Transcriptional profiling of fission yeast RNA polymerase II CTD mutants

ANGAD GARG,1 ANA M. SANCHEZ,1,2 BEATE SCHWER,3 and STEWART SHUMAN1

1Molecular Biology Program, Sloan-Kettering Institute, New York, New York 10065, USA
2Gerstner Sloan Kettering Graduate School of Biomedical Sciences, New York, New York 10065, USA
3Department of Microbiology and Immunology, Weill Cornell Medical College, New York, New York 10065, USA

ABSTRACT
The carboxyl-terminal domain (CTD) of RNA polymerase II consists of tandem repeats of heptapeptide Y1S2P3T4S5P6S7. The CTD recruits proteins that drive or regulate gene expression. The trafficking of CTD-interacting proteins is orchestrated by remodeling CTD primary structure via Ser/Thr/Tyr phosphorylation and proline cis-trans isomerization, which collectively inscribe a CTD code. The fission yeast CTD consists of 29 repeats. To decipher the output of the fission yeast CTD code, we manipulated CTD length and amino acid content and gauged effects of these changes on gene expression. Whereas deleting 11 heptads has no effect on yeast growth, RNA-seq revealed that 25% of protein-coding transcripts were dysregulated. We profiled the transcriptomes of full-length CTD mutants, in which: all Tyr1 residues were replaced by Phe; all Ser2, Thr4, or Ser7 positions were changed to Ala; and half of the essential CTD code “letters” Pro3, Ser5, and Pro6 were mutated to Ala. Overlapping RNA-seq profiles suggested that a quarter of the up-regulated mRNAs and half of the down-regulated mRNAs seen in full-length CTD mutants might be attributable to a decrement in wild-type heptad number. Concordant mutant-specific profiles were observed for Y1F, S2A, and T4A cells, and for P6P6A and S5S5A cells, suggesting that Tyr1–Ser2–Thr4 and Ser5–Pro6 comprise distinct “words” in the CTD code. The phosphate regulon, which is repressed by IncRNA-mediated transcription interference, is de-repressed by CTD mutations P6P6A and S5S5A. De-repression of pho1 in P6P6A and S5S5A cells depends on cleavage and polyadenylation factor subunits Swd22 and Ppn1 and termination factor Rhn1, signifying that Pro6 and Ser5 mutations elicit precocious IncRNA 3′-processing/termination.

Keywords: CTD code; RNA 3′-processing; gene expression; transcription termination

INTRODUCTION
The carboxyl-terminal domain (CTD) of the Rpb1 subunit of RNA polymerase II (Pol2) consists of tandem repeats of a consensus heptapeptide Y1S2P3T4S5P6S7. The CTD provides a landing pad for the recruitment of proteins that regulate transcription initiation, elongation and termination, modify chromatin structure, and catalyze or regulate RNA capping, splicing, and polyadenylation (Corden 2013; Eick and Geyer 2013; Jeronimo et al. 2013). The CTD primary structure is tuned during the transcription cycle by phosphorylation and dephosphorylation of the heptad serine, threonine, and tyrosine residues and by cis–trans isomerization of the prolines. These variations inscribe a “CTD code” that conveys information about the transcription machinery, which is “read” by a slew of CTD-binding proteins. Insights into CTD coding principles have been gained by: (i) elucidating how individual proteins recognize the CTD, and (ii) genetically manipulating the CTD primary structure (e.g., by varying CTD length and amino acid content) and assessing the effects of these changes on cell physiology.

The number of CTD heptad repeats (n) varies among taxa, with avian and rodent Plasmodium species (n = 8) and the microsporidian Encephalitozoon cuniculi (n = 15) on the low end and mammals (n = 52) and Hydra (n = 56) at the high end. Cell growth in any given taxon is contingent on a minimum CTD length, which is less than the native CTD heptad number, and the minimum heptad number increases with evolutionary complexity of the
organism. For example, the fission yeast *Schizosaccharomyces pombe* CTD has 29 repeats (Fig. 1A). The junction CTD segment to the body of Rpb1 consists of four repeats that deviate in size and/or sequence from the consensus heptad; this segment is referred to as the CTD “rump.” Distal to the rump is an array of 25 heptad repeats that conform perfectly to the YSPTSPS consensus, with the single exception of an alanine in lieu of Pro3 in the fifth heptad downstream from the rump. The CTD length requirements for fission yeast growth were assessed by serially truncating the heptads from the carboxyl terminus of Rpb1, which revealed that: (i) Eight repeats (comprising the rump plus four consensus heptads) was lethal; (ii) 10-13 repeats (rump plus six to nine consensus heptads) resulted in slow growth and cold-sensitive phenotypes; and (iii) 16 (rump plus six to nine consensus heptads) revealed that: (i) Eight repeats (comprising the rump plus four consensus heptads) was lethal; (ii) 10-13 repeats (rump plus six to nine consensus heptads) resulted in slow growth and cold-sensitive phenotypes; and (iii) 16 or more repeats (rump plus 12 or more consensus heptads) sufficed for normal growth under all laboratory conditions tested (Schneider et al. 2010).

The in vivo requirements for all amino acids of the Y1S2P3T4S5P6S7 repeat (each of the seven amino acids being a “letter” in the CTD code) were gauged initially by introducing alanine and conservative substitutions in lieu of Tyr1, Ser2, Pro3, Thr4, Ser5, Pro6, or Ser7 residues of consensus heptads of the CTD array in the context of a truncated version of the Rpb1 CTD containing 18 repeats (Fig. 1A) that supported normal cell growth (Schwer and Shuman 2011; Schwer et al. 2012). Those studies revealed that: (i) Tyr1, Ser5, and Pro6 are essential for viability of fission yeast, by the criterion that alanine substitution is lethal, whereas Ser2, Thr4, and Ser7 are inessential; and (ii) Y1F, Y1F + S2A, Y1F + S7A, S2A + T4A, S2A + S7A, and T4A + S7A mutants are viable, signifying that phenylalanine is functional in lieu of Tyr1 and Ser5 is the only strictly essential phosphorylation site. The singular essentiality of Ser5–PO4 in fission yeast, and of Pro6 that is required for Ser5 phosphorylation, is linked to recruitment of the mRNA capping apparatus to the Pol2 elongation complex, insofar as the requirements for Ser5–PO4 and Pro6 can be bypassed by covalently fusing the capping enzyme to Pol2 (Schwer and Shuman 2011; Schwer et al. 2012).

The ability of *S. pombe* to grow when the Tyr1, Ser2, Thr4, or Ser7 residues of consensus heptads are uniformly replaced by a nonphosphorylatable side chain accorded with an initial transcriptome analysis by poly(A)+ RNA-seq, which showed that only a small fraction of fission yeast mRNAs is dysregulated by CTD phospho-site mutations (Schwer et al. 2014). To wit, CTD mutations S2A, Y1F, S7A, and T4A elicited greater than or equal to twofold dysregulation of 4.4%, 1.4%, 1.2%, and 0.14% of the annotated fission yeast protein-coding RNAs, respectively. The majority of the mRNAs dysregulated in Y1F cells were coordinateley affected by S2A, suggesting that Tyr1–Ser2 constitutes a two-letter “word” in the CTD code. RNA-seq and cell-based functional assays showed that Y1F and S2A mutations elicited increased expression of genes encoding proteins involved in iron uptake (Schwer et al. 2014), implying that Tyr1 and Ser2 transduce a repressive signal to the iron regulon. Different CTD cues were found to govern phosphate homeostasis in fission yeast (Schwer et al. 2014, 2015; Chatterjee et al. 2016; Sanchez et al. 2018a). Inability to place a Ser7–PO4 mark (as in S7A) resulted in de-repression of a phosphate uptake (PHO) regulon in phosphate-replete medium. In contrast, indelible installation of a Ser7–PO4 mimetic (S7E) hyper-repressed PHO gene expression in phosphate-replete cells and attenuated induction during starvation. The PHO regulon was de-repressed by ablation of the CTD Ser5–PO4 mark, achieved either by mutating Ser5 in all consensus heptads to alanine, or replacing all Pro6 residues with alanine (which

![FIGURE 1. RNA-seq profiling of fission yeast cells with 18 versus 29 rpb1-CTD heptad repeats. (A) The amino acid sequences of the rpb1-CTD heptad arrays with 29 (full length; FL) and 18 repeats are shown. (B) The number of dysregulated coding and noncoding genes in rpb1-CTD-18 compared to rpb1-CTD-FL are specified on the top. A volcano plot of the P-value estimate plotted as a function of log2 fold-change for dysregulated coding genes is shown below. Examples of genes with high fold-change and/or low P-values are indicated.](www.rnajournal.org)
impedes Ser5 phosphorylation [Schwer et al. 2012]). These findings implicated Ser5, Pro6, and Ser7 as a three-letter code word that exerts a repressive effect on the PHO genes via serine phosphorylation. Subsequent studies showed that the fission yeast PHO genes are repressed in phosphate-replete cells by transcription in cis of 5′ flanking IncRNAs and that CTD mutations exert their effects on phosphate homeostasis by affecting the propensity of the nascent Pol2 IncRNAs to polyadenylate and terminate prior to reaching the flanking PHO mRNA promoters (for review, see Shuman 2020).

Several potential caveats pertain to the initial CTD phospho-site mutagenesis strategy: (i) the mutational effects were studied in the context of a CTD that contains 18 repeats rather than the native 29 repeats; (ii) mutations were introduced only into the consensus heptads; and (iii) the nonconsensus rump heptads were spared. Thus, it is possible that the transcriptomic or phenotypic impact of mutating CTD amino acids might be different in the context of the full length CTD and when the mutations apply to the rump as well as the consensus heptads. To address these caveats, we have constructed a series of viable S. pombe rpb1 mutant strains in which the native CTD length was maintained as 29 heptads (four rump and 25 consensus repeats) and the Tyr1, Ser2, Thr4, or Ser7 in every heptad was replaced by Phe, Ala, Ala, or Ala, respectively (Sanchez et al. 2018b). The full-length CTD phospho-site mutations impacted Pho1 expression similarly to what was seen in the truncated CTD contexts. We also constructed viable mixed heptad rpb1-CTD mutants—P3•P3A and P6•P6A—in which Pro3 or Pro6 was changed to alanine in every other heptad of the full-length CTD repeat array (Sanchez et al. 2020).

In the present study, we interrogate the transcriptomes of the expanded set of fission yeast rpb1-CTD strains as follows: (i) by gauging the effect of CTD truncation in a wild-type CTD heptad context; and (ii) by profiling full-length CTD mutants vis-à-vis an isogenic full-length wild-type control. Our results reveal multiple genes dysregulated by CTD truncation (affecting 25% of the annotated protein-coding transcripts) and coherent sets of genes that are coordinately impacted by individual CTD coding letters.

RESULTS AND DISCUSSION

RNA-seq profiling of fission yeast cells with 18 versus 29 CTD heptad repeats

Although truncation of the Rpb1 CTD to 18 repeats (rump plus 14 consensus heptads; Fig. 1) had no apparent effect on vegetative growth under standard laboratory conditions (Schneider et al. 2010), we wanted to assess the effect of CTD truncation on gene expression by comparing the transcriptome of the rpb1-CTD-18 strain to that of an identically marked full-length wild-type CTD control strain. We performed RNA-seq on poly(A)+ RNA isolated from cells with full-length and truncated CTDs. cDNAs obtained from three (full-length CTD) or four (truncated CTD) biological replicates, using RNA from cells grown to mid-log phase in YES medium at 30°C, were sequenced for each strain. A cutoff of plus or minus twofold change in normalized transcript read level and an adjusted P-value of ≤0.05 were the criteria applied to derive an initial list of differentially expressed annotated loci in the truncated versus full-length CTD. We then focused on differentially expressed genes with average normalized read counts ≥100 in either strain in order to eliminate transcripts that were expressed at very low levels in vegetative cells. We thereby identified: (i) 715 annotated protein-coding RNAs and 147 annotated noncoding RNAs that were up-regulated by these criteria in rpb1-CTD-18 cells compared to rpb1-CTD-FL cells; and (ii) 587 coding RNAs and 96 noncoding RNAs that were down-regulated (Fig. 1B). In toto, 25% (1302/5118) of the fission yeast protein-coding genes were affected by deletion of 11 consensus CTD heptads. The lists of UP and DOWN protein-coding genes are compiled in Supplemental Table S1. A volcano plot of the dysregulated coding genes is shown in Figure 1B. Gene Ontology analysis indicated that genes involved in nucleic acid metabolism were enriched in the up-regulated set (P < 1.52 × 10−45) and genes involved in cell wall organization or biosynthesis were enriched in the down-regulated set (P < 1.11 × 10−17).

With respect to the three-gene PHO regulon (Carter-O’Connell et al. 2012), tgp1 (glycerophosphodiester transporter) was up-regulated 2.6-fold in CTD-18 versus CTD-FL, whereas pho84 (phosphate transporter) was down-regulated 2.4-fold. Expression of pha1 (acid phosphatase) was not affected by the CTD-18 truncation, nor was pho7 (the DNA-binding transcription factor that drives expression of the three PHO genes). Among the genes of the iron homeostasis regulon (Labbé et al. 2007; Rustici et al. 2007), frp1 (ferric reductase), fip1 (iron permease), fio1 (iron transport oxidase), str3 (siderophore-iron transporter), sxx1 (sulfiredoxin), and shu1 (heme importer) were unaffected by the CTD-18 truncation; str1 (siderophore-iron transporter) was up-regulated by threefold in CTD-18 cells; sib1 and sib2 (ferriochrome biosynthesis genes) were down-regulated by 2.2-fold and 3.2-fold, respectively. The fep1 gene, which encodes the DNA-binding transcription factor that represses the iron regulon under iron-replete conditions (Pelletier et al. 2002), was down-regulated twofold by the CTD-18 truncation.

Transcriptomic impact of halving the Pro3, Ser5, and Pro6 content of the full-length CTD

Pro3 and Pro6 are essential for fission yeast viability, insofar as mutating them to alanine in every one of the 14
consensus heptads of \( rpb1-CTD-18 \) is lethal (Schwer and Shuman 2011). However, \( S. \) \textit{pombe} is viable when Pro3 or Pro6 is changed to alanine in every other heptad, including the rump, in the context of the full-length CTD (Fig. 2), signifying that reduced proline content is tolerated and that Pro3 and Pro6 need not be present in consecutive heptads (Sanchez et al. 2020). An initial survey of the effects of the \( P3\bar{ }P3A \) and \( P6\bar{ }P6A \) full-length CTD alleles on Pho1 acid phosphatase activity showed that pho1 expression was de-repressed in phosphate-replete \( P6\bar{ }P6A \) cells but unaffected in \( P3\bar{ }P3A \) cells (Sanchez et al. 2020). Systematic variation in the CTD Ser5 content of \( rpb1-CTD-18 \) by installing \( S5A \) mutations in half or more of the heptads (either alternating or en bloc) revealed that fission yeast were viable with as few as two Ser5-containing consensus heptads (plus the rump) and 12 SSA-containing heptads (Schwer et al. 2012). Ser5 status was found to be a tunable determinant of \( pho1 \) expression, that is, serial decrements in the number of consensus Ser5 heptads from seven to two elicited a progressive increase (via de-repression) in Pho1 activity in phosphate-replete cells (Schwer et al. 2015). These results underscored that there are functional distinctions between the Pro3 and Pro6 letters of the fission yeast CTD code, and that Ser5 and Pro6 content have parallel effects on phosphate homeostasis, but they do not provide a panoramic view of the role of the CTD prolines and Ser5 in gene expression.

To address how Ser5 content in the native full-length CTD affects gene expression globally, we constructed a viable strain in which Ser5 was replaced by alanine in every other heptad in \( rpb1-CTD-29 \), including the rump (Fig. 2). Poly(A)\(^+\) RNA-seq was performed on the \( P3\bar{ }P3A \), \( P6\bar{ }P6A \), and \( S5\bar{ }S5A \) strains (three biological replicates) and lists of genes expressed differentially (greater than

![Figure 2](image-url)
or equal to twofold up or down) compared to the full-length wild-type CTD control were compiled. Sets of 278, 287, and 277 protein-coding genes were down-regulated in P3•P3A, P6•P6A, and S5•SSA cells, respectively (Supplemental Table S2). Of these, 212 genes were coordinately down in all three chimeric CTD-Ala mutants (Fig. 2; Supplemental Table S3). A set of 50 genes was coordinately down-regulated in P6•P6A and S5•SSA cells, but not in the P3•P3A mutant (Fig. 2; Supplemental Table S3). Far fewer genes (n = 9) were down in P3•P3A and P6•P6A cells, but not in S5•SSA. Whereas 52 mRNAs were down uniquely in P3•P3A cells, fewer genes were uniquely down-regulated in P6•P6A (n = 16) and S5•SSA (n = 10).

Because the number of wild-type CTD heptads is reduced in the chimeric CTD-Ala mutants, we wanted to compare the sets of down genes in P3•P3A, P6•P6A, and S5•SSA cells to the 587 that were down-regulated by the rpb1-CTD-18 CTD truncation. The overlaps with CTD-18 were extensive, embracing 193/278 (69%), 174/287 (61%), and 167/277 (60%) of the down genes in P3•P3A, P6•P6A, and S5•SSA cells, respectively (Fig. 3). Indeed, 156/212 genes (74%) that were down-regulated in all three chimeric CTD-Ala mutants were also down in CTD-18 cells, including the sib1 and sib2 genes of the iron regulon (Fig. 4A; Supplemental Table S3). We surmise that more than half of the complement of down-regulated mRNAs seen when Pro3, Pro6, and Ser5 content was halved might be attributable to a decrement in wild-type CTD heptad number.

Instructive findings emerged from the sets of 59, 67, and 77 coding genes that were up-regulated in P3•P3A, P6•P6A, and S5•SSA cells, respectively, whereby 54 genes were coordinately up in P6•P6A, and S5•SSA cells, of which 39 were also up in the P3•P3A strain (Fig. 2; Supplemental Table S3). The up-regulated gene overlap with CTD-18 included 18/59 (31%) in P3•P3A, 17/67 (25%) in P6•P6A, and 18/77 (23%) in S5•SSA, respectively (Fig. 3). One-third (13/39) of the genes that were coordinately up-regulated in all three chimeric CTD-Ala mutants were also up in CTD-18 cells, among which was the iron regulon gene str1 (Fig. 4A). Thus, a lower fraction of the up-regulation of coding genes elicited by reduced Pro3, Pro6, and Ser5 content might be caused by a decrement in wild-type heptad number compared to the down-regulated genes discussed above.

The tgp1 mRNA of the PHO regulon, which was up-regulated 23-fold in S5•SSA and P6•P6A cells, was the most strongly affected UP transcript shared by these two mutants (Fig. 4A; see Fig. 4B for read densities across the tandem nc-tgp1–tgp1 lncRNA-mRNA locus). tgp1 was up-regulated to a much lesser extent in the P3•P3A strain (by 2.3-fold) and the CTD-18 truncation mutant (by 2.6-fold), from which we infer that reduction in CTD Ser5 and Pro6 content exerts a specific effect on tgp1 de-repression. Thirty-nine coding genes were coordinately up-regulated in S5•SSA and P6•P6A cells, but not in CTD-18 cells (Supplemental Table S3), 26 of which were also up-regulated in the P3•P3A strain. The phosphate homeostasis genes pho1 and pho84 were specifically up in S5•SSA (by 6.5-fold and 2.3-fold) and P6•P6A (by sevenfold and 2.4-fold) but were unaffected in P3•P3A (Fig. 4A). Read densities across the tandem prt2–pho84 and prt–pho1 lncRNA-mRNA loci are shown in Figure 4B; note that the y-axis scales are different for wild-type and mutant strains. Thus, all three genes of the PHO regulon were de-repressed by S5•SSA and P6•P6A.

Three of the iron regulon genes were up-regulated in P6•P6A cells but not in S5•SSA or P3•P3A (Fig. 4A). These were str3 (up 14-fold), flo1 (3.4-fold), and fis1 (2.9-fold). strx1 was up approximately twofold in all three chimeric CTD-Ala mutants (though not in CTD-18). str2 was up 2.2- to 2.6-fold in P3•P3A and S5•SSA cells but not in P6•P6A mutant. fep1 was not dysregulated in any of the chimeric CTD-Ala mutants.

De-repression of pho1 by S5•SSA and P6•P6A subunits and Rhn1

Expression of pho1, pho84, and tgp1 is actively repressed during growth in phosphate-rich medium by the transcription in cis of a long noncoding (lnc) RNA from the respective 5′ flanking genes prt, prn2, and nc-tgp1 (Shuman 2020). The current model for the repressive arm of fission yeast
phosphate homeostasis is that transcription of the upstream lncRNA interferes with expression of the downstream mRNA genes by displacing the activating transcription factor Pho7 from its binding site(s) in the mRNA promoters that overlap the lncRNA transcription units (Shuman 2020). This model is supported by findings that: (i) mutations of CPF (Cleavage and Polyadenylation Factor) subunits and Rhn1 result in hyper-repression of pho1 expression. Acid phosphatase activity (a gauge of Pho1 enzyme level that correlates with phosphate homeostasis genes) was quantified by incubating suspensions of serial dilutions of the phosphate-replete cells for 5 min with p-nitrophenyl phosphate and assaying colorimetrically the formation of p-nitrophenol. In agreement with the RNA-seq data, the de-repressive effect of S5•S5A and P6•P6A on pho1 depends similarly on CPF and Rhn1, we mated marked S5•S5A and P6•P6A strains with differently marked knock-out strains lacking the Dis2, Ctf1, Ppn1, or Swd22 subunits of the CPF complex (Vanoosthuyse et al. 2014), a strain with a catalytically dead (C13S) version of the Ssu72 protein that lacks the transcription termination factor Rhn1. Viable haploid double-mutant progeny grew fairly well on YES agar at 30°C, except for the P6•P6A ssu72-C13S strain, which was slow growing, as reflected in colony size (Fig. 5A). The S5•S5A and P6•P6A single-mutants, a wild-type CTD control, and the S5•S5A CPF/Rhn1 and P6•P6A CPF/Rhn1 double-mutants growing exponentially at 30°C in liquid culture under phosphate-replete conditions were assayed for Pho1 expression. Acid phosphatase activity (a gauge of Pho1 enzyme level that correlates with pho1 mRNA levels) was quantified by incubating suspensions of serial dilutions of the phosphate-replete cells for 5 min with p-nitrophenylphosphate and assaying colorimetrically the formation of p-nitrophenol. In agreement with the RNA-seq data, the basa...
suggest that de-repression of the PHO regulon by diminishing CTD Ser5 and Pro6 content is a consequence of CPF/Rhn1-dependent precocious termination of interfering lncRNA transcription.

Transcriptomic impact of pan-T4A mutation of the full-length CTD

A full-length rpb1-CTD-T4A allele in which every Thr4 was replaced by alanine (Fig. 6) results in hyper-repression of pho1 and synthetic lethality with ppn1Δ and swd22Δ (Sanchez et al. 2018b), prompting the conclusion that the Thr4-PO4 (or the Thr4 hydroxyl) plays a positive role in 3′-processing/termination. Here we performed poly(A)⁺ RNA-seq analysis of the rpb1-CTD-T4A strain (three biological replicates) and thereby identified: (i) 100 protein-coding RNAs and 63 noncoding RNAs that were up-regulated by T4A compared to full-length wild-type rpb1-CTD cells; and (ii) 277 coding RNAs and 153 noncoding RNAs that were down-regulated (see Supplemental Table S4 for list of affected protein-coding genes). Because the number of T4A dysregulated genes in the full-length CTD was much higher than what was noted previously for the truncated CTD-18-T4A strain, we checked the overlap of genes affected in full-length CTD-T4A and truncated wild-type CTD-18 cells, which showed that 26/100 coding RNAs up-regulated by T4A and 153/277 coding RNA down-regulated by T4A were similarly dysregulated up and down in CTD-18 cells (Fig. 7), hinting that one-quarter to one-half of the observed effects in T4A might be related to a reduction in the number of wild-type repeats. With respect to phosphate gene regulation, pho1 was among the set of genes specifically down (by 8.6-fold) in T4A cells but not in CTD-18 (Fig. 4A). pho84 was more repressed in the

FIGURE 5. De-repression of pho1 by S5*SSA and P6*P6A depends on CPF subunits and Rhn1. (A) Strains with the specified genotypes were spot-tested for growth on YES at 30°C. (B) S. pombe strains bearing the indicated CPF/Rhn1 loss of function alleles in context of the specified rpb1-CTD backgrounds were grown in liquid culture at 30°C and assayed for acid phosphatase activity. Each datum in the bar graph is the average of assays using cells from at least three independent cultures ± SEM.

FIGURE 6. Transcriptome profiles of Y1F, S2A, and T4A CTD mutants. (Left panel) The amino acid sequences of the CTD heptad arrays of Y1F, S2A, and T4A full-length rpb1-CTD alleles are shown, with mutated residues indicated in red (Y1F), blue (S2A), and green (T4A) font. (Right panel) Overlaps of protein-coding genes that were down-regulated (DOWN) or up-regulated (UP) in the indicated CTD mutants compared to CTD-FL.
T4A strain (ninefold) compared to CTD-18 (2.4-fold) (Fig. 4A). Read densities across the tandem prt2-pho84 and prt-pho1 IncRNA-mRNA loci are shown in Figure 4B and affirm that both mRNAs were down-regulated in T4A cells versus wild-type. (Note that the y-axis scales in Figure 4B are different for the wild-type and T4A strains.) The genes specifically up-regulated in T4A cells (but not in CTD-18) included several constituents of the Fep1-repressible iron homeostasis regulon: fio1 (7.6-fold), srx1 (sixfold), fip1 (4.6-fold), and isu1 (2.2-fold). str1 was up-regulated to a greater degree in T4A cells (15-fold) than in the CTD-18 strain (threefold) (Fig. 4A). fep1 expression was unaffected by T4A. We surmise that the Thr4 letter of the CTD code is needed for Fep1 repression of a subset of the known iron responsive genes. The effect of T4A in de-repressing the iron regulon genes overlapped with that of P6•P6A (Fig. 4A).

Transcriptome profiles of pan-Y1F and pan-S2A CTD mutations

Full-length rpb1-CTD alleles in which every Tyr1 was replaced by phenylalanine or every Ser2 was replaced by alanine (Fig. 6) are synthetically lethal with pnp1Δ and swd22Δ (Sanchez et al. 2018b), implicating the Tyr1–PO4 and Ser2–PO4 marks—or the Tyr1 and Ser2 hydroxyls—in 3′-processing/termination. Here we performed RNA-seq analysis of the full-length rpb1-CTD-Y1F strain, which identified: 119 protein-coding RNAs and 65 noncoding RNAs that were up-regulated by Y1F compared to full-length wild-type rpb1-CTD cells; and 125 coding RNAs and 64 noncoding RNAs that were down-regulated (see Supplemental Table S4 for list of affected protein-coding genes). 33/119 (28%) of the coding RNAs up-regulated by Y1F and 73/125 (58%) of the coding RNAs down-regulated by Y1F were similarly dysregulated up and down in CTD-18 cells (Fig. 7). RNA-seq of the full-length rpb1-CTD-S2A strain highlighted: 111 protein-coding RNAs and 65 noncoding RNAs that were up-regulated by CTD-S2A; and 141 coding RNAs and 52 noncoding RNAs that were down-regulated (Supplemental Table S4). 25/111 (23%) of the coding RNAs up-regulated by S2A and 68/141 (48%) of the coding RNA down-regulated by S2A were dysregulated up and down in CTD-18 cells. These overlaps with CTD-18 echo the findings for the CTD-T4A strain; to wit, one-quarter (UP genes) to one-half (DOWN genes) of the observed coding transcriptome effects in Y1F and S2A cells might be related to a reduction in the number of wild-type CTD repeats. As was the case for the chimeric P3•P3A, S5•S5A, and P6•P6A CTD mutants, a lower fraction of the up-regulated coding genes in Y1F, S2A, and T4A cells overlapped with CTD-18 compared to the down-regulated genes.

The concordance of the Y1F and S2A coding transcriptome profiles (and that of T4A) is shown in the Venn diagrams in Figure 6. A set of 67 mRNAs were coordinately down-regulated in Y1F and S2A cells (P < 2.5 × 10−77), of which 60 were also down in T4A cells. Y1F and S2A cells shared an ensemble of 38 up-regulated mRNAs (P < 2.9 × 10−36), 16 of which were also up in T4A cells. Gene lists are provided in Supplemental Table S5. Anent the PHO genes, note that the Y1F and S2A CTD alleles did not elicit the strong hyper-repression of pho1 that was seen in the T4A mutant (Fig. 4A). pho84 expression was down in Y1F cells to the same extent as seen in CTD-18 cells and tgo1 RNA was up in S2A cells to the same degree as in the CTD-18 mutant (Fig. 4A).

RNA-seq of fission yeast with a pan-(S2A + S7A) mutation of the full-length CTD

We had previously profiled the transcriptome of a fission yeast strain in which Ser2 and Ser7 of each consensus heptad were both mutated to alanine in the context of CTD-18 (Schwer et al. 2014). The CTD-18-(S2A + S7A) mutant exhibited up-regulation of several components of the iron regulon, recapitulating what was observed in the CTD-18-(S2A) strain. CTD-18-(S2A + S7A) cells also displayed elevated acid phosphatase activity vis-à-vis wild-type CTD-18 and CTD-18-(S2A) (Schwer et al. 2015). Here we constructed a full-length CTD-(S2A + S7A) mutant in which all Ser2 and Ser7 positions were replaced by alanine (Fig. 8). RNA-seq identified: (i) 95 protein-coding RNAs and
141 noncoding RNAs that were up-regulated in CTD-(S2A + S7A) cells compared to wild-type rpb1-CTD-FL cells; and (ii) 347 coding RNAs and 54 noncoding RNAs that were down-regulated (Supplemental Table S4). 17/95 (18%) of the coding RNAs up-regulated by S2A + S7A and 43/347 (12%) of the coding RNA down-regulated by S2A + S7A were dysregulated up and down in CTD-18 cells (Fig. 8).

Whereas the extent of overlap of full-length S2A + S7A up-regulated genes with CTD-18 was in keeping with those of the other full-length CTD mutants described above, the percent overlap of the S2A + S7A down-regulated genes with CTD-18 was much less. We surmise that simultaneous loss of the Ser2–PO4 and Ser7–PO4 marks (or the respective serine hydroxyls) exerts a singular impact in reducing the expression of ~300 mRNAs independent of a reduction in the number of intact CTD heptads.

The PHO genes tgp1 and pho1 were up-regulated by fivefold and threefold, respectively, in S2A + S7A cells (Fig. 4A). The iron regulon was globally up-regulated in S2A + S7A cells as follows: str3 (5-fold), shu1 (51-fold), fio1 (sixfold), fip1 (fivefold), frp1 (fivefold), and srx1 (threefold). The effect of S2A + S7A on the iron genes was similar to that of P6•P6A, except that frp1 and shu1 were uniquely up-regulated in the S2A + S7A strain.

FIGURE 8. Transcriptome profile of a CTD-(S2A + S7A) mutant. (Left panel) The amino acid sequence of the S2A + S7A CTD. (Right panel) Overlaps of dysregulated genes of CTD-(S2A + S7A) and CTD-18 are shown.

Conclusions
The Pol2 CTD, composed of a consensus Y^S^P^T^S^P^S^S heptapeptide repeat array, is critical for orchestrating transcriptional events. Our RNA-seq analysis of fission yeast rpb1-CTD mutants unveils three key contributions of the CTD primary structure to the gene expression landscape. First, whereas truncation of the CTD to 18 repeats had no effect on fission yeast vegetative growth (Schneider et al. 2010), 25% of protein-coding genes were dysregulated by greater than or equal to twofold, highlighting the genome-wide impact of CTD length on the fission yeast transcriptome. In contrast, microarray analysis of mRNA expression in a budding yeast strain in which the CTD was truncated to 13 repeats showed that only one protein-coding gene was up-regulated by greater than or equal to twofold versus wild-type and fewer than 10 protein-coding genes were down-regulated by greater than or equal to twofold (Aristizabal et al. 2013). Second, 25%–50% of the fission yeast protein-coding genes that were dysregulated by full-length CTDs with mutated heptad repeats overlapped with those that were affected by CTD truncation. We suggest that this reflects a reduction in the number of wild-type consensus repeats within the full-length CTD array. Third, our data illuminate shared patterns of gene dysregulation by specific CTD mutants, which fortify our suggestions that Ser5–Pro6 and Tyr1–Ser2–Thr4 comprise distinct "words" in the CTD code that exert different impacts on gene expression programs.

The fission yeast phosphate homeostasis regulon is especially sensitive to CTD heptad missense mutations. The PHO genes are repressed in phosphate-replete cells by transcriptional interference via upstream flanking lncRNAs. The efficiency of precocious lncRNA 3′-processing and transcription termination is a major factor in tuning expression of the PHO mRNAs in phosphate-replete cells. Previous studies had established that the de-repressive effects of CTD phospho-site mutation S7A on pho1 expression were dependent on CPF subunits and termination factor Rnh1. Here we show that the chimeric CTD mutants S5•S5A and P6•P6A also exert their pho1 de-repression via CPF and Rhn1. Collectively, findings here and previously (Sanchez et al. 2018b) implicate the Tyr1–Ser2–Thr4 word as a positive effector of 3′-processing/termination and the Ser5–Pro6–Ser7 triplet as a negative influence on precocious 3′-processing/termination.

MATERIALS AND METHODS

CTD mutations
To introduce mutations of the full-length CTD in-frame to amino acid 1550 of S. pombe Rpb1, synthetic DNA segments (GenScript) were inserted into a rpb1 CTD integration cassette marked with natMX (Schneider et al. 2010). The plasmids
harboring the mutated integration cassettes were linearized and transformed into a diploid S. pombe strain (ade6–m216 ade6–m210 leu1–32 leu1–32 ura4–D18 ura4–D18 his3–D1 his3–D1). Nourseothricin-resistant transformants were selected and correct integrations at one of the rpb1::natMX allele were confirmed by Southern blotting. A segment of the rpb1::natMX allele was amplified by PCR and sequenced to verify that the desired mutations were present in all CTD repeats. The heterozygous rpb1+/rpb1::natMX diploids were then sporulated, subjected to random spore analysis and nourseothricin-resistant haploids harboring full-length CTD alleles were isolated.

Mutational effects on fission yeast growth

Cultures of S. pombe strains were grown in YES liquid medium until A600 reached 0.6–0.8. The cultures were adjusted to A600 of 0.1 and 3 µL aliquots of serial fivefold dilutions were spotted on YES agar. The plates were photographed after incubation for 2 d at 34°C, 2.5 d at 30°C and 37°C, 4 d at 25°C, 6 d at 20°C.

Double mutants

Standard genetic methods were used to generate haploid strains harboring mutations/deletions in two differently marked genes. In brief, pairs of haploids with missense or null mutations were mixed on malt agar to allow mating and sporulation and then the mixture was subjected to random spore analysis. Spores (~1500) were plated on YES agar and on media selective for marked mutant alleles; the plates were incubated at 30°C for up to 5 d to allow slow growing progeny to germinate and form colonies. At least 500 viable progeny were screened by replica-plating for the presence of the second marker gene, or by sequentially replica-plating from YES to selective media. Growth phenotypes of viable double-mutants were assessed in parallel with the individual mutants and wild-type cells at different temperatures (20°C to 37°C) as described above.

Acid phosphatase activity

Cells were grown at 30°C in YES liquid medium. Aliquots of exponentially growing cultures were harvested with water, and resuspended in water. To quantify acid phosphatase activity, reaction mixtures (200 µL) containing 100 mM sodium acetate (pH 4.2), 10 mM p-nitrophenyl phosphate, and cells (ranging from 0.01 to 0.1 A600 units) were incubated for 5 min at 30°C. The reactions were quenched by addition of 1 mL of 1 M sodium carbonate, the cells were removed by centrifugation, and the absorbance of the supernatant at 410 nm was measured. Acid phosphatase activity is expressed as the ratio of A410 (p-nitropheno l production) to A600 (cells). The data shown in graphs are averages (±SEM) of at least three assays using cells from three independent cultures.

Transcriptome profiling by RNA-seq

RNA was isolated from S. pombe cells grown in liquid YES medium at 30°C to an A600 of 0.5 to 0.6. Cells were harvested by centrifugation and total RNA was extracted via the hot phenol method. The integrity of total RNA was gauged with an Agilent Technologies 2100 Bioanalyzer. The Illumina TruSeq stranded mRNA sample preparation kit was used to purify poly(A)+ RNA and to carry out the subsequent steps of poly(A)+ RNA fragmentation, strand-specific CDNA synthesis, indexation, and amplification. Indexed libraries were normalized and pooled for paired-end sequencing performed by using an Illumina HiSeq 4000 system. FASTQ files bearing paired-end reads of length 51 bases were mapped to the S. pombe genome using HISAT2-2.1.0 with default parameters (Kim et al. 2015). The resulting SAM files were converted to BAM files using Samtools (Li et al. 2009). Count files for individual replicates were generated with HTSeq-0.10.0 (Anders et al. 2015) using exon annotations from Pombase (GFF annotations, genome-version ASM294v2; source "ensembl"). The data sets (comprising 24 to 40 million paired reads per sample) had an overall alignment rate of 93%–98% to genomic loci. RPKM analysis and pairwise correlations were performed as described previously (Schwer et al. 2014). Read densities for individual genes were highly reproducible between biological replicates, with Pearson coefficients of 0.973 to 0.989. Differential gene expression and fold change analysis was performed in DESeq2 (Love et al. 2014). Cut-off for further evaluation was set for genes that had an adjusted P-value (Benjamini-Hochberg corrected) of ≤0.05, had an average normalized count of ≥100 in the wild-type or mutant strains, and were up or down by at least twofold in comparison to wild-type.

DATA DEPOSITION

The RNA-seq data in this publication have been deposited in NCBI’s Gene Expression Omnibus and are accessible through GEO Series accession number GSE160852 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE160852).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH) grants R35-GM126945 (S.S.) and R01-GM134021 (B.S.).

Received January 11, 2021; accepted February 11, 2021.

REFERENCES

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169. doi:10.1093/bioinformatics/btu638
Aristizabal MJ, Negri GL, Benschop JJ, Holstege FC, Krogan NJ, Kobot MS. 2013. High-throughput genetic and gene expression analysis of the RNAPII-CTD reveals unexpected connections to SRB10/CDK8. PLoS Genet 9: e1003758. doi:10.1371/journal.pgen.1003758
Carter-O’Connell I, Peel MT, Wykoff DD, O’Shea EK. 2012. Genome-wide characterization of the phosphate starvation response in Schizosaccharomyces pombe. BMC Genomics 13: 697. doi:10.1186/1471-2164-13-697
Chatterjee D, Sanchez AM, Goldgur Y, Shuman S, Schwer B. 2016. Transcription of lncRNA prt, clusteredprtRNA sites for Mmi1 binding, and RNA polymerase II CTD phospho-sites govern the repression of pho1 gene expression under phosphate-replete conditions in fission yeast. RNA 22: 1011–1025. doi:10.1261/ma.056515.116

Corden JL. 2013. RNA polymerase II C-terminal domain: tethering transcription to transcript and template. Chem Rev 113: 8423–8455. doi:10.1021/cr400158h

Eick D, Geyer M. 2013. The RNA polymerase II carboxy-terminal domain (CTD) code. Chem Rev 113: 8456–8490. doi:10.1021/cr400071f

Jeronimo C, Bataille AR, Robert F. 2013. The writers, readers, and functions of the RNA polymerase II C-terminal domain code. Chem Rev 113: 8491–8522. doi:10.1021/cr4001397

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12: 357. doi:10.1038/nmeth.3317

Labbé S, Pelletier B, Mercier A. 2007. Iron homeostasis in the fission yeast Schizosaccharomyces pombe. Biometals 20: 523–537. doi:10.1007/s10534-006-9056-5

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079. doi:10.1093/bioinformatics/btp352

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550. doi:10.1186/s13059-014-0550-8

Pelletier B, Beaudoin J, Mukai Y, Labbé S. 2002. Fep1, an iron sensor regulating iron transporter gene expression in Schizosaccharomyces pombe. J Biol Chem 277: 22950–22958. doi:10.1074/jbc.M202682200

Rustici G, van Bakel H, Lackner DH, Holstege FC, Wijmenga C, Bähler J, Brazma A. 2007. Global transcriptional responses of fission and budding yeast to changes in copper and iron levels: a comparative study. Genome Biol 8: R73. doi:10.1186/gb-2007-8-5-r73

Sanchez AM, Shuman S, Schwer B. 2018a. Poly(A) site choice and Pol2 CTD Serine-5 status govern lncRNA control of phosphate-responsive tgp1 gene expression in fission yeast. RNA 24: 237–250. doi:10.1261/ma.063966.117

Sanchez AM, Shuman S, Schwer B. 2018b. RNA polymerase II CTD interactome with 3′ processing and termination factors in fission yeast and its impact on phosphate homeostasis. Proc Natl Acad Sci 115: E10652–E10661. doi:10.1073/pnas.1810711115

Sanchez AM, Garg A, Shuman S, Schwer B. 2019. Inositol pyrophosphates impact phosphate homeostasis via modulation of RNA 3′ processing and transcription termination. Nucleic Acids Res 47: 8452–8469. doi:10.1093/nar/gkz567

Sanchez AM, Garg A, Shuman S, Schwer B. 2020. Genetic interactions and transcriptomics implicate fission yeast CTD prolyl isomerase Pin1 as an agent of RNA 3′ processing and transcription termination that functions via its effects on CTD phosphatase Ssu72. Nucleic Acids Res 48: 4811–4826. doi:10.1093/nar/gkaa212

Schneider S, Pei Y, Shuman S, Schwer B. 2010. Separable functions of the fission yeast Spt5 CTD in capping enzyme binding and transcription elongation overlap with those of the RNA polymerase II CTD. Mol Cell Biol 30: 2353–2364. doi:10.1128/MCB.00116-10

Schwer B, Shuman S. 2011. Deciphering the RNA polymerase II CTD code in fission yeast. Mol Cell 43: 311–318. doi:10.1016/j.molcel.2011.05.024

Schwer B, Sanchez AM, Shuman S. 2012. Punctuation and syntax of the RNA polymerase II CTD code in fission yeast. Proc Natl Acad Sci 109: 18024–18029. doi:10.1073/pnas.1208995109

Schwer B, Bitton DA, Sanchez AM, Bähler J, Shuman S. 2014. Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast. Proc Natl Acad Sci 111: 4185–4190. doi:10.1073/pnas.1321842111

Schwer B, Sanchez AM, Shuman S. 2015. RNA polymerase II CTD phospho-sites Ser5 and Ser7 govern phosphate homeostasis in fission yeast. RNA 21: 1770–1780. doi:10.1261/ma.052555.115

Schwer B, Sanchez AM, Shuman S. 2020. Inactivation of fission yeast Erh1 de-represses pho1 expression: evidence that Erh1 is a negative regulator of prt IncRNA termination. RNA 26: 1334–1344. doi:10.1261/ma.076463.120

Shuman S. 2020. Transcriptional interference at tandem IncRNA and protein-coding genes: an emerging theme in regulation of cellular nutrient homeostasis. Nucleic Acids Res 48: 8243–8254. doi:10.1093/nar/gkaa630

Vanoosthuyse V, Legros P, van der Sar SJ, Yvert G, Toda K, Le Bihan T, Watanabe Y, Hardwick K, Bernard P. 2014. CPF-associated phosphatase activity opposes condensin-mediated chromosome condensation. PLoS Genet 10: e1004415. doi:10.1371/journal.pgen.1004415