MOLECULAR AND CELLULAR BIOLOGY, Sept. 2005, p. 7665–7674 Vol. 25, No. 17
0270-7306/05/$08.00+0 doi:10.1128/MCB.25.17.7665–7674.2005


# Role of the Mammalian RNA Polymerase II C-Terminal Domain (CTD) Nonconsensus Repeats in CTD Stability and Cell Proliferation†

Rob D. Chapman,* Marcus Conrad, and Dirk Eick

*GSF-Research Centre for Environment and Health, Institute for Clinical Molecular Biology and Tumour Genetics, Marchioninistr. 25, D-81377 Munich, Germany*

Received 18 April 2005/Returned for modification 6 June 2005/Accepted 14 June 2005

**The C-terminal domain (CTD) of mammalian RNA polymerase II (Pol II) consists of 52 repeats of the consensus heptapeptide YSPTSPS and links transcription to the processing of pre-mRNA. The length of the CTD and the number of repeats diverging from the consensus sequence have increased through evolution, but their functional importance remains unknown. Here, we show that the deletion of repeats 1 to 3 or 52 leads to cleavage and degradation of the CTD from Pol II in vivo. Including these repeats, however, allowed the construction of stable, synthetic CTDs. To our surprise, polymerases consisting of just consensus repeats could support normal growth and viability of cells. We conclude that all other nonconsensus CTD repeats are dispensable for the transcription and pre-mRNA processing of genes essential for proliferation.**

The unique C-terminal domain (CTD) of the RNA polymerase II large subunit (Pol II LS/Rpb1) is highly conserved through evolution, apparently increasing in length and diversifying in structure with the complexity of organism (45). Of the 52 mammalian CTD repeats, 21 obey the evolutionary conserved, consensus heptapeptide $Y^1S^2P^3T^4S^5P^6S^7$; the remaining 31, nonconsensus repeats display a variety of substitutions at positions 2, 4, 5, and 7 (see Fig. 1) (11). The CTD is essential for transcription initiation and elongation in the context of chromatin (for a review, see reference 43). It is striking, in that it is almost entirely composed of residues that can be modified either through phosphorylation ($Y^1$, $S^2$, $T^4$, $S^5$, and $S^7$) (4, 49), isomerization ($P^3$ and $P^6$) (33), or glycosylation (S and T) (23; for an overview, see reference 7). These modifications may influence the ability of factors to bind CTD, as well as their activities. Such factors include components of the capping machinery (10, 21, 35), serine/arginine-rich (SR) proteins (6, 34, 38, 46, 47), splicing factors (32), elongation factors (19, 36), including cyclin T1 (48), and factors involved in 3′RNA processing and transcription termination (15, 24, 28, 44). The CTD-interacting domains (CTD-ID) of several factors have been identified (3, 12). These also include the proline-rich sequence-binding WW and FF domains (8, 32). The CTD tandem repeat structure could serve to bind multiple factors simultaneously, with the nonconsensus repeats providing further specificity for factor binding. The CTD may be phosphorylated on more than 50 sites in vivo (39); the specific phosphorylation status of serine-2 and -5 correlates with the binding and displacement of CTD-binding factors at different stages during the transcription cycle (26); a clear function for the phosphorylation of tyrosine-1, threonine-4, and serine-7 has yet to be shown (5, 9).

Why the CTD sequence is 100% conserved across mammals is a mystery: deletion to 31 repeats interferes with cell viability (30); however, mice homozygous for a deletion to 39 repeats, although smaller than wild-type littermates, are viable (29). Several extensive studies in mammalian cells have attempted to deconstruct the CTD, in order to understand the importance of the consensus and nonconsensus repeats in vivo (2, 16). The nonconsensus 52nd CTD repeat (CTD52) contains an unusual 10-amino-acid extension, which includes a constitutively phosphorylated CKII site (9) and is the binding site for the Abl1 tyrosine kinase (5). Deletion of CTD52 results in the increased degradation of such mutants to the CTD-less Pol IIb form and a reduced ability to transcribe (9), splice, and 3′-end process pre-mRNA (17). The consequences of removal of the CTD in vivo are severe: a CTD-less polymerase cannot initiate on the endogenous chromatin template; efficient coupling of RNA processing to the transcription complex no longer occurs (30, 31).

We now show that CTD52 is not the only element within the CTD that can affect its stability. Through a sophisticated genetic analysis, we demonstrate that a functional relationship between repeats 1 to 3 and CTD52 serves to regulate the stability of the Pol II by preventing its degradation to the IIb form. All other repeats could be deleted without inducing degradation. This important finding explains inconsistencies in earlier data but more usefully paves the way for a rational approach to analyze CTD function. We have now produced stable mutants composed entirely of perfect, "consensus" YSPTSPS repeats and found that a mutant containing 55 of such repeats was able to functionally replace the endogenous polymerase, in respect to growth and viability. These data suggest that, in the context of cultured cell lines, the nonconsensus CTD repeats may have more to do with efficiency than the regulation of processes essential for proliferation.

* Corresponding author. Mailing address: GSF-Institut fuer Klinische Molekularbiologie und Tumorgenetik, Hematologikum, Marchioninistr. 25, D-81377 Munich, Germany. Phone: 49-89-7099-529. Fax: 49-89-7099-500. E-mail: rob.chapman@gsf.de.

† Supplemental material for this article may be found at http://mcb.asm.org/.

## MATERIALS AND METHODS

**Cell lines and cell culture.** Suspension cells were grown in RPMI 1640 medium supplemented with 10% fetal calf serum, 100 U of penicillin/ml, 100 μg of streptomycin/ml, and 2 mM L-glutamine (Invitrogen), which is referred to below as growth medium. Stable cell lines were grown in the presence of 1 mg of G418/ml and 0.1 μg of tetracycline/ml as previously described (9, 30). Raji is an Epstein-Barr virus-positive Burkitt's lymphoma cell line. Cells were transfected with Pol II LS* expression vectors by electroporation (10 μg of plasmid DNA/$10^7$ cells; 960 μF, 250 V), and polyclonal cell lines were established after selection with G418 and tetracycline. For the expression of the various recombinant large subunits of Pol II, $10^7$ cells were washed three times with 20 ml of phosphate-buffered saline supplemented with 1% fetal calf serum and subsequently resuspended in 10 ml of growth medium. Monolayers of HeLa cells were propagated on tissue culture dishes in Dulbecco modified Eagle medium (Invitrogen) supplemented with 10% fetal calf serum (PAA). At 24 h after transfection or induction, α-amanitin (2 μg/ml, final concentration; Roche) was added to the medium to inhibit the endogenous Pol II. Protein samples were prepared after 48 h of α-amanitin treatment. After electroporation and during viability assays, medium was supplemented with 50 μM α-thioglycerol to compensate for low cell density.

**Preparation and analysis of proteins.** For Western blot analysis, whole-cell extracts were prepared by using Laemmli buffer (30,000 cells/0.1 ml) supplemented with a protease-inhibitor cocktail plus EDTA, as directed by the manufacturer (Roche). Separation was performed by using 6% sodium dodecyl sulfate-polyacrylamide gels before transfer to a polyvinylidene difluoride membrane (Immobilon-P or -FL; Millipore).

**Antibodies.** The following antibodies were used during the present study: high-affinity rat anti-HA monoclonal antibody 3F10 (Roche), horseradish peroxidase-linked goat anti-rat immunoglobulin G conjugate (Sigma), and horseradish peroxidase-, Alexa-Fluor 680-linked goat anti-rat immunoglobulin G conjugate (Molecular Probes). Blots were visualized by using either enhanced chemiluminescence (GE Healthcare) or the Odyssey IR scanner (Licor).

**Plasmids.** For the production of CTD mutants, a vector containing exon 29 of Pol II LS (pUC19-CTD) was modified. To produce C-terminal fusions with a truncated CTD52 (CTD52′), pUC19-CTD was digested with StyI and AgeI and then treated with calf intestinal phosphatase before insertion of a linker composed of the annealed phosphorylated primers A (5′-CAAGGGCTACAGCCTCACCAGCCCAGCCATCAGCCCAACCGGTTAGC-3′) and B (5′-GGCCGCTAACCGGTTGGGCTGATGGCTGGGCTGGTGAGGCTGTAGCC-3′). Fusions of the distal CTD region with small acidic or basic peptides were produced by insertion of phosphorylated linkers incorporating the encoding sequences between the AgeI and NotI sites. hTRX1 was cloned into the AgeI site after PCR of HeLa cDNA using the primers C (5′-AAAGACCGGTGTGAAGCAGATCGAGAGCAAGACTGC-3′) and D (5′-TTAAGCCGGCTTAGACTAATTCATTAATGGTGGC-3′) and digestion with AgeI and NgoMIV. Similarly, a fusion with enhanced green fluorescent protein (EGFP) was produced by digesting PCR product with AgeI and NotI (primer E, 5′-AAACGACCGGTTTTGTGAGCAAGGGCGAGGAGCTGTTCA-3′, and primer F, 5′-AAAAGGAAAGCGGCCGCGTCACTTGTACAGCTCGTCCATGCCGAG-3′).

Unique cut sites were introduced into pUC19-CTD*wt* and pUC19-Δ50 by mutagenesis PCR (see Fig. S1 in the supplemental material): BspEI in intron 28 (primer G, 5′-CATGTTTGTATGTCCGGAGCTTCACCAAC-3′, and primer H, 5′-GTTGGTGAAGCTCCGGACATACAAACATG-3′); AgeI in intron 28 (primer I, 5′-CATGTTTGTATGACCGGTGCTTCACCAAC-3′, and primer J, 5′-GTTGGTGAAGCACCGGTCATACAAACATG-3′); AvrII in exon 29, repeat 2 (primer K, 5′-CCTATAGGTGGTGCTATGTCTCCTAGGTACTCACCGACATCACCAGCC-3′, and primer L, 5′-GGCTGGTGATGTCGGTGAGTACCTAGGAGACATAGCACCACCTATAGG-3′); NheI in exon 29, repeat 3 (primer M, 5′-GCTATACACCCCAGAGCCCCGCTAGCTACTCCCCTACTTCACC-3′, and N, 5′-GGTGAAGTAGGGGAGTAGCTAGCGGGGCTCTGGGGTGTATAGC-3′). In brief, 0.1 μg of plasmid was amplified over 18 cycles by using *Pfu* (Promega) and the relevant primer pair. The PCR was DpnI digested, heat inactivated, and transformed directly into bacteria. The cloning strategy used enzyme isoschizomers, whose cut sites were destroyed when ligated together. The choice of enzymes used meant that the resulting sequence modification produced no change in the CTD amino acid sequence. The "seamless" deletion of repeats 1 to 3 was thereby achieved through digestion of the modified pUC19-CTD vectors with AvrII and NheI, followed by religation: no new residues were introduced by using this strategy. Phosphorylated linkers encoding repeat 52 flanked by two consensus repeats (primer O, 5′-CTAGCTACTCACCCACTAGCCCCAGCTACAGCCTCACCAGCCCAGCCATCAGCCCAGATGACAGCGATGAGGAGAACTCACCCACTAGCC-3′, and primer P, 5′-CTAGGGGCTAGTGGGTGAGTAGTTCTCCTCATCGCTGTCATCTGGGCTGATGGCTGGGCTGGTGAGGCTGTAGCTGGGGCTAGTGGGTGAGTAG-3′), or the Abl1CTD-ID (primer Q, 5′-CTAGCTACAGCATCCAGCAAATGAGGAACAAGTTTGCCTTCCGAGAGGCCATCAACAAGG-3′, and primer R, 5′-CTAGCCTTGTTGATGGCCTCTCGGAAGGCAAACTTGTTCCTCATTTGCTGGATGCTGTAG-3′) in sense or antisense. Mutants were either transferred to a tetracycline-regulated vector, as previously described (9), or by removal of mutant CTD with BspEI and NotI, and insertion via the new AgeI site in intron 28 and NotI site in exon 29. In the same way, mutants from J. Corden were recloned from supplied vectors by using primers G and S (5′-ATAAGAATGCGGCCGCACAGGCCACCGAGGTTGTCTGAC-3′), followed by digestion with BspEI and NotI. Synthetic CTDs were produced by using a linker composed of primers T (5′-CTAGCTACAGCCCTACCTCTCCTAGCTACTCGCCCACCTCCCCTAGCTACTCGCCAACCTCTCCTTCCTATTCTCCAACATCACCTAGGGATGAGGAGAACTGAC-3′) and U (5′-CTTGGTCAGTTCTCCTCATCCCTAGGTGATGTTGGAGAATAGGAAGGAGAGGTTGGCGAGTAGCTAGGGGAGGTGGGCGAGTAGCTAGGAGAGGTAGGGCTGTAG-3′) inserted between NheI and StyI. Consecutive cloning of a four-repeat, NheI-ClaI fragment from the resulting vector, back into itself, via AvrII-ClaI increased the number of repeats. A mutant with 35 consensus repeats (Δ3-52 Sub 35C) resulted from unexpected recombination in bacteria. Fusion of this mutant with 1 of 20 repeats (Δ3-52 Sub 20C) yielded the mutant Δ4-52 Sub 55C: CTD52 was added to synthetic mutants in the form of a linker (primer V, 5′-CTAGCTACAGCCTCACCAGCCCAGCCATCAGCCCAGATGACAGCGATGAGGAGAACTGAGC-3′, and primer W, 5′-GGCCGCTCAGTTCTCCTCATCGCTGTCATCTGGGCTGATGGCTGGGCTGGTGAGGCTGTAG-3′). For expression in HeLa cells, mutants were removed from their pUC19-CTD subclone as NgoMIV and ClaI fragments and cloned between the NgoMIV and ClaI sites of pSTC-TK HA Pol II LS* (18), a constitutive expression vector. All vectors were sequenced before usage (SEQUISERVE.DE).

**RNA analysis.** HeLa cells were harvested by using TriFast RNA extraction reagent (Peqlab). Then, 10 μg of total RNA was separated on a 1% agarose formaldehyde gel and transferred to nitrocellulose membrane (Hybond N+; GE Healthcare). Membranes were hybridized with $^{32}$P-labeled probes (HighPrime; Roche). The accession numbers for probes used are as follows: PES, BC032489; YY1, M77698; WDR12, NM_018256; NST, NM_014366; and c-Myc, D10493.

## RESULTS

**A system for the conditional expression and analysis of Pol II LS mutants.** We have established an episomal vector system in which a genomic clone of the mouse RNA Pol II large subunit (Pol II LS/Rbp1) is under the control of a tetracycline-regulated promoter (9, 30). This clone has been modified to include a hemagglutinin (HA) tag at its N terminus and a point mutation conferring resistance to the Pol II inhibitor α-amanitin (LS*) (Fig. 1) (18). Unique restriction enzyme sites were introduced at either side of exon 29 (encoding CTD) to allow the wild-type (*wt*) sequence to be exchanged for mutated CTD clones. Deletions, insertions, and mutations were performed in a manner not to introduce foreign residues into the CTD sequence other than those desired. After stable selection, the expression of mutants was induced by removal of tetracycline from the cell medium. The ability of Pol II LS* carrying mutant CTDs to replace the function of the endogenous Pol II LS can be assessed by culturing cells in medium containing α-amanitin, to which the endogenous Pol II LS is sensitive. The dependency of cells on our α-amanitin-resistant polymerases was controlled by readdition of tetracycline, as previously demonstrated (9, 30). Cultures die over the days after shutoff of the mutants' expression. It is unlikely that survival is due to a recombination event, since our mutants are episomally maintained and thus are present in many copies (30).

The number and complexity of mutants described in the present study necessitate schematics for each figure: some mu-

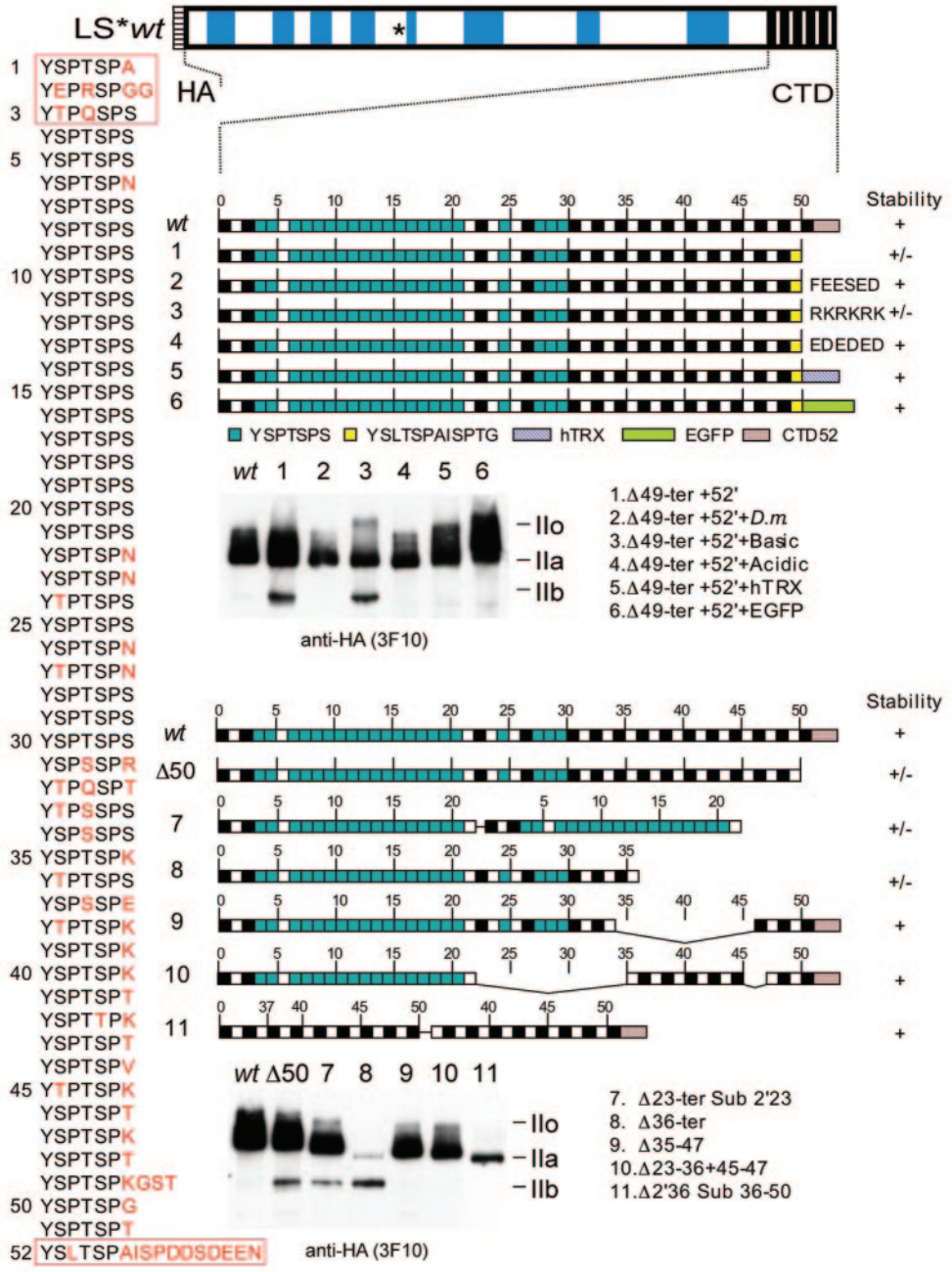


FIG. 1. The mammalian Pol II CTD is composed of 52 repeats of the consensus heptapeptide YSPTSPS (left-hand panel). Deviations from the consensus sequence, largely in the distal half of the CTD, are marked red. Repeats 1 to 3 and 52 (CTD52), the focus of the present study, are boxed in red. A genomic clone of Pol II LS*wt was manipulated to introduce a variety of deletions and substitutions; an N-terminal HA tag allows identification after Western transfer (upper panel). The schematics highlight the essential features of each mutant: repeats conforming to the consensus YSPTSPS are turquoise, nonconsensus repeats are black or white, and CTD52 is red. PAGE reveals the phosphorylated (IIo), nonphosphorylated (IIa), and CTD-less (IIb) forms of Pol II LS*. Mutant 1 was produced by replacing repeats 51 and 52 with a sequence encoding part of repeat 52 (YSLTSPAISPTG; yellow). Mutants 2 to 6 were produced by cloning the given acidic and basic polypeptides or domains into 52′ of mutant 1. Mutants 7 to 11 display a variety of deletions and substitutions, as published previously. Stable Raji cell lines were produced for each mutant; after expression, cells were cultivated 48 h in the presence of α-amanitin before cells were harvested, and the different forms produced by each mutant separated by polyacrylamide gel electrophoresis, followed by Western analysis. The effect of these mutations on degradation to the Pol IIb form is visible in each blot and summarized as "Stability" in the right-hand column. A *wt* clone, and a deletion mutant (Δ50) known to degrade to IIb were included as controls.

tants appear several times in the different figures for the purpose of comparison. As a control, a stable polymerase (*wt*), and a polymerase susceptible to degradation (Δ50) were included alongside the various, numerically labeled mutants. The number assigned to each mutant may change in subsequent figures; therefore, a key to the mutants is provided for each figure, in compliance with the original nomenclature used by J. Corden and coworkers (2).

**A six-residue motif at the C terminus of the CTD prevents degradation to Pol IIb.** It was previously found that mutants lacking the 17-amino-acid, 52nd repeat (CTD52) became degraded to the 180-kDa Pol IIb form in vivo (9). To further investigate this phenomenon, a mutant was produced containing 49 repeats plus part of CTD52 (Fig. 1, mutant 1), to which the sequences for a variety of polypeptides were cloned (Fig. 1, mutants 2 to 6). Three mutants were produced ending in either the acidic C-terminal sequence from the *Drosophila melanogaster* CTD (mutant 2) or a repetitive basic- or acid-amino acid hexapeptide (mutants 3 or 4, respectively). Expression analysis of these mutants (Fig. 1, upper panel) reveals both the nonphosphorylated (IIa) and slower-migrating, phosphorylated (IIo) forms. The CTD-less Pol IIb form appears only with mutants 1 and 3, where either the C-terminal seven amino acids of CTD52 are missing, or replaced by basic residues. These data suggest that it is the acidic nature of the C-terminal amino acids and not their specific sequence that is important in preventing degradation. In addition, fusions with larger domains, such as human thioredoxin 1 (hTRX) and EGFP were produced (mutants 5 and 6, respectively) and also provide protection against degradation. The degree to which a mutant is degraded is noted as "stability" in each figure and is scored as follows: +, no degradation; ±, partial degradation; and −, complete degradation.

**Manipulation of internal CTD repeats does not induce cleavage to IIb.** To confirm that this effect is specific to the C-terminal amino acids of CTD52, five mutants with a variety of deletions and substitutions were obtained from J. Corden and recloned into our system (2). Their composition and expression patterns are shown in Fig. 1, lower panel. Mutants 7 and 8 both lack CTD52 and are both degraded to IIb. Mutants 9 to 11 contain severe deletions or substitutions within the CTD, but CTD52 is present. These mutants were not degraded, supporting the importance of CTD52 for CTD stability. In contrast, neither the length of the CTD, nor the ratio of consensus to nonconsensus repeats affected stability. The reduction in the amount of mutant 8 (both IIa and IIo forms) compared to other mutants increases with time cultured in the presence of α-amanitin. Its 36 repeats and instability are probably limiting to its own expression (30). Similarly, mutant 11 consistently demonstrated low levels of the IIo form compared to IIa. This mutant diverges largely from the consensus sequence at position 7 (Ser/Lys) in almost every repeat, suggesting that the choice of residue at this position may influence CTD modifications.

**Removal of CTD repeats 1 to 3 also induces degradation to Pol IIb.** Size comparison of the IIb form with truncated Pol II mutants suggests that the CTD is proteolytically cleaved at, or up to a site within the first few CTD repeats. Interestingly, repeat 2 has a unique sequence not present elsewhere in the CTD, a potential unique site for proteolytic cleavage. To investigate whether this is indeed the case, repeats 1 to 3 were deleted in both *wt* and CTD-truncated (Δ50) polymerases, as depicted in Fig. 2 (mutants 1 and 5). If a protease cuts the CTD in this region, its deletion should rescue the phenotype of Δ50. Expression profiles show, however, that this is not the case. Moreover, deletion of repeats 1 to 3 can itself induce degradation to the Pol IIb form, even in the presence of CTD52 (Fig. 2, mutant 1). This is most intriguing, given that other internal mutations had no effect on stability (Fig. 1, mutants 7 to 11). In the absence of both repeats 1 to 3 and CTD52 (mutant 5), the effect appears to be synergistic: almost all of the polymerase is converted to the IIb form, a finding suggestive of a functional interaction between these two sequences. To investigate the possibility of a specific interaction, point mutations of the most likely candidate, the unique repeat 2, were produced (residues were mutated back to the consensus sequence: E2S, R4T, and GG7S) in both wt and Δ50 backgrounds, but no increased degradation could be observed (data not shown). It is possible, however, that these mutations were too subtle or that repeat 2 alone is not sufficient for CTD stability.

**A random peptide can functionally replace repeats 1 to 3.** To further test the hypothesis that a functional interaction between repeats 1 to 3 and CTD52 is required for CTD's stability, two linkers were produced, encoding a motif from Abl1 previously shown to bind CTD52 (AblCTD-ID) (3), and a random sequence of the same length (the Abl linker in antisense). Could the AblCTD-ID motif replace repeats 1 to 3? Strikingly, in the *wt* background both the AblCTD-ID and the random linker replacement could prevent the degradation pattern seen when repeats 1 to 3 were deleted (Fig. 2, mutants 2 and 3). To our great surprise, the AblCTD-ID completely prevented degradation of a mutant where both repeats 1 to 3 and CTD52 are absent (mutant 6). In contrast, the random replacement (mutant 7) in the same Δ50 background could not prevent degradation to IIb, although the CTD seems to be more stable compared to the deletion of repeats 1 to 3 (mutant 5).

Notably, introduction of the AblCTD-ID (mutants 2 and 6) causes a shift in the migration pattern of both *wt* and Δ50 polymerases to a IIo form of higher apparent molecular weight. This band is highly reactive with the anti-phospho-CTD antibodies H5 and H14 (data not shown). We cannot exclude that it is this hyperphosphorylation event that protects these mutants from degradation. A previous study (17) described the importance of CTD52 for the splicing and 3′ processing of pre-mRNA and showed that its exact position within the CTD is irrelevant for its function. We therefore replaced repeats 1 to 3 with a linker consisting of CTD52 sandwiched between two consensus repeats (Fig. 2, mutants 4 and 8). In the *wt* background (mutant 4), this linker can replace repeats 1 to 3 without inducing degradation. In the background of Δ50 (mutant 8), this linker prevents the complete degradation seen with mutant 5 but cannot provide full protection from degradation.

In summary, these data suggest that degradation is only induced in the *wt* background when several consensus repeats replace the nonconsensus repeats 1 to 3 (mutant 1).

**EGFP prevents degradation to IIb in the absence of both repeats 1 to 3 and 52.** In Fig. 1, we demonstrated that the six C-terminal amino acids of CTD52 could be functionally replaced by acidic amino acids or by a fusion with hTRX or EGFP. It is not obvious, however, how these mutants can prevent degradation. Can EGFP also rescue the degradation induced by deletion of repeats 1 to 3? Figure 3A shows a panel of mutants with both single and double deletions of repeats 1 to 3 and CTD52. EGFP prevents degradation to IIb in all combinations. The fusion with EGFP decreases the rate of migration in sodium dodecyl sulfate-polyacrylamide gel electrophoresis compared to *wt*; unusually, a large proportion of

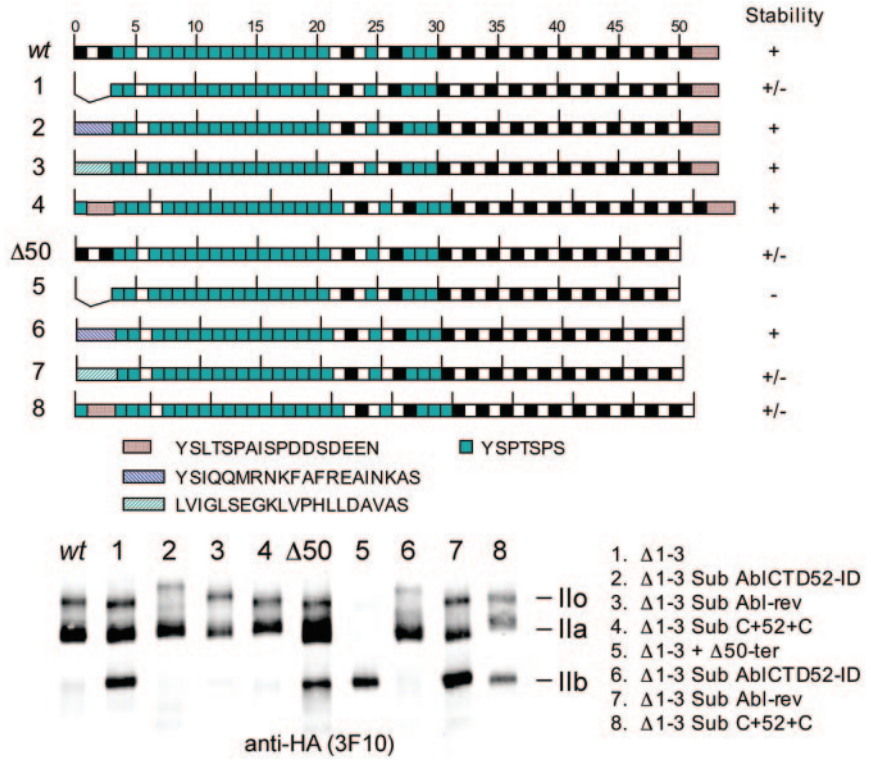


FIG. 2. Replacement of repeats 1 to 3 by consensus repeats induces degradation to Pol IIb. Deletion and substitution mutants of repeats 1 to 3 were created both in the *wt* and Δ50 backgrounds. Mutants 2 and 6 contain a substitution of repeats 1 to 3 for a motif from Abl1 (CTD-ID), known to interact with CTD52 (sequences shown below). Similarly, mutants 3 and 7 contain a random linker of the same length. In mutants 4 and 8, repeats 1 to 3 were replaced by repeat 52 flanked at each side by a consensus repeat. Stable Raji cell lines were produced for each mutant; after expression, cells were cultivated 48 h in the presence of α-amanitin before being harvested, and the different forms produced by each mutant were separated by polyacrylamide gel electrophoresis, followed by Western analysis. The Western blot reveals the different forms produced by each mutant; stability is summarized in the right-hand column.

these mutants remains in a form migrating somewhere between Pol IIa and IIo. However, this finding seems not to be important in vivo since Raji B-cell lines expressing these mutants could be cultured indefinitely in the presence of α-amanitin.

**Acidic heptapeptides cannot prevent degradation in the absence of repeats 1 to 3.** Do the mutants ending in acidic heptapeptides behave like EGFP fusions or the *wt* CTD when repeats 1 to 3 are deleted? These mutants exhibit the same expression profile as those with or without CTD52 (Fig. 3B). The basic hexapeptide cannot prevent degradation in the presence of repeats 1 to 3 (mutant 7), and in their absence Pol II is totally degraded to IIb (mutant 8). The acidic hexapeptide, and that from *D. melanogaster* polymerase (mutants 9+10 and 11+12, respectively) behave like the *wt* polymerase and a polymerase lacking repeats 1 to 3. This further suggests that it is the acidic nature of the seven C-terminal amino acids of repeat 52, rather than their specific sequence, that is important. Importantly, these sequences protect against degradation differently from EGFP.

**The seven C-terminal amino acids of Pol II CTD are dispensable for pre-mRNA processing.** To confirm the results seen in Raji cells, CTD mutants fused to EGFP were subcloned back into the vector pSTC-TK-Pol II LS*, originally supplied by Walter Schaffner (18) (Fig. 4A). Expression from this vector is constitutive and under the control of the thymidine kinase promoter. For microscopy, an adherent cell line was chosen. Stable HeLa cell lines were produced by selection of these mutants with α-amanitin. No difference in the nuclear distribution of the various mutants could be seen (data not shown). Northern blot analysis of total RNA extracted from these cell lines was compared to nontransfected control cells cultured in α-amanitin for 48 h (Fig. 4B). The control cell line (K) shows degradation of total cellular RNA, likely resulting from exposure to α-amanitin. The expression of five endogenous genes was examined. The PES1 and YY1 mRNAs show partial stability in the presence of α-amanitin in all of the HeLa cell lines tested here. The lack of slow-migrating pre-mRNA in the control indicates that fresh Pol II transcription is blocked in these cells. If the seven C-terminal amino acids of the *wt* CTD or repeats 1 to 3 were required for efficient pre-mRNA processing we would expect an accumulation of precursor compared to processed RNA. This is not the case: both mutants behave like Pol II *wt* EGFP.

**Artificial CTDs allow the improved study of consensus and nonconsensus repeats.** Overlooking the effect of CTD mutation on stability has compromised previous studies of CTD function. Given our findings that repeats 1 to 3 and 52 are essential for stability, but not viability and pre-mRNA processing, it was possible to produce stable CTD mutants where

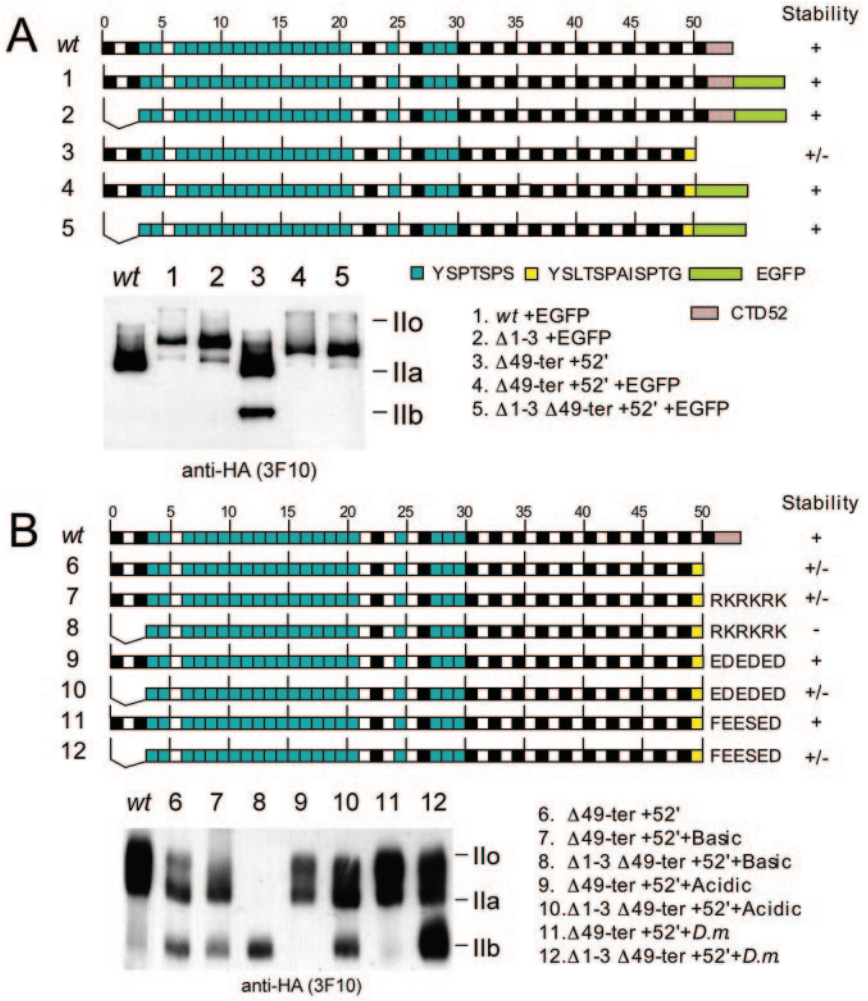


FIG. 3. Fusions with EGFP, but not acidic polypeptides, prevent degradation in the absence of repeats 1 to 3 and 52. (A) EGFP was cloned at the C terminus of mutants containing deletions of repeats 1 to 3, CTD52, and in combination. (B) Repeats 1 to 3 were deleted in the background of truncated CTD52 and fusions thereof with acidic or basic hexapeptides. The sequence used in mutants 11 and 12 corresponds to the C-terminal amino acids of *D. melanogaster* CTD. Stable Raji cell lines were produced after expression, cells were cultivated 48 h in the presence of α-amanitin before being harvested, and the different forms produced by each mutant were separated by polyacrylamide gel electrophoresis, followed by Western analysis; stability is summarized in the right-hand column.

repeats 4 to 51 were replaced by synthetic CTD repeats. Although EGFP could also provide stability, the mechanism by which it does so is apparently different from that used by repeats 1 to 3 and 52 and could potentially affect our results. The aim was to create a mutant with repeats 1 to 3, 48 YSPTSPS repeats, and CTD52—a total of 52 repeats, like *wt*. However, upon sequencing we discovered that recombination events in bacteria had altered the expected repeat number. Figure 5A shows three such mutants containing 20, 35, and 55 pure, consensus, YSPTSPS repeats (mutants 1, 2, and 3, respectively). The schematic shows the position of consensus repeats (turquoise) in each mutant compared to *wt* and Δ50. Expression analysis confirms that including repeats 1 to 3 and 52 prevents degradation. These mutants were compared against each other and to two mutants known not to support long-term viability: Δ50 and a mutant containing 34 nonconsensus repeats (mutant 4, which also contains repeats 1, 2, and 52) (2, 9). An analysis of their ability to support the proliferation (B) and viability (C) of cell lines under α-amanitin selection reveals a general effect: as previously observed (9, 30), cells growing with α-amanitin-resistant polymerases initially grow slowly, with reduced viability; this is probably a negative effect of overexpression. Raji cells overexpress c-*myc* and regularly undergo apoptosis; normal cell cultures often contain up to 5% dead cells. The cumulative cell number accounts for cell culture expansion over the course of the experiment, whereas viability measures the amount of cell death. Taken together, these data reveal the ability of CTD mutants to support the long-term growth and survival of cells. It must be noted, however, that differences in cumulative cell number do not necessarily mean reduced cell division, since it does not take into account cell death.

Comparison of an α-amanitin-resistant polymerase with a *wt* CTD to a mutant containing 55 consensus repeats revealed an additional "adjustment" phase, after which the rate at which viable cells were produced synchronized with that of the *wt*

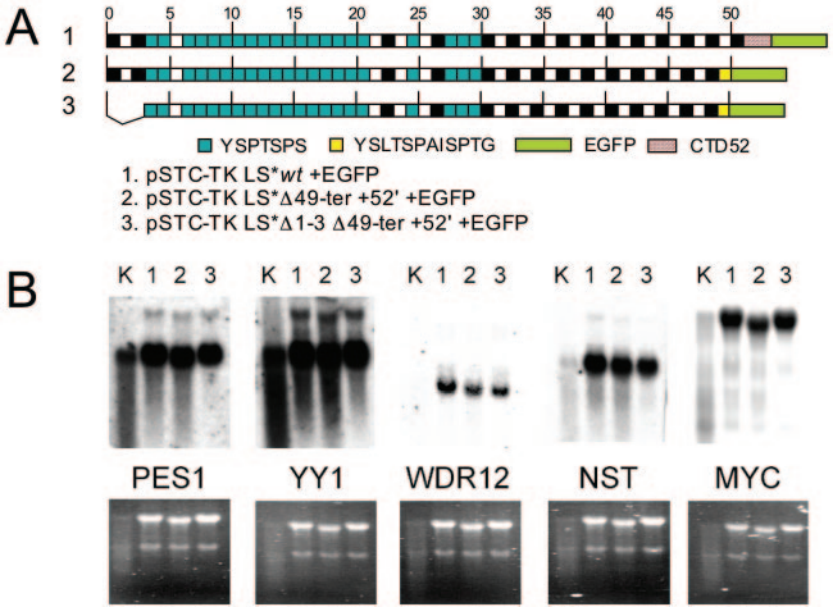


FIG. 4. Deletion of repeats 1 to 3 and CTD52 has no apparent effect on RNA processing, cell growth, and viability. (A) Deletion mutants of repeats 1 to 3 and CTD52 fused with EGFP were recloned into the expression vector pSTC-TK under the control of the constitutive thymidine-kinase promoter. HeLa cells transfected with these mutants were cultured in the presence of α-amanitin for several weeks, after which total RNA was extracted. As a control (K), nontransfected cells were cultured in α-amanitin for 48 h before RNA extraction. (B) The abundance of five different RNAs examined (upper panel), and the 18S and 28S rRNAs (lower panel) are shown.

(cumulative cell number). The *wt* CTD contains 21 consensus repeats and is viable; however, mutant 1 has 20 consensus repeats and cannot sustain cell viability. Mutant 3 has 55 consensus repeats and can support long-term growth and viability. These data show that, despite their initial differences, the length of the CTD is more important than the number of consensus repeats. Interestingly, despite containing 10 repeats more than mutant 1, cells expressing mutant 4 have, in comparison, a much reduced ability to support growth and viability (Fig. 5B and C). The consensus repeats must therefore have a unique function that cannot be performed by the nonconsensus repeats. Mutant 2, with 35 consensus repeats, has an intermediate phenotype: cells can be cultured for many weeks, but viability remains low. This experiment was repeated several times, and always the same mutants survived or died.

These data suggest that, at least in cell culture, a CTD composed of 55 consensus repeats can fulfill all essential functions for proliferation.

## DISCUSSION

The CTD functionally connects transcription and the pre-mRNA processing apparatus (reviewed in references 14, 20, 22, 40, and 50). To address the importance of the CTD sequence for this function, we performed an extensive genetic analysis of unique sequences within the mammalian CTD. We identified two regions which, when removed, lead to the degradation of Pol II to its CTD-less, IIb form. Given the severe effect of CTD removal on transcription and pre-mRNA processing, it is possible that inducing CTD degradation may have a role in an as-yet-undiscovered stress response; it has recently been shown that BRCA1/BARD1 can induce the ubiquitination and degradation of Pol II LS after DNA damage by UV (25). Degradation to IIb is specific to the removal of the repeats 1 to 3 and CTD52 only and not any other sequence within the CTD. Removal of both sequences simultaneously is more severe than individually, suggesting a relationship between the two regions. Mutagenesis of CTD52 revealed a requirement for acidic amino acids at its C terminus, irrelevant of their specific sequence. The sequence of repeats 1 to 3 shows specificity only in that it is not composed of consensus repeats: it can be replaced by any random sequence but not by YSPTSPS repeats. The lack of sequence specificity in regulating this phenomenon, suggests that repeats 1 to 3 act as a spacer between the "Linker" region and the CTD rather than being the binding site for a specific factor (for Pol II architecture, see reference 13). Deletion of repeats 1 to 3 shifts consensus repeats in to their place and could cause a different CTD structure, thereby exposing the CTD to proteolytic attack or preventing a factor binding to the Linker that is involved in regulating CTD stability.

In an unusual mutant where repeats 1 to 3 were exchanged for the CTD52-binding motif of Abl1 (Abl1CTD-ID), degradation to the IIb form could be totally prevented in the absence of CTD52, a unique property among our many mutants. This mutant also exhibited a different modification pattern, which may protect against degradation. However, whether this relates to phosphorylation per se or a specific modification is currently difficult to determine; also, some mutants demonstrate reduced amounts of the IIo form but are nevertheless stable. Since the activity of Abl1 on Pol II is dependent on double-stranded DNA breaks and its specific inhibitor STI571 or ionizing radiation fail to induce IIb (data not shown), it is unlikely that Abl1 tyrosine kinase activity is responsible for this

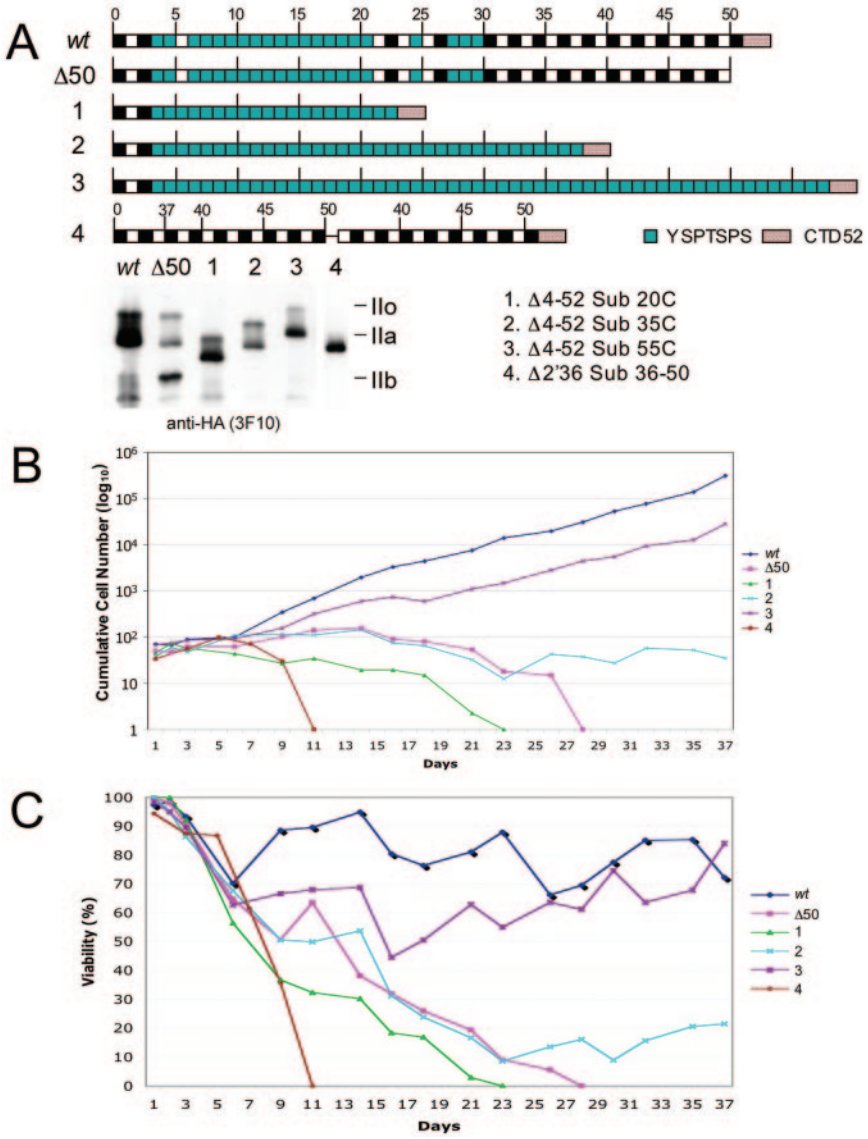


FIG. 5. Repeats 4 to 51 can be replaced by consensus repeats and still support growth and viability. (A) Artificial CTDs were produced containing 20 (mutant 1), 35 (mutant 2), and 55 (mutant 3) consensus repeats in place of repeats 4 to 51. The ability of these mutants to replace the endogenous Pol II LS was tested by cultivation of cells in the presence of 2 μg of α-amanitin/ml. The cumulative growth (cell number $\times 10^4$; B) and viability (C) of these cells were monitored over a period of weeks. A mutant known to be viable (*wt*) and two nonviable mutants (Δ50 and mutant 4) were included as controls. The number of living ($N_1$) and dead cells ($N_d$) was determined by trypan blue staining. The percentage of viable cells ($V$) was calculated by using the formula $V = 100 \times N_1(N_1 + N_d)$. Cumulative growth was calculated by multiplying the number of cells present by the factor by which the culture was split over the course of the experiment.

effect. Further work is required to determine what this motif binds in addition to CTD52.

The severe degradation caused by removal of repeats 1 to 3 and CTD52 could be rescued by EGFP but not an acidic polypeptide at the C terminus. This would suggest that EGFP prevents degradation via a different mechanism to CTD52. Our results indicate that degradation occurs via an as-yet-unidentified exopeptidase, which cuts back from the C terminus: the 40-kDa fragment that should result from a single cleavage could not be detected with anti-CTD antibody, whereas a CTD-EGFP fragment induced by the cleavage of a tobacco etch virus site inserted between Linker and CTD could be detected (data not shown). The structure of EGFP may prevent access of the exopeptidase to CTD, whereas CTD52 may depend on interactions with other proteins to prevent degradation.

Importantly, the stability provided by EGFP allowed the production of cell lines expressing mutants lacking repeats 1 to 3 and CTD52 under α-amanitin. That these cell lines could be maintained indefinitely (>2 months in culture) demonstrates that these motifs are not essential to Pol II function. Analysis of their ability to process endogenous RNAs revealed no difference compared to a *wt* EGFP mutant. Our data are at odds with previous work demonstrating CTD52 to be essential for pre-mRNA processing (17). This can be explained through the different systems used to analyze Pol II function. In our system,

we analyzed the effect of CTD mutation on the ability of Pol II to transcribe and process the RNA from endogenous genes, whereas other studies used transiently transfected reporter assays. The degradation of Pol II LS to the IIb form produces a CTD-less polymerase that, although unable to transcribe chromatin, can transcribe "naked" transfected reporter genes (30). Although the CTD has been shown to bind many factors in vitro, it is not so clear whether the interaction with the CTD is always essential for their function: a CTD-less polymerase can stimulate capping in the absence of CTD, but the presence of CTD improves the efficiency of the reaction fourfold (35); of the CTD-interacting proteins described to date, only the yeast U1snRNP component Prp40p has demonstrated splicing activity, but as yet no metazoan homologue has been discovered (32, 37); CTD enhances but is not essential for RNA cleavage and polyadenylation in vivo (28); a polymerase with just five repeats is able to induce the albeit reduced expression of luciferase protein from a transiently transfected reporter in vivo (30).

The importance of these processes occurring cotranscriptionally is a topic of hot debate (1, 27, 37). There is good evidence that the CTD localizes components of the splicing machinery to sites of transcription (31, 34), but why the extreme conservation of the CTD sequence across mammals, when the only major phenotype of the mouse with a 13 repeat deletion was its small size, and the higher rate of neonatal lethality (29)? Several other studies suggest that certain CTD functions are dependent on its length (41), but that specific nonconsensus sequences may enhance this function (42). Through this work we have found that those nonconsensus sequences not affecting stability can be functionally replaced by consensus repeats. The effect on cell viability and growth appears to be dependent on the number of repeats: the greater the number of consensus repeats, the greater the rate of proliferation and cell survival. As previously noted, a truncation of the CTD to 31 repeats or less cannot support long-term viability (30), but a mouse with 39 repeats could be produced (29). In the present study, a mutant with 39 repeats (of which 35 are consensus) demonstrated a growth defect and starkly reduced viability; however, cells could be cultured continuously under α-amanitin selection. In contrast, despite initially exhibiting a growth deficiency compared to *wt*, a mutant with 59 repeats (of which 55 are consensus) showed no long-term differences compared to *wt*, suggesting that the cells may compensate for a reduced function. Nevertheless, since these cells can be cultured continuously with viable "consensus" polymerases, we can assume that they are able to perform all of the essential functions required of them. The highly conserved composition of the mammalian CTD consensus and nonconsensus repeats is perhaps not entirely essential for life but rather "optimized" for efficient function or other, as-yet-unknown, purposes that could convey a survival advantage (e.g., stress responses). Small improvements in the efficiency of such an important enzyme could account for its high evolutionary conservation. A mutant composed of nonconsensus repeats had a severe growth defect compared to mutants of similar length, containing consensus repeats. That these repeats differ from the consensus sequence almost uniquely at Ser7 is suggestive of an important role for this position. Interestingly, the crystal structures of several factors bound to the CTD thus far have not indicated that this position plays a definitive role in their binding. It is possible that the nonconsensus repeats are important for the expression of specific genes (perhaps during cellular differentiation), alternative splicing, or stress responses. Producing transgenic mice with "all-consensus" repeats would resolve this discussion. The mice produced by J. Corden's group (29) have an apparent defect in growth and not development, but other, more subtle defects may have gone undetected.

Using our system, synthetic CTDs can be produced and their function tested in vivo on the endogenous chromatin template. For the first time, this approach allows CTDs of the same length, but of different composition, to be accurately compared for their abilities to support transcription and the processing of RNA, while avoiding the adverse effects of variations in length and instability.

## ACKNOWLEDGMENTS

We thank E. Kremmer for antibodies, and SEQUISERVE for optimizing the sequencing of repetitive CTD templates. We are grateful to A. Meinhardt and M. Meisterernst for valuable criticisms and suggestions.

This study was supported by grant SFB/Transregio 5 from the Deutsche Forschungsgemeinschaft.

## REFERENCES

1. **Ares, M., Jr., and N. J. Proudfoot.** 2005. The Spanish connection: transcription and mRNA processing get even closer. Cell **120:**163–166.
2. **Bartolomei, M. S., N. F. Halden, C. R. Cullen, and J. L. Corden.** 1988. Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. Mol. Cell. Biol. **8:**330–339.
3. **Baskaran, R., G. G. Chiang, and J. Y. Wang.** 1996. Identification of a binding site in c-Abl tyrosine kinase for the C-terminal repeated domain of RNA polymerase II. Mol. Cell. Biol. **16:**3361–3369.
4. **Baskaran, R., M. E. Dahmus, and J. Y. Wang.** 1993. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. Proc. Natl. Acad. Sci. USA **90:**11167–11171.
5. **Baskaran, R., S. R. Escobar, and J. Y. Wang.** 1999. Nuclear c-Abl is a COOH-terminal repeated domain (CTD)-tyrosine (CTD)-tyrosine kinase-specific for the mammalian RNA polymerase II: possible role in transcription elongation. Cell Growth Differ. **10:**387–396.
6. **Bourquin, J. P., I. Stagljar, P. Meier, P. Moosmann, J. Silke, T. Baechi, O. Georgiev, and W. Schaffner.** 1997. A serine/arginine-rich nuclear matrix cyclophilin interacts with the C-terminal domain of RNA polymerase II. Nucleic Acids Res. **25:**2055–2061.
7. **Buratowski, S.** 2003. The CTD code. Nat. Struct. Biol. **10:**679–680.
8. **Chang, A., S. Cheang, X. Espanel, and M. Sudol.** 2000. Rsp5 WW domains interact directly with the carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. **275:**20562–20571.
9. **Chapman, R. D., B. Palancade, A. Lang, O. Bensaude, and D. Eick.** 2004. The last CTD repeat of the mammalian RNA polymerase II large subunit is important for its stability. Nucleic Acids Res. **32:**35–44.
10. **Cho, E. J., T. Takagi, C. R. Moore, and S. Buratowski.** 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev. **11:**3319–3326.
11. **Corden, J. L., D. L. Cadena, J. M. Ahearn, Jr., and M. E. Dahmus.** 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA **82**:7934–7938.
12. **Corden, J. L., and M. Patturajan.** 1997. A CTD function linking transcription to splicing. Trends Biochem. Sci. **22:**413–416.
13. **Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, and R. D. Kornberg.** 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. Science **288:**640–649.
14. **Cramer, P., A. Srebrow, S. Kadener, S. Werbajh, M. de la Mata, G. Melen, G. Nogues, and A. R. Kornblihtt.** 2001. Coordination between transcription and pre-mRNA processing. FEBS Lett. **498:**179–182.
15. **Dichtl, B., D. Blank, M. Sadowski, W. Hubner, S. Weiser, and W. Keller.** 2002. Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. EMBO J. **21:**4125–4135.
16. **Fong, N., and D. L. Bentley.** 2001. Capping, splicing, and 3′ processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev. **15:**1783–1795.
17. **Fong, N., G. Bird, M. Vigneron, and D. L. Bentley.** 2003. A 10 residue motif at the C terminus of the RNA Pol II CTD is required for transcription, splicing and 3′ end processing. EMBO J. **22:**4274–4282.

18. **Gerber, H. P., M. Hagmann, K. Seipel, O. Georgiev, M. A. West, Y. Litingtung, W. Schaffner, and J. L. Corden.** 1995. RNA polymerase II C-terminal domain required for enhancer-driven transcription. Nature **374:**660–662.
19. **Goldstrohm, A. C., T. R. Albrecht, C. Sune, M. T. Bedford, and M. A. Garcia-Blanco.** 2001. The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. Mol. Cell. Biol. **21:**7617–7628.
20. **Hirose, Y., and J. L. Manley.** 2000. RNA polymerase II and the integration of nuclear events. Genes Dev. **14:**1415–1429.
21. **Ho, C. K., V. Sriskanda, S. McCracken, D. Bentley, B. Schwer, and S. Shuman.** 1998. The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. **273:**9577–9585.
22. **Howe, K.** 2002. RNA polymerase II conducts a symphony of pre-mRNA processing activities. Biochim. Biophys. Acta **1577:**308–324.
23. **Kelly, W. G., M. E. Dahmus, and G. W. Hart.** 1993. RNA polymerase II is a glycoprotein: modification of the COOH-terminal domain by O-GlcNAc. J. Biol. Chem. **268:**10416–10424.
24. **Kim, M., N. J. Krogan, L. Vasiljeva, O. J. Rando, E. Nedea, J. F. Greenblatt, and S. Buratowski.** 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature **432:**517–522.
25. **Kleiman, F. E., F. Wu-Baer, D. Fonseca, S. Kaneko, R. Baer, and J. L. Manley.** 2005. BRCA1/BARD1 inhibition of mRNA 3′ processing involves targeted degradation of RNA polymerase II. Genes Dev. **19:**1227–1237.
26. **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.
27. **Kornblihtt, A. R., M. de la Mata, J. P. Fededa, M. J. Munoz, and G. Nogues.** 2004. Multiple links between transcription and splicing. RNA **10:**1489–1498.
28. **Licatalosi, D. D., G. Geiger, M. Minet, S. Schroeder, K. Cilli, J. B. McNeil, and D. L. Bentley.** 2002. Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. Mol. Cell **9:**1101–1111.
29. **Litingtung, Y., A. M. Lawler, S. M. Sebald, E. Lee, J. D. Gearhart, H. Westphal, and J. L. Corden.** 1999. Growth retardation and neonatal lethality in mice with a homozygous deletion in the C-terminal domain of RNA polymerase II. Mol. Gen. Genet. **261:**100–105.
30. **Meininghaus, M., R. D. Chapman, M. Horndasch, and D. Eick.** 2000. Conditional expression of RNA polymerase II in mammalian cells. Deletion of the carboxyl-terminal domain of the large subunit affects early steps in transcription. J. Biol. Chem. **275:**24375–24382.
31. **Misteli, T., and D. L. Spector.** 1999. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol. Cell **3:**697–705.
32. **Morris, D. P., and A. L. Greenleaf.** 2000. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. **275:**39935–39943.
33. **Morris, D. P., H. P. Phatnani, and A. L. Greenleaf.** 1999. Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3′-end formation. J. Biol. Chem. **274:**31583–31587.
34. **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.
35. **Moteki, S., and D. Price.** 2002. Functional coupling of capping and transcription of mRNA. Mol. Cell **10:**599–609.
36. **Naar, A. M., D. J. Taatjes, W. Zhai, E. Nogales, and R. Tjian.** 2002. Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. Genes Dev. **16:**1339–1344.
37. **Neugebauer, K. M.** 2002. On the importance of being co-transcriptional. J. Cell Sci. **115:**3865–3871.
38. **Patturajan, M., X. Wei, R. Berezney, and J. L. Corden.** 1998. A nuclear matrix protein interacts with the phosphorylated C-terminal domain of RNA polymerase II. Mol. Cell. Biol. **18:**2406–2415.
39. **Payne, J. M., and M. E. Dahmus.** 1993. Partial purification and characterization of two distinct protein kinases that differentially phosphorylate the carboxyl-terminal domain of RNA polymerase subunit IIa. J. Biol. Chem. **268:**80–87.
40. **Proudfoot, N. J., A. Furger, and M. J. Dye.** 2002. Integrating mRNA processing with transcription. Cell **108:**501–512.
41. **Rosonina, E., and B. J. Blencowe.** 2004. Analysis of the requirement for RNA polymerase II CTD heptapeptide repeats in pre-mRNA splicing and 3′-end cleavage. RNA **10:**581–589.
42. **Ryan, K., K. G. Murthy, S. Kaneko, and J. L. Manley.** 2002. Requirements of the RNA polymerase II C-terminal domain for reconstituting pre-mRNA 3′ cleavage. Mol. Cell. Biol. **22:**1684–1692.
43. **Sims, R. J., III, R. Belotserkovskaya, and D. Reinberg.** 2004. Elongation by RNA polymerase II: the short and long of it. Genes Dev. **18:**2437–2468.
44. **Steinmetz, E. J., N. K. Conrad, D. A. Brow, and J. L. Corden.** 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature **413:**327–331.
45. **Stiller, J. W., and B. D. Hall.** 2002. Evolution of the RNA polymerase II C-terminal domain. Proc. Natl. Acad. Sci. USA **99:**6091–6096.
46. **Tanner, S., I. Stagljar, O. Georgiev, W. Schaffner, and J. P. Bourquin.** 1997. A novel SR-related protein specifically interacts with the carboxy-terminal domain (CTD) of RNA polymerase II through a conserved interaction domain. Biol. Chem. **378:**565–571.
47. **Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden.** 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc. Natl. Acad. Sci. USA **93:**6975–6980.
48. **Zhang, F., M. Barboric, T. K. Blackwell, and B. M. Peterlin.** 2003. A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. Genes Dev. **17:**748–758.
49. **Zhang, J., and J. L. Corden.** 1991. Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. J. Biol. Chem. **266:**2290–2296.
50. **Zorio, D. A., and D. L. Bentley.** 2004. The link between mRNA processing and transcription: communication works both ways. Exp. Cell Res. **296:**91–97.