Supporting Online Material

Materials and Methods

Transfections and RNase protections

Transfections were carried out as outlined in (*1*) using VAI as an internal control. The specific activity of VAI riboprobe was not necessarily the same in different experiments. Endogenous Rpb1 is inhibited by addition of α-amanitin after transfection (*1*).

Products of RNase protection of transcripts from the U2G template correspond to properly initiated RNA that has undergone 3' box-directed processing to give the same 3' end as endogenous pre-U2 (this RNA lacks the signals to be processed to the mature 3' end in the cytoplasm (*1*)) (Proc), properly initiated but unprocessed RNA (Unproc) and unprocessed RNA which may result from pol II "reading around" the plasmid and back over the gene due to inefficient termination (*1*) (-111Unproc) (Figures 1A, S1A, S2A). Transcripts from endogenous U2 snRNA genes give products corresponding to U2 snRNA with a mature 3' end (U2), transcripts that have undergone processing directed by the 3' box (pre-U2) and transcripts that have escaped processing (U2Unproc) (Figures 1B, S1A). Products of RNase protection of transcripts from pCMV-hnRNPK correspond to RNA processed at the polyadenylation site (Proc) and unprocessed RNA (Unproc) (Figures 1A, S1A, S2B).

The % RNA (Figures 1A, S1B, S2A, S2B, S2D) was calculated by adding the values for Proc and Unproc from phosphorimager quantitation after correction to VAI and normalizing to $(Con)^{25}$ or $(Con)^{48}$. Unproc/Proc is the ratio of the values from phophorimager quantitation corrected for the length of the product.

RT-PCR analysis

2μg of total RNA from Raji cells expressing α-amanitin-resistant Rpb1 (*2*) was reverse transcribed with random hexamers and U2 pre-snRNA and hElf-1 mRNA were quantitated by real-time PCR analysis using specific primers (primers amplify the region +61/+196 of the U2 gene (Accession number U57614) and +1297/+1579 of the hElf-1 gene (Accession number NM_172373)). The results were corrected to the level of 7SK snRNA (primers amplify +11/+190 (Accession number NR_001445)). Error bars indicate the range of values from three independent experiments.

CTD mutations

Mutations in Ser2 and Ser5 in a background of 25 repeats were produced by building on shorter mutant CTDs (*3*) as described (*4*). $S7A^{25}$ was produced using a similar strategy based on restriction sites recognized by NgoMIV. Production of $S7A^{48}$ is described (*2*). $(Con)^{25}$ contains 25 consensus repeats in the CTD and is 1-25 NotI (*4*). $(Con)^{48}$ contains 48 consensus repeats in the CTD (*2*).

Chromatin Immunoprecipitation

293 cells were transfected with Rpb1 constructs and endogenous Rpb1 was inhibited by the addition of α-amanitin (*1*). ChIP analysis (*4*) was carried out 48hrs later with antibodies to the N-terminus of Rpb1 (Santa Cruz, N-20) and TAP-tag (rabbit pre-immune serum) (Figure 2A) or PTFγ (*5*) (Figure S3B). Error bars indicate the range of values from three independent experiments. Oligos used for ChIP amplified the -920/-717,

-40/+33 and +940/+1149 regions of U2 genes, the -103/-41 region of U1 gene (Accession number J00318), the -102/+72 region of the γ-actin gene (Accession number M19283) and the -119/+30 region of the GAPDH gene (Accession number AY340484). ChIP analysis with the anti-phospho ser7 (4E12) (*6*) was carried out on untransfected 293 cells.

Western blotting

Western blot analysis was carried out essentially as described (*7*) with antibodies to Rpb1 (Santa Cruz, H-224 (Figsures S1C, S2C)), stably expressed HA-tagged Rpb1 (anti-HA antibody, Roche 12CA5 (Figure 1B)), Rpb2 (Abcam, ab10338 (Figures S1C, S2C)), phospho-serine 5 (Covance, H14 (Figures S1C, S2C, S3C, S3D), phospho-serine 2 3E10 (*6*) (Figures S1C, S2C, S3C, S3D), Oct-1 (Santa Cruz, sc-232 (Figures 1B, S2C), TBP (Santa Cruz, sc273 (Figure S1C), phospho-serine 7 4E12 (*6*) (Figures S3C, S3D) and Int11 (RC68) (*8*) (Figure 2C). For Western blots of protein from whole cells (Figures S1C, S2C), cell pellets were resuspended directly in 2x SDS loading dye and boiled before loading. Protein samples from GST-CTD pull down or phospho-GST-CTD (Figures 2C, S3C, S3D) were boiled after the addition of an equal volume of 2x SDS loading dye before loading.

GST-CTD pull down

1.5mg of the P11 0.5M KCl fraction (pre-cleared with GST beads for 1 h at 4ºC) was incubated at 4ºC for 3 h with 20μg of each GST fusion protein bound to 30μl of GST beads (*9*). The beads were washed with 5x 1ml 0.1M KCl HEGN (*9*) 1x 1ml 0.3M KCl HEGN, 1x 1ml 0.5M KCl HEGN and an additional wash with 0.1M KCL HEGN. Bound protein was eluted by incubating twice with 0.1ml 0.3% sarkosyl in 0.1M KCl HEGN for 15 min at 4ºC. The same amount of each fusion protein was recovered after pull down. Eluate was analysed by Western blot using anti-RC68 antibody (*8*). $1/20^{th}$ of the eluate was loaded and $1/200^{th}$ of the input was loaded for comparison. We have therefore indicated that $1/10^{th}$ equivalent of the input was loaded (Figure 2C). Phosphorylation and subsequent purification of GST-CTD was carried out as described (*10*) using HeLa cell nuclear extract.

Nuclear run on analysis

Nuclear run on analysis was carried out as previously described (*1*). For analysis of the endogenous U2 genes, three contiguous 80-mer oligos complementary to the region 208-447 of the U2 gene sequence (Accession number U57614) were used, taking the first base pair of the U2 coding region as 1. For analysis of the U1 genes, two 80-mer oligos complementary to the region 5-164 of the U1 coding region were used and for 7SK genes, two 80-mer oligos complementary to the region 125-284 of the 7SK coding region were used. For analysis of Rpb1, three 60-mer oligos complementary to the region 421-600 of Rpb1 mRNA (Accession number NM_000937) were used. AS, which is complementary to the complement of nucleotides 96-175 of the U2 gene sequence and Up, which is complementary to nucleotides -268 to -189 (U2 promoter) were used as negative controls. The endogenous pol III-dependent 7SK gene serves as an α-amanitin-resistant control and the level of transcription of the transfected Rpb1 gene serves as a transfection and protein-coding gene control.

Supplementary Figures

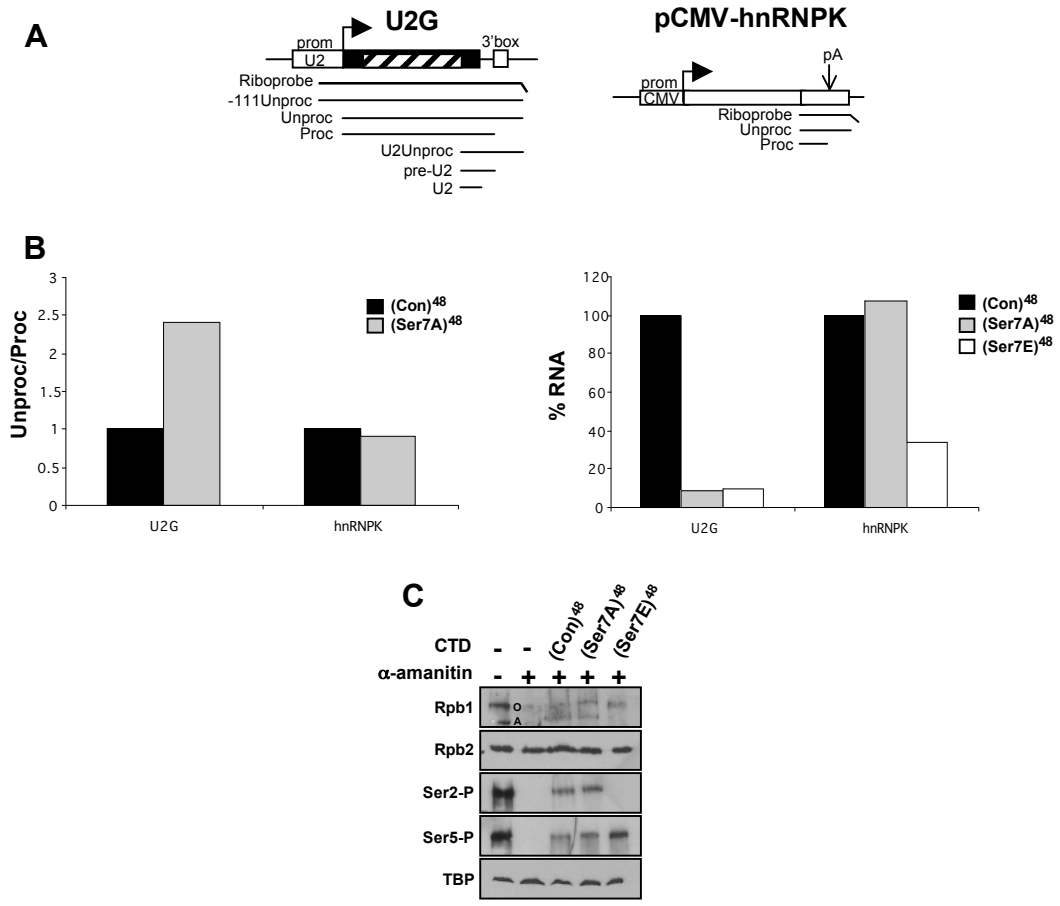

Figure S1. Ser7 is required for expression of snRNA but not protein-coding templates.
(**A**) Diagrammatic representation of the transfected U2G and pCMV-hnRNPK constructs. The position of the riboprobes and expected RNase protection products are shown (see Materials and Methods). For U2G, Proc corresponds to properly initiated RNA that has undergone 3' box-directed processing to give the same 3' end as endogenous pre-U2. Unproc corresponds to properly initiated but unprocessed RNA and -111Unproc corresponds to unprocessed RNA which may result from pol II "reading around" the plasmid and back over the gene (Figures 1A, S2A), U2 corresponds to endogenous U2 snRNA that has the mature 3' end, pre-U2 corresponds to transcripts from endogenous U2 genes where processing directed by the 3' box has occurred and U2Unproc corresponds to unprocessed transcripts from endogenous U2 genes (Figure 1B). For pCMV-hnRNPK, Proc corresponds to RNA processed at the polyadenylation site (pA) and Unproc corresponds to unprocessed RNA (Figures 1A, S2B). (**B**) Graphic representation of the data shown in Figure 1A. (**C**) Western blot analysis of Rpb1, Rpb2, phospho-serine 2, phospho-serine 5 and TBP from transfected cells. The antibodies used are indicated on the left. The hypo- and hyper-phosphorylated forms of wild type Rpb1 are marked as A and O respectively.

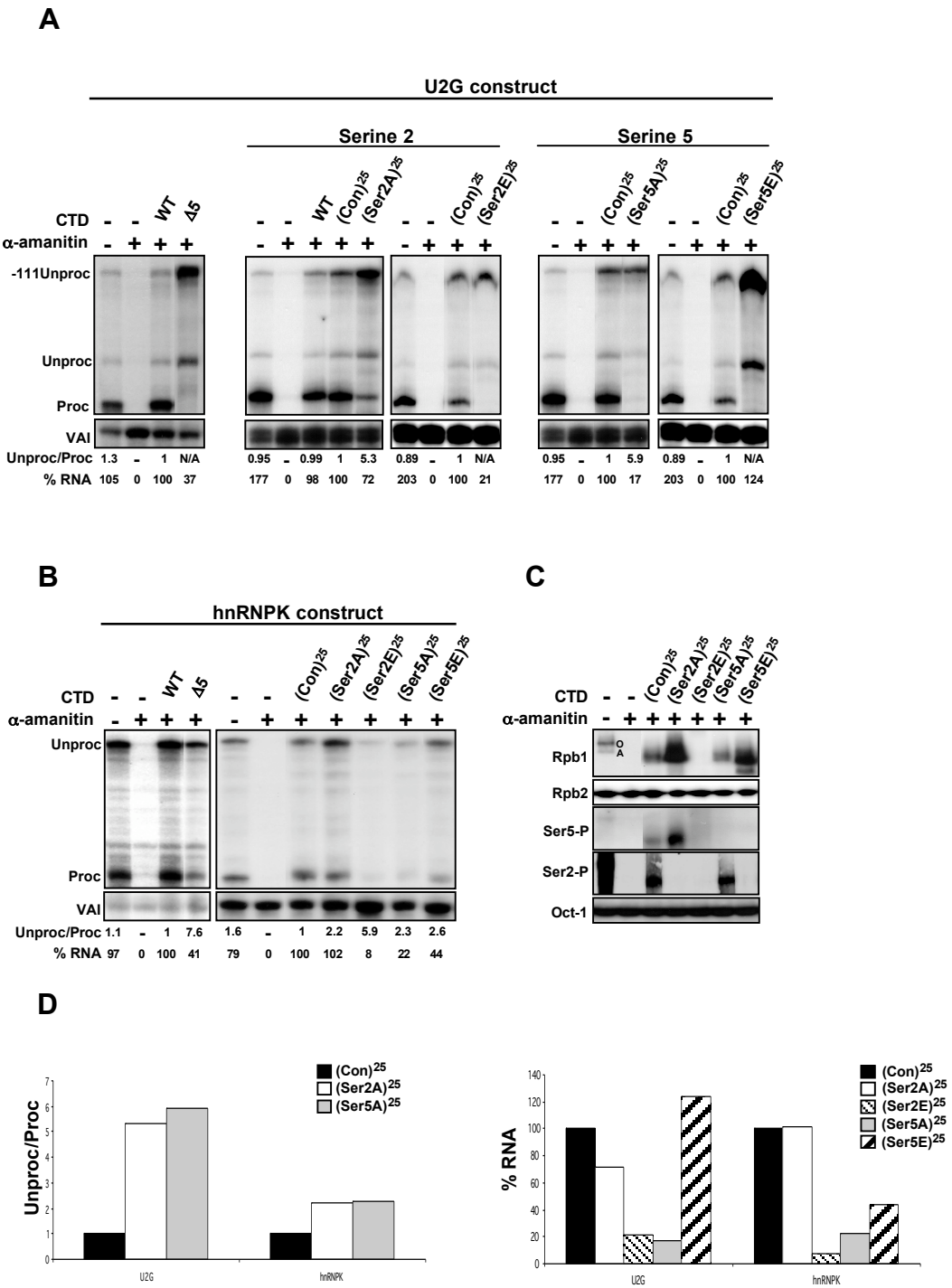

Figure S2. Ser2 and Ser5 of the CTD heptapeptide are required for efficient expression of both snRNA and protein-coding templates.

(**A**) RNase protection analysis of transcripts from the U2G construct after ectopic expression of α-amanitin-resistant Rpb1. (**B**) RNase protection analysis of transcripts from the pCMV-hnRNPK construct. (**C**) Western blot analysis of Rpb1, Rpb2, phospho-serine 5, phospho-serine 2 and Oct-1 from transfected cells. The antibodies used are indicated on the left. (**D**) Graphic representation of the data shown in **A** and **B**.

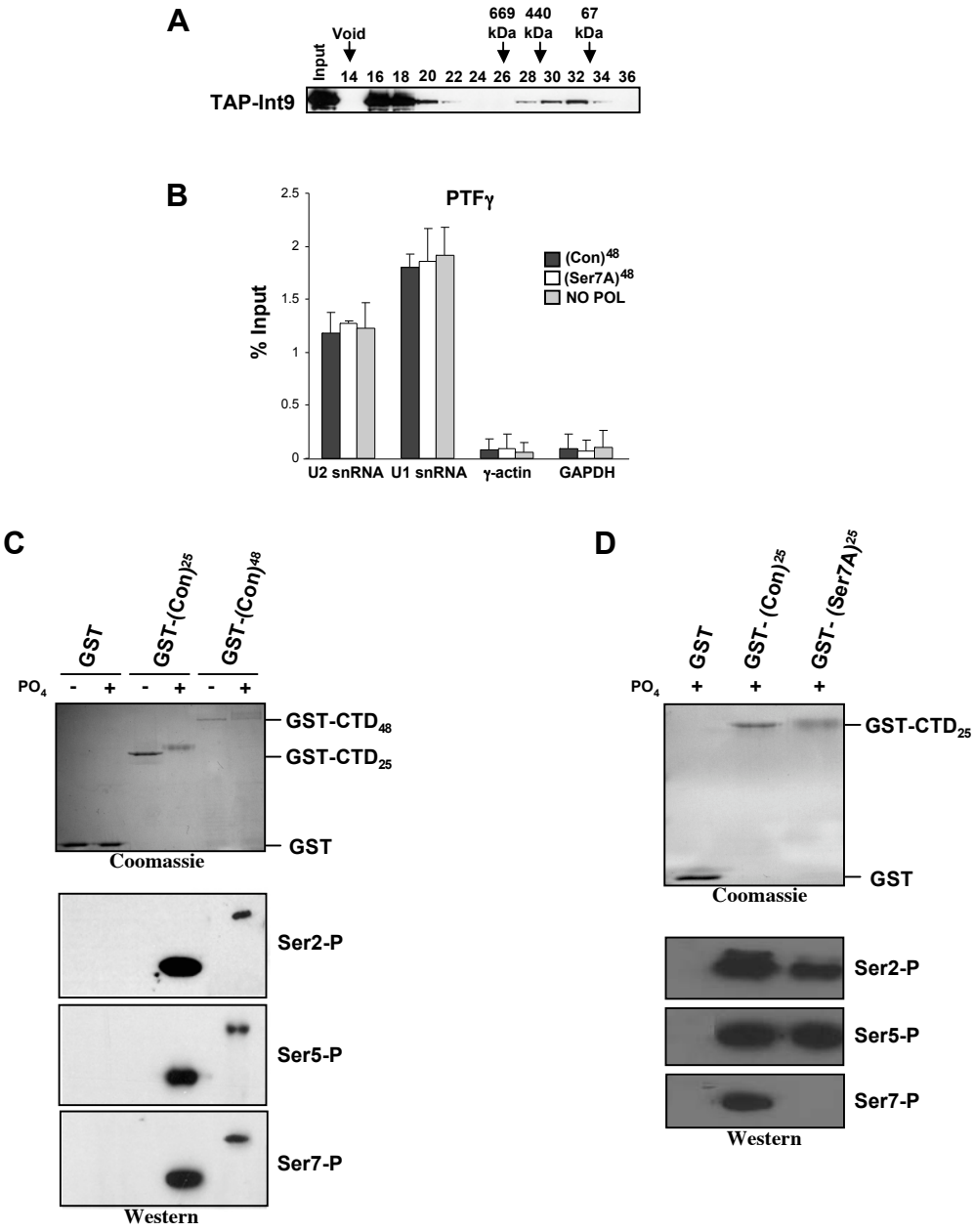

Figure S3. Mutation of Ser7 to alanine affects association of Integrator with snRNA genes.

(**A**) TAP-Int9 in nuclear extract from transfected cells fractionates on a Superose 6 column as expected (*9*). (**B**) ChIP analysis of PTFγ associated with U1, U2, γ-actin and GAPDH promoters. (**C**) Coomassie-staining and Western blot analysis of GST-CTD proteins with 25 or 48 consensus repeats before and after phosphorylation. The antibodies used are indicated on the right here and in (D). (**D**) Coomassie-staining and Western blot analysis of GST-CTD proteins with 25 consensus or S7A repeats after phosphorylation.

Supporting References

1. J. E. Medlin, P. Uguen, A. Taylor, D. L. Bentley, S. Murphy, *Embo J* **22**, 925 (2003).
2. R. D. Chapman, M. Conrad, D. Eick, *Mol Cell Biol* **25**, 7665 (2005).
3. M. L. West, J. L. Corden, *Genetics* **140**, 1223 (1995).
4. J. Medlin *et al.*, *Embo J* **24**, 4154 (2005).
5. J. B. Yoon, S. Murphy, L. Bai, Z. Wang, R. G. Roeder, *Mol Cell Biol* **15**, 2019 (1995).
6. R. D. Chapman *et al.* (*companion manuscript*).
7 M. F. Dubois, V. T. Nguyen, S. Bellier, O. Bensaude, *J. Biol. Chem.* **269**, 13331 (1994)
8. Z. Dominski, X. C. Yang, M. Purdy, E. J. Wagner, W. F. Marzluff, *Mol Cell Biol* **25**, 1489 (2005).
9. D. Baillat *et al.*, *Cell* **123**, 265 (2005).
10. Y. Hirose, J. L. Manley, *Nature* **395,** 93 (1998).