

Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B

Zuyao Ni,^{1†} Jonathan B. Olsen,^{1,2†} Xinghua Guo,¹ Guoqing Zhong,¹ Eric Dongliang Ruan,¹ Edyta Marcon,¹ Peter Young,¹ Hongbo Guo,¹ Joyce Li,¹ Jason Moffat,^{1,2} Andrew Emili^{1,2} and Jack F. Greenblatt^{1,2,*}

¹Banting and Best Department of Medical Research; ²Department of Molecular Genetics; Donnelly Centre; University of Toronto; Toronto, ON Canada

[†]These authors contributed equally to this work.

Key words: RPRD1A, RPRD1B, CID, CTD, RNA polymerase II

Abbreviations: RNAP II, RNA polymerase II; CTD, carboxy terminal domain; CID, CTD interaction domain; S2P, phosphorylated serine 2 in the RNAP II CTD; S5P, phosphorylated serine 5 in the RNAP II CTD; S7P, phosphorylated serine 7 in the RNAP II CTD; RNAP II (HIO), hyper-phosphorylated RNAP II; RNAP II (IIA), hypo-phosphorylated RNAP II; ChIP, chromatin immunoprecipitation; shRNA, short hairpin RNA; IP-WB, immunoprecipitation and western blot; qPCR, quantitative real-time PCR; ORF, open reading frame; AP-MS, affinity purification and mass spectrometry; RNAP I, RNA polymerase I; RNAP III, RNA polymerase III

RNA polymerase II (RNAP II) C-terminal domain (CTD) phosphorylation is important for various transcription-related processes. Here, we identify by affinity purification and mass spectrometry three previously uncharacterized human CTD-interaction domain (CID)-containing proteins, RPRD1A, RPRD1B and RPRD2, which co-purify with RNAP II and three other RNAP II-associated proteins, RPAP2, GRINL1A and RECQL5, but not with the Mediator complex. RPRD1A and RPRD1B can accompany RNAP II from promoter regions to 3'-untranslated regions during transcription *in vivo*, predominantly interact with phosphorylated RNAP II, and can reduce CTD S5- and S7-phosphorylated RNAP II at target gene promoters. Thus, the RPRD proteins are likely to have multiple important roles in transcription.

In eukaryotes, the C-terminal domain (CTD) of the largest subunit, POLR2A, of RNA polymerase II (RNAP II) contains multiple, sometimes degenerate repeats, 26–27 in yeast and 52 in human, of the heptapeptide sequence $Y_1S_2P_3T_4S_5P_6S_7$.¹ During transcription, serine residues on positions 2, 5 and 7 are subjected to dynamic phosphorylation and dephosphorylation by specific kinases and phosphatases.² The CDK7 (Kin28 in yeast) kinase, a subunit of the general transcription factor TFIIF, is known to be responsible for the phosphorylation of S5 (S5P) near gene promoters,³ although lower levels of S5P persist on RNAP II further downstream.⁴ As RNAP II moves away from promoters, S5 is dephosphorylated mainly by Rtr1 in yeast.⁵ It is not known whether the human Rtr1 homolog, RPAP2, which also associates with RNAP II, performs a similar role.⁶ Another yeast phosphatase, Ssu72, a subunit of the cleavage and polyadenylation factor CPF, reverses S5P further downstream and at 3'-ends of genes.⁷ With the decline of S5P in downstream transcribed regions, S2 phosphorylation (S2P) is catalyzed by the CDK9 subunit of the positive elongation complex p-TEFb, as well as CDK12 and CDK13, and gradually increases as RNAP II travels toward the 3'-ends of genes.^{8–12} S2P is reversed by the phosphatase

CTDPI (Fcp1 in yeast).¹³ Like S5P, S7 phosphorylation (S7P) is catalyzed in promoter-proximal regions by TFIIF, as well as by CDK12.^{4,10,14,15} S7P also occurs in promoter-distal regions through the action of Bur1 in yeast, although the enzymes responsible for S7P dephosphorylation remain unknown.¹⁶

Site-specific CTD phosphorylation provides a docking platform for many proteins that regulate various transcription-related events.^{2,3,17} For example, S5P creates a binding site for the capping enzyme Cgt, which modifies the 5'-ends of nascent mRNA.^{18–20} On the other hand, S2P specifically associates with the cleavage factor Pcf11 and the “torpedo” complex subunit Rtt103, resulting in transcriptional termination.^{21–23} Some proteins, such as the helicase RECQL5, can bind to both S2P and S5P CTDs, leading to transcriptional repression.^{24,25} S7P in conjunction with S2P interacts with Integrator, a 3'-end processing complex for snRNA.²⁶ Some CTD binders (e.g., Cgt1, Pcf11 and Rtt103) contain a CTD interaction domain, or CID. Based on the complexity of CTD modification, the variety of binding partners and different possible outcomes, CTD phosphorylation and other modifications were proposed as a possible “CTD code.”⁸ This “code” specifies factor recruitment, which in turn controls

*Correspondence to: Jack F. Greenblatt; Email: jack.greenblatt@utoronto.ca

Submitted: 07/20/11; Revised: 08/17/11; Accepted: 08/18/11

DOI: 10.4161/trns.2.5.17803

Identified proteins	Tagged proteins							
	POLR2D	POLR2E	RPRD1A	RPRD1B	GRINL1A	MED7	MED18	
RPRD1A	140	173	1620	289	261	0	0	RPRD
RPRD1B	96	165	170	642	266	0	0	
RPRD2	30	46	11	71	62	0	0	
GRINL1A	29	48	34	50	2793	3	0	RNAPII
RECQL5	15	0	26	44	23	0	0	
RPAP2	20	12	43	41	52	3	0	
POLR2A	132	176	166	157	270	59	16	RNAPII
POLR2B	162	182	175	163	359	61	17	
POLR2C	171	268	170	162	435	74	19	
POLR2D	358	103	92	119	87	0	0	
POLR2E*	180	239	128	142	229	32	15	
POLR2G	625	158	128	140	194	12	6	
POLR2H*	185	580	177	188	238	22	11	
POLR2I	175	212	98	155	206	36	0	
POLR2J	111	85	73	147	223	0	0	
POLR2L*	49	95	64	88	108	0	5	
POLR1A	0	208	0	0	0	0	0	RNAPI
POLR1B	0	132	0	0	0	0	0	
POLR1C **	0	250	0	0	0	0	0	
POLR1D **	0	166	0	0	0	0	0	RNAPIII
POLR3A	0	92	0	0	0	0	0	
POLR3D	0	69	0	0	0	0	0	
CDK19	0	0	0	0	0	26	1	Mediator
MED1	4	6	0	0	12	151	22	
MED4	6	16	0	0	31	262	45	
MED6	0	0	0	0	17	61	32	
MED7	0	3	0	0	4	161	0	
MED8	8	12	0	0	22	68	48	
MED9	6	12	0	0	9	144	14	
MED10	10	11	0	0	33	146	14	
MED11	2	0	0	0	0	63	29	
MED12	0	0	0	0	0	40	5	
MED12L	0	0	0	0	0	9	0	
MED13	0	0	0	0	0	21	2	
MED13L	0	0	0	0	0	30	3	
MED14	2	7	0	0	12	160	22	
MED15	0	0	0	0	0	7	1	
MED16	0	0	0	0	0	3	0	
MED17	6	12	0	0	4	98	38	
MED18	0	0	0	0	0	25	230	
MED19	7	9	0	0	0	155	11	
MED20	5	0	0	0	0	84	55	
MED21	19	29	0	0	24	120	46	
MED22	4	0	0	0	0	37	0	
MED26	0	4	0	0	10	68	16	
MED27	4	10	0	0	22	103	52	
MED28	4	4	0	0	0	86	21	
MED29	0	0	0	0	0	20	0	
MED30	5	10	0	0	0	54	11	
MED31	6	0	0	0	9	137	26	

Figure 1. RPRD proteins interact with RNAP II and three known RNAP II-associated proteins. Summary of AP-MS interaction data for indicated protein baits. Total matching spectral evidence (spectral counts normalized by molecular weight) recorded for co-purified interacting “prey” proteins are shown as an average from a minimum of two independent biological replicates. Background contaminants present after purification of tagged GFP control proteins were subtracted. *, Shared subunits of RNAP I, II and III; **, Shared subunits of RNAP I and III. Sequest V2.7 was used for the database search with the MS data.

transcription and other processes (e.g., transcription-coupled histone modification).²⁷⁻²⁹ Therefore, identifying RNAP II CTD binding factors and molecular dependencies is an important step in better understanding the mechanisms that control transcription and related processes.

To identify novel RNAP II-interacting proteins, we applied a lentivirus-based affinity purification and tandem mass spectrometry

(AP-MS) approach in HEK293 cells stably expressing an epitope tagged RNAP II subunit (e.g., POLR2D or POLR2E).³⁰ As expected, we recovered most the subunits of RNAP II, as well as shared RNAP I and RNAP III subunits, along with three previously characterized RNAP II interactors, GRINL1A (Gdown1), RECQL5 and RPAP2 (Fig. 1). Unexpectedly, we identified three novel RNAP II-interacting proteins, RPRD1A, RPRD1B and

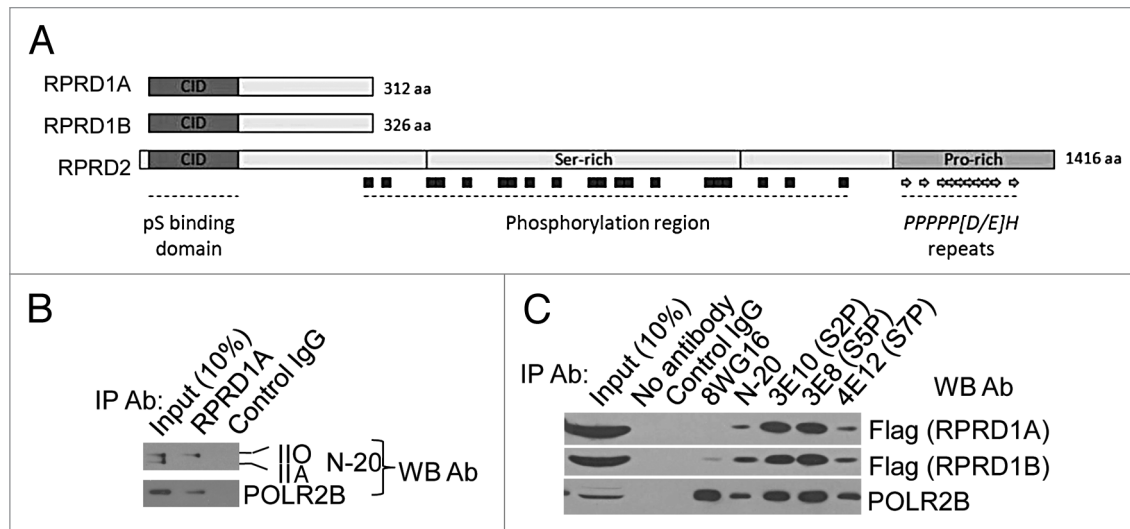


Figure 2. RPRD proteins predominantly interact with phosphorylated RNAP II. (A) RPRD protein architecture. (B and C) Precipitation with the indicated antibodies was performed with soluble protein lysates from HEK293 cells (B) or HEK293 cells stably expressing tagged RPRD1A or RPRD1B (C), followed by protein gel blotting using the indicated probes. Control IgG: TBL1 antibody.

RPRD2 (referred to, generally, as RPRD proteins), with significant supporting mass spectral evidence (see Fig. 1). The RPRD proteins are evolutionarily conserved, represent 3 of the 8 CID-containing proteins annotated in humans (Fig. S1) and appear to be ubiquitously expressed in most or all human cell types and tissues (Fig. S2). They are most closely related to the *S. cerevisiae* protein Rtt103, whose CID has been structurally characterized in complex with an S2P CTD peptide.³¹ RPRD1A and RPRD1B share similar amino acid sequences and domain architectures, while RPRD2 is much larger and contains compositionally biased serine- and proline-rich regions (Fig. 2A), the latter region becoming hyperphosphorylated during M-phase.³² The proline-rich region contains several repeats of the consensus P[PPPP[D/E]H] that is recognized and bound by the WW domains of the transcription factor and tumor suppressor WWOX.³³ Previous studies showed that RPRD1A, also called P15RS, regulates cell cycle genes and attenuates WNT signaling, although the mechanism was unclear.^{34,35} The interaction of RPRD1A with RNAP II has not been reported previously, and the RPRD1B and RPRD2 proteins are functionally uncharacterized.

To confirm these interactions, we applied AP-MS using RPRD1A and RPRD1B as bait proteins. Proteins identified as co-purifying with RPRD1A and RPRD1B included almost all the core RNAP II subunits as well as GRINL1A, which creates a Mediator requirement for transcriptional activation in vitro, the helicase RECQL5, which inhibits transcription by RNAP II in vitro, and the putative CTD-S5 phosphatase RPAP2 (Fig. 1).^{5,6,24,36} RPRD1A and RPRD1B likewise co-purified with each other as well as with RPRD2 (Fig. 1). Conversely, RNAP I, RNAP III, and the Mediator complex, which binds RNAP II with an unphosphorylated CTD,³⁷ were not detected. RPRD proteins were also detected in AP-MS experiments by reciprocal tagging and purification of GRINL1A, but not when Mediator subunits were used as baits (Fig. 1). Results from immunoprecipitation

and protein gel blotting (IP-WB) experiments confirmed interactions of RPRD1A, RPRD1B and GRINL1A with RPAP2 (Fig. S3). Collectively, these data suggest that RPRD proteins associate simultaneously with RNAP II, GRINL1A, RECQL5 and RPAP2, but not with Mediator.

Given the Mediator-exclusive association with RNAP II and the presence of CIDs in the RPRD proteins (Fig. 1), we reasoned that the RPRD proteins may recognize phosphorylated forms of RNAP II. Indeed, IP-WB performed in HEK293 cells probing with anti-RNAP II (N-20) antibody that recognizes both hyper- (IIO) and hypo-phosphorylated RNAP II (IIA), showed that antibody against RPRD1A, whose specificity was confirmed using lentivirus-encoded shRNA to knock down RPRD1A (Fig. S4A), precipitated RNAP II (IIO) (Fig. 2B). CID proteins usually tend to bind to CTD peptides phosphorylated in a specific way.³⁸ To investigate possible CTD phosphoisoform binding specificity of RPRD proteins, we performed an IP-WB experiment after precipitating with monoclonal antibodies specific for S2P (3E10), S5P (3E8) and S7P (4E12).³⁹ Unexpectedly, tagged RPRD1A and RPRD1B both co-immunoprecipitated with all three phosphorylated CTD variants (Fig. 2C). Conversely, the monoclonal antibody 8WG16, which recognizes unmodified CTD, precipitated substantial levels of POLR2B, but very low levels of RPRD1B and no detectable levels of RPRD1A (Fig. 2C). Consistent with this, an antibody against endogenous RPRD1A co-immunoprecipitated all three CTD phosphoisoforms (Fig. S4B). Taken together, these results demonstrate that RPRD1A and RPRD1B predominantly bind to RNAP II with a phosphorylated CTD. Because the CTD may be phosphorylated simultaneously on S2, S5 and S7, we cannot yet be certain from these experiments which CTD repeat phosphoisoform(s) is recognized directly by RPRD1A and RPRD1B.

Next, we sought to examine possible effects of RPRD proteins on RNAP II phosphorylation. Although we found using protein

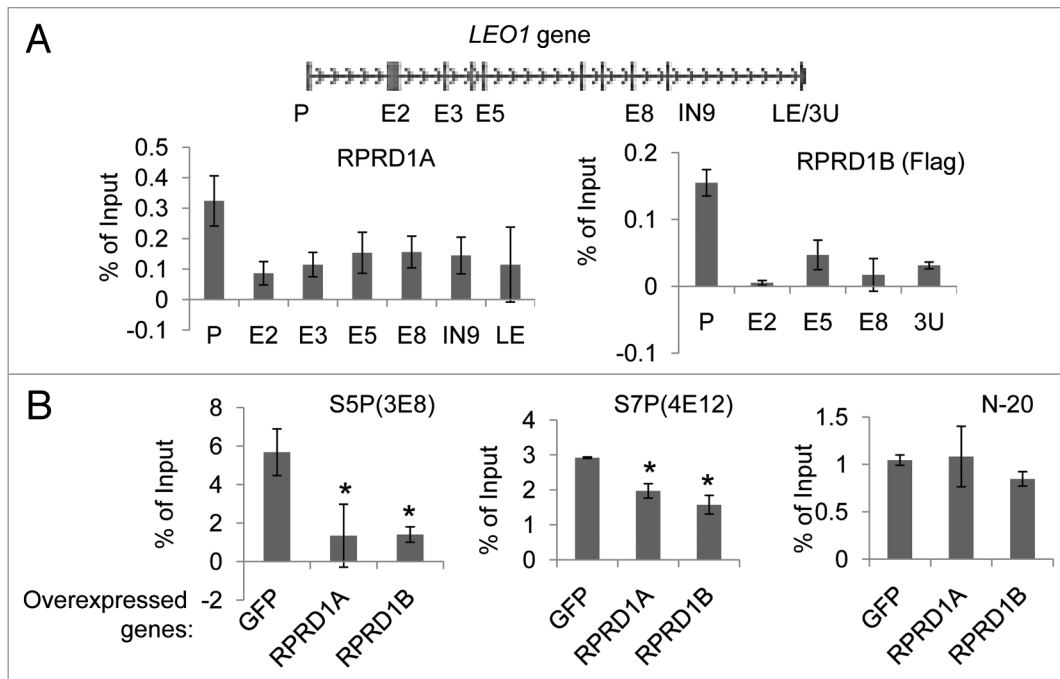


Figure 3. RPRD1A and RPRD1B control the RNAP II phosphorylation state at the *LEO1* gene. (A) Association of RPRD1A and RPRD1B with the *LEO1* gene. ChIP was performed with anti-RPRD1A or anti-Flag antibodies using soluble protein lysates prepared from untransduced HEK293 cells or HEK293 cells expressing tagged RPRD1B, respectively. Immunoprecipitated DNA was quantified by qPCR with primers recognizing the indicated regions of the *LEO1* gene; P, promoter; E, exon; IN, intron; 3U, 3' UTR; LE, last exon. (B) Overexpression of RPRD1A or RPRD1B inhibits RNAP II phosphorylation at the *LEO1* promoter. ChIP experiments using 3E8 and 4E12 antibodies specific for S5P and S7P and N-20 antibody that recognizes all RNAP II phosphoisoforms were performed in HEK293 cells overexpressing the indicated genes. Immunoprecipitated DNA was quantified by qPCR with primers recognizing the *LEO1* promoter region. Promoter occupancies for S5P and S7P were normalized to that of total RNAP II detected by the N-20 antibody and represent averages from a minimum of two biological replicates. *, statistically significant ($p < 0.05$) compared with overexpressed GFP.

gel blotting that overexpression of either RPRD1A or RPRD1B does not globally affect CTD phosphorylation on S2, S5 or S7 in HEK293 cells (data not shown), we could not exclude gene-specific regulation with this experiment. To test this possibility, we first used ChIP to map the occupancy of RPRD1A and RPRD1B proteins at the *LEO1* locus, a gene whose expression is regulated by the RPRD proteins (data not shown). We were able to detect enrichment of both RPRD1A and RPRD1B at the promoter and downstream sequences of the *LEO1* locus, with relatively higher levels at the promoter (Fig. 3A). Since promoters are usually enriched for RNAP II with S5P and S7P,⁴ we performed ChIP experiments to test whether RPRD1A and RPRD1B have any effects on those phosphorylations. Our results showed that ectopic overexpression of either RPRD1A or RPRD1B (Fig. S5) depleted promoter bound forms of RNAP II with S5P and S7P at the *LEO1* promoter (Fig. 3B). Our detection of RNAP II with S2P at the *LEO1* promoter was too low to be reliable (data not shown). Similar results were obtained at the promoter regions of several other genes after overexpressing RPRD1B (Fig. S6). Collectively, these data demonstrate that the RPRD1A and RPRD1B proteins reduce the amounts of RNAP II with S5P and S7P at target gene promoters.

In summary, we identified three novel RNAP II-associated proteins, RPRD1A, RPRD1B and RPRD2. All three RPRD proteins contain CID's. Like other evolutionarily conserved CID-containing proteins that have specific CTD phospho-

epitope preferences,³⁸ RPRD1A and RPRD1B appear to bind preferentially to RNAP II with a phosphorylated CTD. Binding of RPRD1A and RPRD1B to RNAP II results in a specific reduction of RNAP II with S5P and S7P at target gene promoters. Moreover, the co-purification of RPRD proteins with RPAP2 suggests that they may control its putative CTD phosphatase activity. Because of their tight association with RNAP II and ubiquitous expression, RPRD proteins may serve as general transcription factors. The presence of the RPRD proteins in downstream regions, as well as in promoter regions, suggests that they may influence initiation, elongation and/or termination of transcription.

Materials and Methods

HEK293 cells were cultured in Dulbecco's modified Eagle's medium with 10% fetal bovine serum and antibiotics. Lentivirus were produced and used to infect HEK293 cells at a multiplicity of infection <1 as described previously in reference 40. Lentivirus-encoded tagged ORFs and shRNAs were transduced into HEK293 cells as described in reference 30. Transduced cells were selected with puromycin (Sigma) at a concentration of 2 $\mu\text{g}/\text{ml}$ for a minimum of 48 h. Gateway-compatible entry clones were obtained from the following sources: RPRD1A (OHS1770-97649968), RPRD1B (OHS1770-9385393), GRINL1A (OHS1770-9382169), MED7 (OHS1770-9381921),

MED18 (OHS1770-9380635) from the human ORFeome library (Open Biosystems); and POLR2D (IOH13973) and POLR2E (IOH5025) from the UltimateORF collection (Invitrogen). All entry clones were sequence verified. Each ORF was cloned into a lentiviral expression vector as described previously in reference 30.

The following antibodies were used in protein gel blotting, IP-WB and ChIP experiments: Flag (Sigma, F1804), Gal4 (Millipore, 06-262), RPRD1A (P15RS, C-18, Santa Cruz, SC-85089), RPAP2 (ProteinTech Group Inc., 17401-1-AP), POLR2B (Sigma-Aldrich, AV32456), Pol II (8WG16, produced in the Greenblatt lab), Pol II (N-20, Santa Cruz, SC-899), POL II (S2P CTD 3E10, Dirk Eick lab), POL II (S5P CTD 3E8, Dirk Eick lab), POL II (S7P CTD 4E12, Dirk Eick lab), TBL1 (SC-11391, Santa Cruz). See **Table S1** for primer sequences used in qPCR. The shRNA sequence CGG CAG AAA TAG ATG ATA GAA was used to knock down RPRD1A.

Total RNA was extracted by using RNeasy kit (Qiagen) and reverse transcribed into cDNA using SuperScript VILO cDNA Synthesis Kit (Invitrogen) according to the manufacturer's instructions. ChIP assays were performed as described previously in reference 41. To quantify the amount of cDNA and ChIP DNA, real time PCR using primers for each gene (**Table S1**) was performed on a 2 μ l aliquot from a total of 400 μ l or 100 μ l of cDNA or ChIP DNA, respectively, with the SYBR Green kit (Applied Biosystems) using the 7300 Real Time OCR System (Applied Biosystems) in a 10 μ l volume in duplicate. PCRs consisting of 40 cycles of 95°C for 15 sec and 55°C for 30 sec were performed. The copy number was calculated as described previously in reference 41. cDNA values were normalized to β -TUBULIN.

Western blot assays were performed as described previously in reference 30. For IP-WB, cell lysates were incubated overnight at 4°C with 2 μ g of antibody followed by adding 20 μ l of Protein G beads (Sigma) for additional 4 h incubation. After washing with low salt buffer (10 mM TRIS-HCl, pH 7.9, 100 mM NaCl, 0.1% Nonidet P-40), proteins were eluted into protein loading buffer, followed by protein gel blot analysis.

Affinity purification, sample processing and tandem mass spectrometry were performed essentially as previously described in reference 30. Bottom-up shotgun sequencing of tryptic peptide mixtures was performed on an Orbitrap Velos mass spectrometer

(ThermoFisher Scientific, San Jose, CA) with collision activated dissociation (CAD) using a nanoflow reverse phase chromatography and a nanospray ion source (Proxeon). A spray voltage of +2.5 kV was applied. Each high resolution (60,000) full mass scan was followed by 10 sequentially acquired data-dependent MS/MS spectra to derive amino acid sequence information for selected peptide precursor ions. A dynamic target exclusion list was enabled to minimize redundant peptide sampling. RAW files were extracted with the ReAdW program and submitted for database searching using SEQUEST v2.7 and a modified UniProtKB/Swiss-Prot fasta sequence file with additional entries as follows: BSA (Swiss-Prot accession number P02769); GFP (Swiss-Prot accession number P42212); TEV (Swiss-Prot accession number P04517) and streptavidin (Swiss-Prot accession number P22629). Search parameters were set so as to permit one missed cleavage site and a static cysteine modification of +57 (carboxyamidomethylation) using precursor and fragment ion tolerances of ± 3 and 0 m/z, respectively. Protein hits were filtered using the StatQuest program with a confidence threshold of at least 99%.⁴² An additional 20 ppm precursor ion accurate mass filter was implemented to further reduce the false discovery rate. Protein identifications based on the identification of a single peptide were discarded, while common contaminant proteins detected in multiple control GFP purifications were filtered out.

A Student's t-test was used for statistical analysis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank S. Smiley for helpful discussion, D. Roudeva for critical reading of the manuscript, and A.B. Mak, J.A.Y. Fedyshyn, B. Fedyshyn, J.A. Hewel and Y. Li for technical support. This work was supported by grants from the Canadian Institutes of Health Research (to J.M., J.F.G., A.E.), Ontario Research Fund (to J.F.G. and A.E.), Canadian Institutes of Health Research Postdoctoral Fellowship (Z.N.) and Ontario Graduate Scholarship (J.B.O.).

Note

Supplemental material can be found at: www.landesbioscience.com/journals/transcription/article/17803

References

1. Corden JL. Tails of RNA polymerase II. *Trends Biochem Sci* 1990; 15:383-7; PMID:2251729; DOI:10.1016/0968-0004(90)90236-5.
2. Eglhoff S, Murphy S. Cracking the RNA polymerase II CTD code. *Trends Genet* 2008; 24:280-8; PMID:18457900; DOI:10.1016/j.tig.2008.03.008.
3. Komarnitsky P, Cho E, Buratowski S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* 2000; 14:2452-60; PMID:11018013; DOI:10.1101/gad.824700.
4. Kim M, Suh H, Cho E, Buratowski S. Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5 and 7. *J Biol Chem* 2009; 284:26421-6; PMID:19679665; DOI:10.1074/jbc.M109.028993.
5. Mosley AL, Pattenden S, Carey M, Venkatesh S, Gilmore J, Florens L, et al. Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* 2009; 34:168-78; PMID:19394294; DOI:10.1016/j.molcel.2009.02.025.
6. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 2007; 27:262-74; PMID:17643375; DOI:10.1016/j.molcel.2007.06.027.
7. Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. Ssu72 is an RNA polymerase II CTD phosphatase. *Mol Cell* 2004; 14:387-94; PMID:15125841; DOI:10.1016/S1097-2765(04)00235-7.
8. Buratowski S. The CTD code. *Nat Struct Biol* 2003; 10:679-80; PMID:12942140; DOI:10.1038/nsb0903-679.
9. Ahn SH, Kim M, Buratowski S. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* 2004; 13:67-76; PMID:14731395; DOI:10.1016/S1097-2765(03)00492-1.
10. Bartkowiak B, Liu P, Phatnani H, Fuda N, Cooper J, Price D, et al. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev* 2010; 24:2303-16; PMID:20952539; DOI:10.1101/gad.1968210.
11. Buratowski S. Progression through the RNA polymerase II CTD cycle. *Mol Cell* 2009; 36:541-6; PMID:19941815; DOI:10.1016/j.molcel.2009.10.019.
12. Peterlin BM, Price D. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 2006; 23:297-305; PMID:16885020; DOI:10.1016/j.molcel.2006.06.014.

13. Cho EJ, Kobor M, Kim M, Greenblatt J, Buratowski S. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* 2001; 15:3319-29; PMID:11751637; DOI:10.1101/gad.935901.
14. Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher R, et al. TFIIF-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing and termination by RNA polymerase II. *Mol Cell Biol* 2009; 29:5455-64; PMID:19667075; DOI:10.1128/MCB.00637-09.
15. Akhtar MS, Heidemann M, Tietjen J, Zhang D, Chapman R, Eick D, et al. TFIIF kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell* 2009; 34:387-93; PMID:19450536; DOI:10.1016/j.molcel.2009.04.016.
16. Tietjen JR, Zhang D, Rodríguez-Molina J, White B, Akhtar M, Heidemann M, et al. Chemical-genomic dissection of the CTD code. *Nat Struct Mol Biol* 2010; 17:1154-61; PMID:20802488; DOI:10.1038/nsmb.1900.
17. Bentley D. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr Opin Cell Biol* 1999; 11:347-51; PMID:10395561; DOI:10.1016/S0955-0674(99)80048-9.
18. Fabrega C, Shen V, Shuman S, Lima C. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* 2003; 11:1549-61; PMID:12820968; DOI:10.1016/S1097-2765(03)00187-4.
19. Cho EJ, Takagi T, Moore C, Buratowski S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* 1997; 11:3319-26; PMID:9407025; DOI:10.1101/gad.11.24.3319.
20. Schroeder SC, Schwer B, Shuman S, Bentley D. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev* 2000; 14:2435-40; PMID:11018011; DOI:10.1101/gad.836300.
21. Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil J, et al. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol Cell* 2002; 9:1101-11; PMID:12049745; DOI:10.1016/S1097-2765(02)00518-X.
22. Meinhart A, Cramer P. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* 2004; 430:223-6; PMID:15241417; DOI:10.1038/nature02679.
23. Kim M, Krogan N, Vasiljeva L, Rando O, Nedeja E, Greenblatt J, et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 2004; 432:517-22; PMID:15565157; DOI:10.1038/nature03041.
24. Aygün O, Xu X, Liu Y, Takahashi H, Kong S, Conaway R, et al. Direct inhibition of RNA polymerase II transcription by RECQL5. *J Biol Chem* 2009; 284:23197-203; PMID:19570979; DOI:10.1074/jbc.M109.015750.
25. Kanagaraj R, Huehn D, MacKellar A, Menigatti M, Zheng L, Urban V, et al. RECQ5 helicase associates with the C-terminal repeat domain of RNA polymerase II during productive elongation phase of transcription. *Nucleic Acids Res* 2010; 38:8131-40; PMID:20705653; DOI:10.1093/nar/gkq697.
26. Eglhoff S, Szczepaniak S, Dienstbier M, Taylor A, Knight S, Murphy S. The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J Biol Chem* 2010; 285:20564-9; PMID:20457598; DOI:10.1074/jbc.M110.132530.
27. Ng HH, Robert F, Young R, Struhl K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* 2003; 11:709-19; PMID:12667453; DOI:10.1016/S1097-2765(03)00092-3.
28. Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng M, et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* 2003; 11:721-9; PMID:12667454; DOI:10.1016/S1097-2765(03)00091-1.
29. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadian V, et al. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 2003; 23:4207-18; PMID:12773564; DOI:10.1128/MCB.23.12.4207-18.2003.
30. Mak AB, Ni Z, Hewel J, Chen G, Zhong G, Karamboulas K, et al. A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol Cell Proteomics* 2010; 9:811-23; PMID:20305087; DOI:10.1074/mcp.M000002-MCP201.
31. Lunde BM, Reichow S, Kim M, Suh H, Leeper T, Yang F, et al. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* 2010; 17:1195-201; PMID:20818393; DOI:10.1038/nsmb.1893.
32. Cui M, Allen M, Larsen A, Macmorris M, Han M, Blumenthal T. Genes involved in pre-mRNA 3'-end formation and transcription termination revealed by a lin-15 operon Muv suppressor screen. *Proc Natl Acad Sci USA* 2008; 105:16665-70; PMID:18946043; DOI:10.1073/pnas.0807104105.
33. Ingham RJ, Colwill K, Howard C, Dettwiler S, Lim C, Yu J, et al. WW domains provide a platform for the assembly of multiprotein networks. *Mol Cell Biol* 2005; 25:7092-106; PMID:16055720; DOI:10.1128/MCB.25.16.7092-106.2005.
34. Wu Y, Zhang Y, Zhang H, Yang X, Wang Y, Ren F, et al. p15RS attenuates Wnt/ β -catenin signaling by disrupting β -catenin-TCF4 interaction. *J Biol Chem* 2010; 285:34621-31; PMID:20739273; DOI:10.1074/jbc.M110.148791.
35. Liu J, Liu H, Zhang X, Gao B, Wang J, Hu Z. Identification and characterization of P15RS, a novel P15(INK4b) related gene on G₁/S progression. *Biochem Biophys Res Commun* 2002; 299:880-5; PMID:12470661; DOI:10.1016/S0006-291X(02)02684-0.
36. Hu X, Malik S, Negroiu C, Hubbard K, Velalar C, Hampton B, et al. A Mediator-responsive form of metazoan RNA polymerase II. *Proc Natl Acad Sci USA* 2006; 103:9506-11; PMID:16769904; DOI:10.1073/pnas.0603702103.
37. Max T, Søgaard M, Svestrup J. Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J Biol Chem* 2007; 282:14113-20; PMID:17376774; DOI:10.1074/jbc.M701345200.
38. Lima CD. Inducing interactions with the CTD. *Nat Struct Mol Biol* 2005; 12:102-3; PMID:15702066; DOI:10.1038/nsmb0205-102.
39. Chapman RD, Heidemann M, Albert T, Mailhammer R, Flatley A, Meisterernst M, et al. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* 2007; 318:1780-2; PMID:18079404; DOI:10.1126/science.1145977.
40. Moffat J, Gruenberg D, Yang X, Kim S, Kloepper A, Hinkle G, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 2006; 124:1283-98; PMID:16564017; DOI:10.1016/j.cell.2006.01.040.
41. Ni Z, Bremner R. Brahma-related gene 1-dependent STAT3 recruitment at IL-6-inducible genes. *J Immunol* 2007; 178:345-51; PMID:17182572.
42. Kislinger T, Rahman K, Radulovic D, Cox B, Rossant J, Emili A. PRISM, a generic large scale proteomic investigation strategy for mammals. *Mol Cell Proteomics* 2003; 2:96-106; PMID:12644571; DOI:10.1074/mcp.M200074-MCP200.