNA PolII from pThoc1-depleted cells, despite the fact that RNA PolII protein levels remain unchanged. Hence, the density of RNA PolII along the *lacZ* gene was reduced upon depletion of pThoc1. Together, the above data are consistent with a role for pThoc1 in transcriptional elongation. However, the data did not rule out the possibility of promoter-independent effects on earlier steps of transcription, including initiation.

Hpr1p-deficient yeast cells are more sensitive to transcriptional elongation inhibitors like 6-azauracil (5). To more directly test whether pThoc1 plays a role in transcriptional elongation, we have tested the sensitivity of pThoc1-depleted cells to the transcriptional elongation inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). Cells transfected with control siRNA show a dose-dependent decrease in cell viability after treatment with DRB (Fig. 5A). Cells depleted of pThoc1 are three- to fivefold more sensitive to DRB at a given dose. In contrast, both control and pThoc1-depleted cells are equally sensitive to the protein synthesis inhibitor cycloheximide. Interestingly, treatment with DRB specifically reduces accumulation of soluble pThoc1 protein. pThoc1 levels decrease significantly within 4 h of DRB treatment, despite the fact that a 24-h treatment with cycloheximide affects pThoc1 levels only slightly (Fig. 5B). The effect of DRB on pThoc1 levels is specific since p53 levels increase upon DRB treatment (data not shown), consistent with a recent report by O'Hagan and Ljungman (21). This observation is consistent with specific loss of pThoc1 stability or solubility upon DRB treatment. It is noteworthy that Hpr1p abundance declines in yeast upon genetic depletion of Cdc73p, a component of the Hpr1p-containing Paf1-RNA PolII complex that is distinct from Srbp-containing RNA PolII complexes (4).

Impairment of transcriptional elongation in Hpr1p-deficient yeast is mediated by cotranscriptionally formed, RNase H-sensitive DNA-RNA hybrids (13). To test whether the transcription defect caused by pThoc1 depletion is also associated with increased DNA-RNA hybrid formation, we extracted nucleic acid from pThoc1-depleted cells and assayed for RNase H-sensitive *lacZ* RNA by RT-PCR and slot blotting. Under the conditions used, RT-PCR of DNase I-treated nucleic acid (total RNA) with primers spanning the middle of the *lacZ* gene yield similar signals from pThoc1-depleted and control cells (Fig. 5C). Upon RNase H treatment, the RT-PCR signal for the pThoc1-depleted cells is reduced to background levels, while the signal for control cells remains detectable. This suggests that upon pThoc1 depletion a significant fraction of *lacZ* RNA spanning the analyzed region is sensitive to RNase H and thus at least partially in the form of a DNA-RNA hybrid. To quantitate the relative extent of DNA-RNA hybrid formation, we slot blotted nuclease-treated nucleic acid and probed using the full-length *lacZ* cDNA. In control cells, the RNase H-resistant RNA signal (RNase H then DNase I) is 48% of the total *lacZ* RNA signal (DNase I) as quantitated by a phosphorimager analysis (Fig. 4D and E). In pThoc1-depleted cells, however, the RNase H-resistant signal is only 27% of the total *lacZ* RNA signal. Hence, there is an increase in the proportion of RNase H-sensitive *lacZ* RNA upon pThoc1 depletion, consistent with increased formation of DNA-RNA hybrids. In wild-type yeast cells, the fraction of *lacZ* RNA detected in

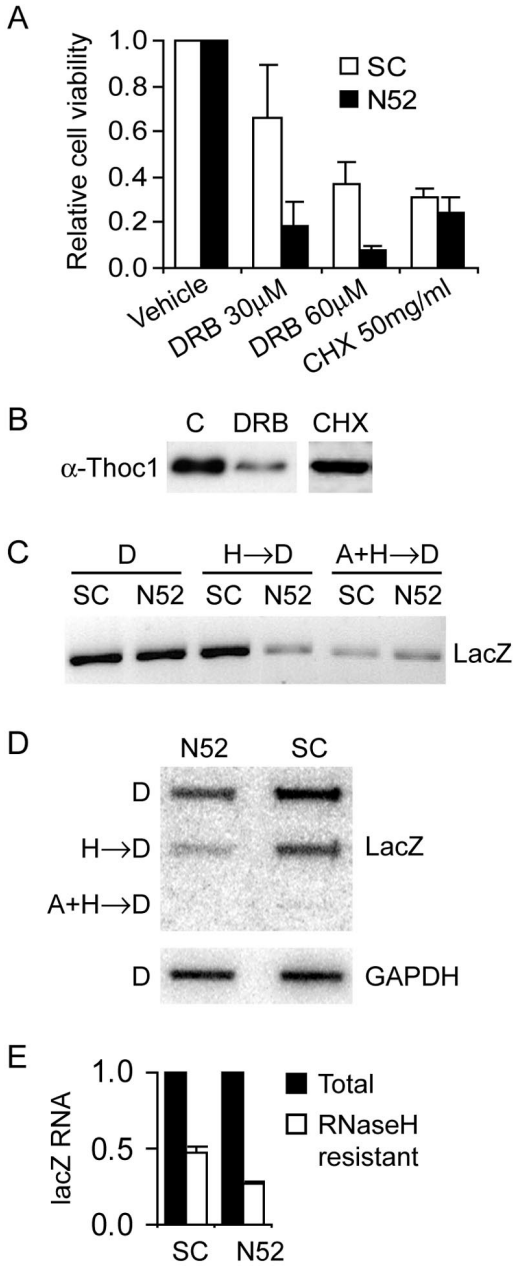

FIG. 5. pThoc1-depleted cells exhibit a transcriptional elongation defect similar to that observed in Hpr1p-deficient yeast. (A) HeLa cells treated with the indicated siRNA were exposed to cycloheximide (CHX), increasing concentrations of DRB, or a vehicle control for 24 h. The data show the mean percent viable cells after treatment, normalized to the vehicle control, for at least three experiments. Error bars represent the standard deviation from the mean. (B) HeLa cells treated with DRB for 4 h or cycloheximide (CHX) for 24 h were analyzed for pThoc1 by Western blotting using pThoc1 antibody (α-pThoc1). (C) HeLa cells treated with the indicated siRNA and infected with the CMV-*lacZ* adenovirus were extracted, and nucleic acid was isolated. Aliquots of the nucleic acid were treated with DNase I (D) to reveal total RNA, RNase H and then DNase I (HD) to specifically eliminate RNA within DNA-RNA hybrids, or RNase A plus RNase H and then DNase I (A+HD) as a background control. *lacZ* RNA levels in these samples were analyzed by RT-PCR using the M primer pair indicated in Fig. 4A. (D) Nucleic acid isolated and treated as in panel C was slot blotted and probed with the complete *lacZ* cDNA. Separate aliquots of DNase I-treated nucleic acid were blotted and analyzed for GAPDH RNA to serve as a loading control. The radioactive signal was visualized and quantitated using a phosphorimager. (E) Radioactive signals from panel D were quantitated using a phosphorimager and expressed relative to the total RNA signal (DNase treated) for each sample. The data are the mean and standard deviation of two independent experiments.

DNA-RNA hybrids is lower than that measured in control cells here (~5% versus ~50%) (13). This could be due to the different methods employed to detect DNA-RNA hybrids (RNase A resistance versus RNase H sensitivity), differences in the structure of the *lacZ* genes used, or biological differences between yeast and human cells.

Secondary phenotypes associated with pThoc1 depletion. Hpr1p-deficient yeast cells are associated with additional phenotypes, including a reduced growth rate (26). We have measured cell proliferation in several cell lines in the presence or absence of normal levels of pThoc1. The growth rate of pThoc1-depleted 293 cells is reduced for 3 days posttransfection compared to control cells (Fig. 6A). Between 3 and 4 days posttransfection, the growth rate of N52-transfected cells increase to near normal levels, coincident with recovery of normal pThoc1 levels (data not shown). Four additional cell lines have been transfected with N52 or control siRNA, and the effects on cell proliferation were measured. In three of the cell lines, N52 transfection is accompanied by a decrease in cell accumulation (Fig. 6B). One cell line, C33a, grows at a similar rate whether transfected with N52 or SC. However, N52 siRNA transfection does not efficiently deplete pThoc1 in this cell line (Fig. 6C). Hence, in all cell lines examined, there is a correlation between pThoc1 levels and cell growth rate.

Hpr1p-deficient yeast cells are more sensitive to DNA damage than wild-type strains (2), are synthetic lethal with mutations in the topoisomerase genes (25), and have defects in NER (12). We have explored the possibility that human cell lines depleted of pThoc1 may exhibit analogous phenotypes. About half as many cells pretreated with N52 siRNA survive 20 μM cisplatin-induced DNA damage or 1 μM camptothecin-induced topoisomerase inhibition as control cells (Fig. 7A). To test the effects of pThoc1 depletion on NER, we have used a host cell reactivation assay (17) with UV-irradiated, GFP-expressing adenovirus. UV-irradiated virus is used to infect pThoc1-depleted or control cells; GFP expression requires host cell-dependent nucleotide excision repair of UV-induced pyrimidine dimers. Consistent with results described above, pThoc1 depletion does not have a significant effect on the expression of GFP from unirradiated adenovirus. At higher doses of UV, however, GFP expression in pThoc1-depleted cells is undetectable 24 h postinfection as assayed by Western blotting and fluorescence microscopy (Fig. 7B and C). In contrast, control cells infected with the same irradiated adenovirus and analyzed at the same time postinfection express a detectable level of GFP. Hence, a greater level of GFP gene reactivation is observed in cells with normal levels of pThoc1. This is consistent with a defect in NER upon pThoc1 depletion. Given the sensitivity of pThoc1-depleted cells to other forms of DNA damage described above, it is likely that other DNA repair pathways are also affected by pThoc1.

Functional interaction between pThoc1 and other putative TREX components. Mutations of individual TREX genes yield similar phenotypes in yeast cells (6, 12, 30, 33). Further, the

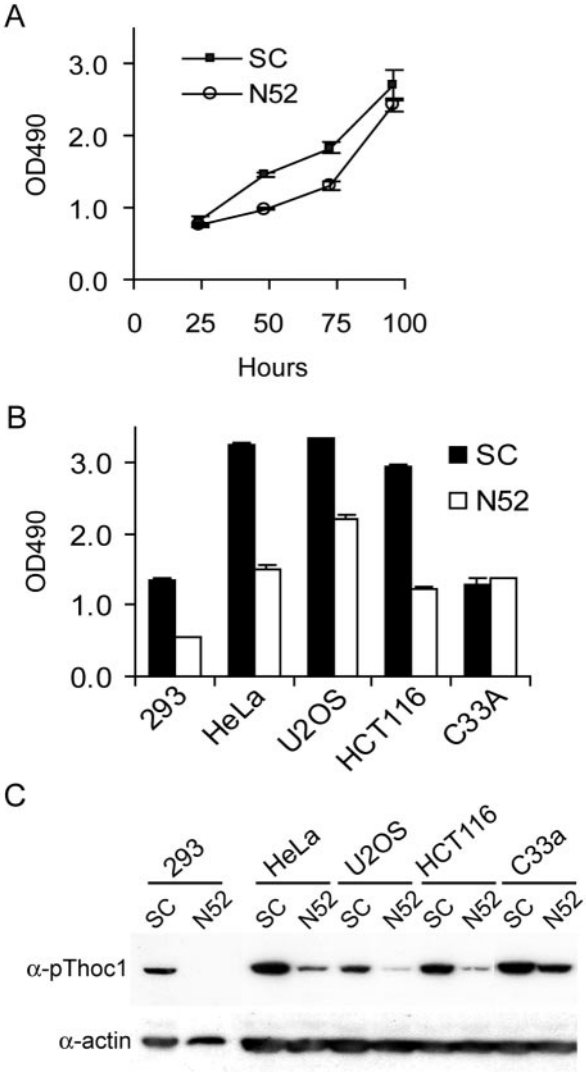

FIG. 6. Depletion of pThoc1 decreases growth rate in multiple cell lines. (A) 293 cells treated with pThoc1 (N52) or control (SC) siRNA were monitored for accumulation of viable cells at the indicated times posttransfection by XTT assay. The data points are the mean optical density at 490 nm (OD490) values for three experiments. Error bars represent the standard deviation from the mean. Similar results were obtained using a trypan blue dye exclusion assay (data not shown). (B) The indicated cell lines were transfected with siRNA as in panel A and assayed for viable cell accumulation 72 h posttransfection as in panel A. (C) Total cell extracts were prepared from cell lines transfected with the indicated siRNA as in panel B, and pThoc1 and β-actin levels were assayed by Western blotting using antibody specific for pThoc1 (α-pThoc1) or β-actin (α-actin).

stoichiometry of yeast TREX proteins is important since either reduced or excess protein can compromise TREX function (29). We have tested whether similar effects are observed in human cells. siRNAs targeted against UAP56 or Tho2 have been used to deplete their respective RNAs, and the effects on cell growth or bacterial *lacZ* expression have been measured. siRNA transfection is effective in reducing UAP56, Tho2, or pThoc1 RNA as assayed by RT-PCR (Fig. 8A). Depletion of UAP56, Tho2, or pThoc1 RNA reduces bacterial *lacZ* expression and the growth rate relative to control siRNA-transfected cells (Fig. 8B). Multiple independent siRNAs targeting different sequences of UAP56 or Tho2 give similar results. We have also tested the effects of pThoc1 overexpression on *lacZ* expression. Transfection of a small amount of pThoc1 expression plasmid actually increases β-galactosidase activity above control levels (Fig. 7C), suggesting that pThoc1 may be rate limiting for efficient *lacZ* expression. Transfection of increasing amounts of pThoc1 plasmid, however, decreases β-galactosi-

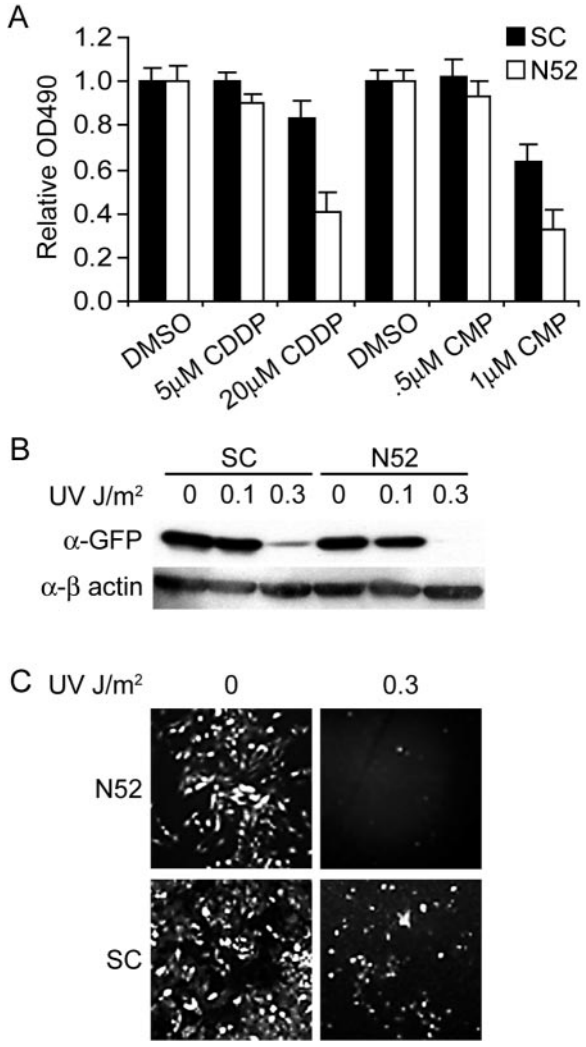

FIG. 7. pThoc1 depletion causes increased sensitivity to DNA damage and defects in nucleotide excision repair. (A) HeLa cells transfected with the indicated siRNA were treated with various concentrations of cisplatin (CDDP), camptothecin (CMP), or a vehicle control 72 h posttransfection. Accumulation of viable cells was measured 1 day later using XTT assay. The data are the mean optical density at 490 nm (OD490) of three experiments, normalized to the vehicle control. The error bars represent the standard deviation from the mean. Note the increased sensitivity of pThoc1-depleted cells to both agents. (B) HeLa cells treated with the indicated siRNA were infected with the CMV-GFP adenovirus previously irradiated with UV light (UV). Total cell extracts were prepared 1 day later, and the relative levels of GFP or β-actin were measured by Western blot analysis using the listed antibodies. Note the decreased reactivation of GFP expression in pThoc1-depleted cells at higher doses of UV. (C) Reactivation of GFP expression from UV-irradiated adenoviral DNA was visualized by fluorescent microscopy. siRNA treatment and irradiation were as in panel B. DMSO, dimethyl sulfoxide.

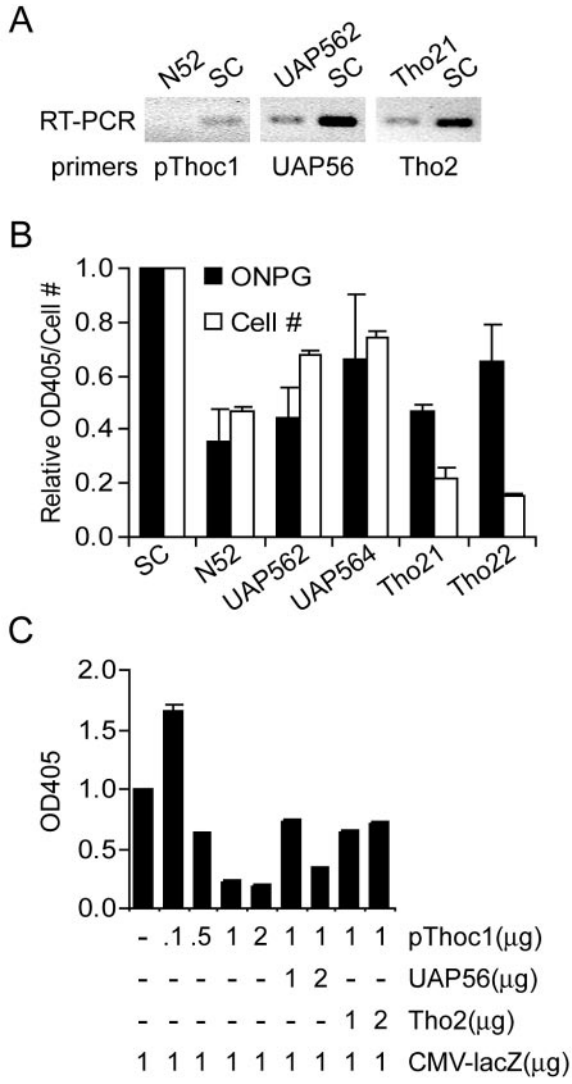

FIG. 8. pThoc1 functionally interacts with other putative members of the TREX complex. (A) siRNA directed against UAP56 (UAP562), Tho2 (Tho21), pThoc1 (N52), or the control (SC) was transfected into HeLa cells, and the effect on the respective RNA level was analyzed by RT-PCR. Note that each targeted siRNA reduced the level of the respective RNA relative to control siRNA transfection. (B) HeLa cells were treated with siRNA targeted against pThoc1 (N52), UAP56 (UAP562 and 4), Tho2 (Tho21 and 2), or control (SC), and the effect on *lacZ* expression (ONPG) or cell growth (Cell #) was determined by ONPG or XTT assay, respectively. The data represent the mean optical density at 405 nm (OD405) (ONPG) or OD490 (XTT) of duplicate experiments normalized to the SC control. The error bars indicate the standard deviation from the mean. The data are representative of multiple independent experiments. (C) Cells were transfected with the indicated amounts of expression plasmids engineered to express pThoc1 (pThoc1), UAP56 (UAP56), Tho2 (Tho2), or *lacZ* (CMV-lacZ). The effects of the exogenously expressed TREX proteins on *lacZ* expression were measured by ONPG assay. The data show the mean OD405 of three transfections. Error bars represent the standard deviation from the mean.

dase activity in a dose-dependent manner to about 20% of control levels ($P = 0.009$ by Student's t test). Coexpression of UAP56 or Tho2 with pThoc1 rescues pThoc1-mediated repression of *lacZ* expression ($P = 0.002$ and 0.04, respectively). Increased UAP56 or Tho2 presumably relieves repression of *lacZ* expression by binding excess pThoc1 and preventing titration of endogenous TREX components into substoichiometric complexes.

DISCUSSION

The primary conclusions of this work are that human pThoc1 regulates transcriptional elongation of a subset of genes and that it may function as the human orthologue of yeast Hpr1p in a protein complex analogous to yeast TREX. As in yeast, this human protein complex physically links elongating RNA PolII and RNA processing factors. This conclusion is significant for a number of reasons. First, it is currently unknown whether metazoan species with long, intron-containing genes use a mechanism analogous to TREX for coupling transcriptional elongation and RNA processing. Second, cellular phenotypes caused by disruption of this mechanism have not been reported in higher eukaryotic cells. Third, the observation that TREX regulates the expression of only a subset of genes, presumably by influencing cotranscriptional mRNP biogenesis, suggests the possibility that TREX assists in specifying posttranscriptional operons that may facilitate RNA processing of transcripts from coordinately regulated genes (14). Finally, the previously documented association between pThoc1 and Rb1 (8, 9) suggests the possibility that Rb1 may have a previously unappreciated role in regulating transcriptional elongation and RNA processing.

Several lines of evidence support the primary conclusions. First, pThoc1 forms stable interactions with UAP56, Tho2, and elongating forms of RNA PolII in intact cells. pThoc1 also colocalizes with UAP56 in subnuclear speckles. Our data confirm and extend findings from in vitro pulldown assays using UAP56 as bait (30). Since the role of UAP56 in RNA splicing and export has been previously documented (18), the stable interactions detected here are consistent with the hypothesis that these proteins form complexes in cells that physically link transcription and RNA processing. Additional biochemical characterization is required to discover whether these protein interactions reflect a single complex analogous to yeast TREX or multiple complexes.

It is noteworthy that pThoc1 interacts predominantly with a serine 2 phosphorylated form of RNA PolII that migrates in gels differently from the RNA PolIIo form typically associated with elongating transcription complexes. This pThoc1-associated RNA PolII makes up a relatively minor fraction of the total N-20- or H5-immunoreactive RNA PolII in these cells, consistent with the small percentage of genes affected by loss of pThoc1. These data suggest the possibility that pThoc1 interacts primarily with an elongating form of RNA PolII containing a set of posttranslational modifications distinct from those of the RNA PolIIo form.

A second line of evidence in support of the primary conclusion is that many of the same phenotypes observed in Hpr1p-deficient yeast are recapitulated in human cells depleted of pThoc1. Both Hpr1p-deficient yeast and pThoc1-depleted human cells show promoter-independent defects in the expression of an overlapping set of genes, including bacterial *lacZ* and endogenous ACC1. In both humans and yeast, the proteins associate with promoter-distal regions of actively tran-

scribing genes. Both Hpr1p-deficient yeast and pThoc1-depleted human cells are especially sensitive to chemical inhibitors of transcriptional elongation (6-azauracil and DRB, respectively). Finally, the molecular mechanism contributing to the observed transcriptional elongation defect, formation of DNA-RNA hybrids, appears to be similar in both yeast and human cells. Further support for functional analogy between human pThoc1 and yeast Hpr1p is provided by the overlapping set of cellular phenotypes associated with their depletion. In particular, both human and yeast cells show defects in cell growth, increased sensitivity to DNA damage, increased sensitivity to loss of topoisomerase activity, and defects in nucleotide excision repair. Finally depletion of pThoc1 or other putative human TREX proteins yields an overlapping set of phenotypes, as in yeast.

While the data presented here establish a role for pThoc1 in transcriptional elongation, we cannot exclude the possibility that pThoc1 also functions earlier in transcription. ChIP analysis indicates that pThoc1 associates with DNA spanning the TATAA box of the promoters tested. This suggests the possibility that pThoc1 may influence events during initiation or very early during elongation. Our data also indicate that pThoc1 and UAP56 colocalize within nuclear speckles. Phosphorylated forms of RNA PolII also colocalize to these splicing factor-rich nuclear speckles (20). Since active transcription occurs primarily at perichromatin fibrils outside of or at the periphery of nuclear speckles (19), pThoc1 and UAP56 may be preassembled and could associate with RNA PolII before recruitment to actively transcribed genes. Therefore, it is possible that pThoc1 and TREX influence events prior to transcription elongation.

pThoc1 lacks overall, statistically significant primary amino acid sequence similarity to any of the yeast TREX proteins in BLASTP alignments. Further, human pThoc1 and its metazoan orthologues contain protein interaction motifs, such as the death domain, that are apparently not present in fungi (11, 32). Hence, it is likely that differences in the structure and function of metazoan and yeast TREX complexes will be identified notwithstanding the observation that they have remarkably similar effects on transcription, cell growth, and response to DNA damage. To date, limited efforts to characterize metazoan TREX complexes have identified proteins conserved across higher eukaryotes, but not with *S. cerevisiae* proteins (23). Like pThoc1, these proteins may function analogously to yeast TREX proteins despite divergence in the primary amino acid sequence. TREX complex components may be among a class of proteins, like some spliceosomal proteins, whose amino acid sequence in *S. cerevisiae* has diverged well beyond expectation from functionally related proteins in *Schizosaccharomyces pombe* or other eukaryotic species (1). Alternatively, the subunit composition and function of TREX complexes may vary among species or biological contexts, possibly specifying distinct regulatory inputs. Nonetheless, the general mechanism of coupling transcriptional elongation to RNA processing factors through assembly of protein complexes containing both transcription elongation and RNA processing factors is shared from yeast to humans.

ACKNOWLEDGMENTS

The inducible Tet-*lacZ* HeLa cell line was graciously provided by Manfred Gossen (Max Delbruck Center for Molecular Medicine). We thank William Burhans, Margot Ip, Adrian Black, and members of the Goodrich lab for helpful discussions.

This work was supported by grants from the National Institutes of Health (CA-70292) and the Ralph C. Wilson Foundation to D.W.G.

REFERENCES

1. **Aravind, L., H. Watanabe, D. J. Lipman, and E. V. Koonin.** 2000. Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc. Natl. Acad. Sci. USA **97:**11319–11324.
2. **Bennett, C. B., L. K. Lewis, G. Karthikeyan, K. S. Lobachev, Y. H. Jin, J. F. Sterling, J. R. Snipe, and M. A. Resnick.** 2001. Genes required for ionizing radiation resistance in yeast. Nat. Genet. **29:**426–434.
3. **Castano, E., P. Gross, Z. Wang, R. G. Roeder, and T. Oelgeschlager.** 2000. The C-terminal domain-phosphorylated IIO form of RNA polymerase II is associated with the transcription repressor NC2 (Dr1/DRAP1) and is required for transcription activation in human nuclear extracts. Proc. Natl. Acad. Sci. USA **97:**7184–7189.
4. **Chang, M., D. French-Cornay, H. Y. Fan, H. Klein, C. L. Denis, and J. A. Jaehning.** 1999. A complex containing RNA polymerase II, Paf1p, Cdc73p, Hpr1p, and Ccr4p plays a role in protein kinase C signaling. Mol. Cell. Biol. **19:**1056–1067.
5. **Chavez, S., and A. Aguilera.** 1997. The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. Genes Dev. **11:**3459–3470.
6. **Chavez, S., T. Beilharz, A. G. Rondon, H. Erdjument-Bromage, P. Tempst, J. Q. Svejstrup, T. Lithgow, and A. Aguilera.** 2000. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. EMBO J. **19:**5824–5834.
7. **Chavez, S., M. Garcia-Rubio, F. Prado, and A. Aguilera.** 2001. Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **21:**7054–7064.
8. **Doostzadeh-Cizeron, J., R. Evans, S. Yin, and D. W. Goodrich.** 1999. Apoptosis induced by the nuclear death domain protein p84N5 is inhibited by association with Rb protein. Mol. Biol. Cell **10:**3251–3261.
9. **Durfee, T., M. A. Mancini, D. Jones, S. J. Elledge, and W.-H. Lee.** 1994. The amino-terminal region of the retinoblastoma gene product binds a novel nuclear matrix protein that co-localizes to centers for RNA processing. J. Cell Biol. **127:**609–622.
10. **Fan, H. Y., K. K. Cheng, and H. L. Klein.** 1996. Mutations in the RNA polymerase II transcription machinery suppress the hyperrecombination mutant hpr1 delta of Saccharomyces cerevisiae. Genetics **142:**749–759.
11. **Feinstein, E., A. Kimchi, D. Wallach, M. Boldin, and E. Varfolomeev.** 1995. The death domain: a module shared by proteins with diverse cellular functions. Trends Biochem. Sci. **20:**342–344.
12. **Gonzalez-Barrera, S., F. Prado, R. Verhage, J. Brouwer, and A. Aguilera.** 2002. Defective nucleotide excision repair in yeast hpr1 and tho2 mutants. Nucleic Acids Res. **30:**2193–2201.
13. **Huertas, P., and A. Aguilera.** 2003. Cotranscriptionally formed DNA-RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell **12:**711–721.
14. **Keene, J. D., and S. A. Tenenbaum.** 2002. Eukaryotic mRNPs may represent posttranscriptional operons. Mol. Cell **9:**1161–1167.
15. **Kim, M., S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski.** 2004. Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. EMBO J. **23:**354–364.
16. **Komarnitsky, P., E. J. Cho, and S. Buratowski.** 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. **14:**2452–2460.
17. **Lee, D. F., R. Drouin, P. Pitsikas, and A. J. Rainbow.** 2004. Detection of an involvement of the human mismatch repair genes hMLH1 and hMSH2 in nucleotide excision repair is dependent on UVC fluence to cells. Cancer Res. **64:**3865–3870.
18. **Luo, M. L., Z. Zhou, K. Magni, C. Christoforides, J. Rappsilber, M. Mann, and R. Reed.** 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature **413:**644–647.
19. **Misteli, T., and D. L. Spector.** 1998. The cellular organization of gene expression. Curr. Opin. Cell Biol. **10:**323–331.
20. **Mortillaro, M. J., B. J. Blencowe, X. Wei, H. Nakayasu, L. Du, S. L. Warren, P. A. Sharp, and R. Berezney.** 1996. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. Proc. Natl. Acad. Sci. USA **93:**8253–8257.
21. **O'Hagan, H. M., and M. Ljungman.** 2004. Nuclear accumulation of p53 following inhibition of transcription is not due to diminished levels of MDM2. Oncogene **23:**5505–5512.
22. **Reed, R.** 2003. Coupling transcription, splicing and mRNA export. Curr. Opin. Cell Biol. **15:**326–331.
23. **Rehwinkel, J., A. Herold, K. Gari, T. Kocher, M. Rode, F. L. Ciccarelli, M. Wilm, and E. Izaurralde.** 2004. Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. Nat. Struct. Mol. Biol. **11:**558–566.

24. **Rondon, A. G., S. Jimeno, M. Garcia-Rubio, and A. Aguilera.** 2003. Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. J. Biol. Chem. **278:**39037–39043.
25. **Sadoff, B. U., S. Heath-Pagliuso, I. B. Castano, Y. Zhu, F. S. Kieff, and M. F. Christman.** 1995. Isolation of mutants of Saccharomyces cerevisiae requiring DNA topoisomerase I. Genetics **141:**465–479.
26. **Santos-Rosa, H., and A. Aguilera.** 1994. Increase in incidence of chromosome instability and non-conservative recombination between repeats in Saccharomyces cerevisiae hpr1 delta strains. Mol. Gen. Genet. **245:**224–236.
27. **Schneiter, R., C. E. Guerra, M. Lampl, G. Gogg, S. D. Kohlwein, and H. L. Klein.** 1999. The *Saccharomyces cerevisiae* hyperrecombination mutant *hpr1*Δ is synthetically lethal with two conditional alleles of the acetyl coenzyme A carboxylase gene and causes a defect in nuclear export of polyadenylated RNA. Mol. Cell. Biol. **19:**3415–3422.
28. **Strasser, K., and E. Hurt.** 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. EMBO J. **19:**410–420.
29. **Strasser, K., and E. Hurt.** 2001. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature **413:**648–652.
30. **Strasser, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A. G. Rondon, A. Aguilera, K. Struhl, R. Reed, and E. Hurt.** 2002. TREX is a conserved complex coupling transcription with messenger RNA export. Nature **417:**304–308.
31. **Stutz, F., A. Bachi, T. Doerks, I. C. Braun, B. Seraphin, M. Wilm, P. Bork, and E. Izaurralde.** 2000. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA **6:**638–650.
32. **Weber, C. H., and C. Vincenz.** 2001. The death domain superfamily: a tale of two interfaces? Trends Biochem. Sci. **26:**475–481.
33. **West, R. W., Jr., B. Kruger, S. Thomas, J. Ma, and E. Milgrom.** 2000. RLR1 (THO2), required for expressing lacZ fusions in yeast, is conserved from yeast to humans and is a suppressor of SIN4. Gene **243:**195–205.
34. **Zenklusen, D., P. Vinciguerra, J. C. Wyss, and F. Stutz.** 2002. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. Mol. Cell. Biol. **22:**8241–8253.