The interaction between the Spt6-tSH2 domain and Rpb1 affects multiple functions of RNA Polymerase II

Zaily Connell1, Timothy J. Parnell2, Laura L. McCullough1, Christopher P. Hill1, and Tim Formosa1,*

1Dept of Biochemistry, University of Utah School of Medicine 50 N Medical Drive, Rm 4100, Salt Lake City, UT 84112
2Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT 84112

* To whom correspondence should be addressed. Tel: +1 801 581-5435 Fax: +1 801 581-7959 Email: Tim@biochem.utah.edu

Supplemental Data

Table S1 Yeast Strains used

All strains were congenic with the A364a background.

The N-terminal SH2 domain of Spt6 uses a standard SH2-motif pocket to bind the Rpb1 phospho-threonine residue pT1471 (1). To accommodate this non-canonical SH2 ligand, pT1471 and the nearby unmodified tyrosine Y1473 are positioned to mimic a phospho-tyrosine, with the intermediate residue P1472 having an important role in achieving the appropriate geometry of this arrangement. Alanine substitutions of T1471 and Y1473 were tested individually, but the strongest effects on binding and growth phenotypes were observed upon substitution of all three residues, T1471A, P1472A, and Y1473A (1 and Fig S1B). An R1282H substitution in Spt6 was sufficient to disrupt the binding pocket for these residues in vitro (1). Alleles are designated here by the WT residue as a superscript to reduce the length of the name. For example, the Spt6 allele spt6-R1282H is designated spt6R and rpb1-T1471A is designated as rpb1T.

The C-terminal SH2 domain of Spt6 uses a non-canonical binding groove to make several points of contact with the Rpb1 linker (1), and disrupting this interaction required mutating two lysines in the binding pocket (K1355A and K1435A). Similarly, while the primary target of binding was the phosphoserine residue pS1493, fully disrupting this interaction in vitro required substitutions with alterations in properties for several residues in the 1490-1496 region of Rpb1 from ...LMFSPLV... to ...KMKRRK. These are designated here as spt6KK and rpb1FSP.

Compound mutants are designated with a comma separating the N-terminal and C-terminal sites (spt6R,KK and rpb1TPY,FSP). These alleles were introduced into the native genomic SPT6 and RPB1 genes by markerless conversion (2) as described previously (1). A selectable marker was inserted downstream of each gene as indicated in each genotype to allow the mutated alleles to be followed in crosses. Each construct was backcrossed multiple times to ensure that all phenotypes associated with the allele were linked to the mutated locus before use.

| Genotype | Strain name | Genotype |
|----------|-------------|----------|
| WT       | 8127-7-4    | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ |
| spt6R    | 8662-2-2K   | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ spt6-R1282H(+40, KanMX) |
| spt6KK   | 9859-2-2K   | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ spt6-K1355A,K1435A(+40, KanMX) |
| spt6R,KK | 9910-1-4    | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ spt6-K1355A,K1435A(+40, KanMX) |
| rpb1TPY  | 9992-1-3    | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ rpb1-TPY1471AAA(+100, KanMX) |
| rpb1FSP  | 9800-8-4    | MATa ura3Δ0 leu2Δ0 trp1Δ2 his3 lys2-128Δ rpb1-LMFSPLV1490KMKRRRK(+100, KanMX) |
| rpb1<sup>TPY,FSP</sup> | 9998-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-TPY1471AAA,LMF spindle1490KMRRRK(+100, KanMX) |
|----------------------|--------|--------------------------------------------------|

Spt6 ChIP (all four strains), MNase-seq (WT, rpb1<sup>TPY,FSP</sup> only), Figs 7, S7, S8I

| WT | 8127-7-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ |
|---|--------|----------------------------------|
| spt6<sup>KK</sup> | 9910-1-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| rpb1<sup>T,FSP</sup> | 9930-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-T1471A,LMF spindle1490KMRRRK(+100, KanMX) |
| rpb1<sup>T,FSP</sup> | 9834-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-Y1473A,LMF spindle1490KMRRRK(+100, KanMX) |

Growth Phenotypes-Spot dilutions, Fig S1B

| WT | 9995-2-2 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(HphMX) |
|---|--------|--------------------------------------------------|
| rpb1<sup>T</sup> | 9871-6-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(HphMX) rpb1-T1471A(+100, KanMX) |
| rpb1<sup>T</sup> | 9955-1-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(NatMX) rpb1-Y1473A(+100, KanMX) |
| rpb1<sup>TPY</sup> | 10195-3B | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(HphMX) rpb1-TPY1471AAA(+100, KanMX) |
| rpb1<sup>FSP</sup> | 9819-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(NatMX) rpb1-TPY1471AAA,LMF spindle1490KMRRRK(+100, KanMX) |
| rpb1<sup>T,FSP</sup> | 9940-5-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(NatMX) rpb1-T1471A,LMF spindle1490KMRRRK(+100, KanMX) |
| rpb1<sup>T,FSP</sup> | 10007-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(NatMX) rpb1-Y1473A,LMF spindle1490KMRRRK(+100, KanMX) |
| rpb1<sup>T,FSP</sup> | 10003-4-2 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ GAL1pr-flo8-HIS3(HphMX) rpb1-TPY1471AAA,LMF spindle1490KMRRRK(+100, KanMX) |

Primer Extension, ADH2 qPCR, Fig 2D, S2B

| WT | 8127-7-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ |
|---|--------|----------------------------------|
| spt6<sup>KK</sup> | 9910-1-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| rpb1<sup>TPY,FSP</sup> | 9998-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-TPY1471AAA,LMF spindle1490KMRRRK(+100, KanMX) |

qPCR, Figs 2B, S2D, 3A-D, 4B-D, S4E-G,I, S5A-D, S8K

| WT | 8127-7-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ |
|---|--------|----------------------------------|
| rpb1<sup>T</sup> | 9866-1-1 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-T1471A(+100, KanMX) |
| spt6<sup>KK</sup> | 9910-1-4 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| rpb1<sup>TPY,FSP</sup> | 9998-1-3 | MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128Δ rpb1-TPY1471AAA,LMF spindle1490KMRRRK(+100, KanMX) |

qPCR, Figs 2C, S2E,H, S3F-I, 5C-D, S8L

| WT | 10152-1A | MATa leu2-Δ0 his3 lys2-128Δ |
|---|----------|----------------------------------|
| rrp6-Δ | 10212-2D | MATa leu2-Δ0 his3 lys2-128Δ rrp6-Δ::KanMX |
| spt6<sup>KK</sup> | 10205-2D | MATa leu2-Δ0 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, NatMX) |
| rrp6-Δ | 10212-4D | MATa leu2-Δ0 his3 lys2-128Δ rrp6-Δ::KanMX spt6-
| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| spt6<sup>R,KK</sup> | R1282H,K1355A,K1435A(+40, NatMX) |
| rpb1<sup>TPY,FSP</sup> | 10183-3C | MATa leu2-∆0 his3 lys2-128Δ rpb1-TPY1471AAA,LMFSPLV1490KMKRRRK(+100, KanMX) |
| rrp6-Δ | 10213-22C | MATa leu2-∆0 his3 lys2-128Δ rrp6-Δ(Δ::KanMX) rpb1-TPY1471AAA,LMFSPLV1490KMKRRRK(+100, NatMX) |

qPCR, Figs 2E, S2K, 4E

| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| WT | 8127-7-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ |
| spt6<sup>R,KK</sup> | 9910-1-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| rpb1<sup>G1097C</sup> | 9139-H | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ rpb1-G1097C(+100, HphMX) |
| spt6<sup>R,KK</sup> | 10118-1B | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ rpb1-G1097C(+100, HphMX) spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| rpb1<sup>H1085Q</sup> | 9518 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ rpb1-H1085Q(+100, KanMX) |
| spt6<sup>R,KK</sup> | 10119-3D | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ rpb1-H1085Q(+100, KanMX) spt6-R1282H,K1355A,K1435A(+40, NatMX) |

qPCR, +/- 6AU, Figs S2J, S6

| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| WT | 10152-1A | MATa leu2-∆0 his3 lys2-128Δ |
| spt6<sup>R,KK</sup> | 10158-5B | MATa leu2-∆0 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, NatMX) |
| rpb1<sup>TPY,FSP</sup> | 10183-3C | MATa leu2-∆0 his3 lys2-128Δ rpb1-TPY1471AAA,LMFSPLV1490KMKRRRK(+100, KanMX) |

qPCR, Figs S3B-E,J,K, 5E-F, S5F-K, S8M

| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| WT | 8127-7-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ |
| spt6<sup>R,KK</sup> | 9910-1-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| spt6<sup>F249K</sup> | 9506-1-3 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ spt6-F249K(-424, KanMX) |
| pob3<sup>Q308K</sup> | 9539-1-1 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ pob3-Q308K(+34, NatMX) |
| spt16<sup>TT</sup> | 9332-5-3 | MATa ura3 leu2 trp1 his3 lys2-128Δ spt16-11(+124, KanMX) |

qPCR, Fig S3K

| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| WT | 8127-7-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ |
| spt6<sup>R,KK</sup> | 9910-1-4 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ spt6-R1282H,K1355A,K1435A(+40, KanMX) |
| spt6<sup>F249K</sup> | 9506-1-3 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ spt6-F249K(-424, KanMX) |
| rpb1<sup>TPY,FSP</sup> | 9998-1-3 | MATa ura3-∆0 leu2-∆0 trp1-∆2 his3 lys2-128Δ rpb1-TPY1471AAA,LMFSPLV1490KMKRRRK(+100, KanMX) |

qPCR, +/- 6AU Figs 6A-D

| Gene 1 | Gene 2 | Mutation/Condition |
|--------|--------|--------------------|
| WT | 10152-1A | MATa leu2-∆0 his3 lys2-128Δ |
| Target Gene | Gene Region detected | Oligo Name | Oligo Sequence (5’ to 3’) |
|-------------|----------------------|------------|--------------------------|
| **ADE12**   |                      |            |                          |
| TSS1        | TF1951               | TAGAAGCTTACCTCGACCGAG |
| TSS1        | TF1952               | GCCATTGTTACCACCGGCG  |
| mRNA        | TF1953               | AGCTAAAGGTTTGGAAGAACGCA |
| mRNA        | TF1954               | CACCGGGTTGATCATTACC   |
| **ADE17**   |                      |            |                          |
| mRNA        | TF2088               | CAGCAGCAGCATCGTTCAG   |
| mRNA        | TF2089               | ATTCTATCAGCACCACGGGC  |
| **ADH1**    |                      |            |                          |
| Primer extension | TF1984      | Cy5-GCCTTTGGCTTTGGAACCTGGA |
| TSS1        | TF2105               | AAGCTATACCAAGCATACAAATCAAC |
| mRNA        | TF2107               | ACGAATCCCACCGGTAAGTTG |
| TSS1 or mRNA | TF2106               | ACCTTCGTGACCACCGACTA |
| **ADH2**    | (region with low homology to ADH1) |            |                          |
| mRNA        | TF2108               | CGCAGTCGTTAAGGCTACCA  |
| mRNA        | TF2109               | CCCACGTAAGCCCGACAATG  |
| **ARO9**    |                      |            |                          |
| mRNA        | TF2090               | TGAGCTAACCATTCACGTT   |
| mRNA        | TF2091               | CATTGAATCGTGGACCCGC   |
| **BIO2**    |                      |            |                          |
| TSS1        | TF1939               | TGCTCCCATACTTGGCTAAAATT |
| TSS1        | TF1940               | CGATTTAACTGCAGCATTGG  |
| mRNA        | TF1941               | TGCCGCAGTTAATCGACAAC  |
| mRNA        | TF1942               | TGCACTTTGGTTGGATCGTG  |
| **CHA1**    |                      |            |                          |
| mRNA set 1  | TF2092               | AGTATGCAGCGTTGGTGGAG  |
| mRNA set 1  | TF2093               | TCCACCCCCACATAGGGAT   |
| mRNA set 2  | TF2097               | CTGCTGCAAACAGCATGCAA  |
| mRNA set 2  | TF2098               | AGGCACCACTACAGTAACC   |
| **IMD2**    |                      |            |                          |
| TSS1-CUT    | TF1912               | GGCTTATACATTTTACCTAGTATGCTGG |
| TSS1-CUT    | TF1913               | TGTAATAAACAATAATGCATTGGATAGCAGT |
| mRNA        | TF1914               | GCTGGGTTGGATGTTCGTCA  |
| mRNA        | TF1915               | CAATTCTCAACCGTCCGCAC  |
| Primer extension | TF1885          | Cy5-TGTAGTCTCTAATGGCGCGCC |
| **IMD3**    |                      |            |                          |
| Upstream CUT | TF2023             | GTGTACTGGGAACTTTTCTCTTTTTCAC |
| Upstream CUT | TF2024             | CCTGTCGATGACTTTTCTCTTAG |
| Gene | Region | TF | Sequence |
|------|--------|----|----------|
| mRNA | TF2025 | ACAGAATCGAAATGGCCATCTTC |
| mRNA | TF2026 | ACCGGAAAATCCAAATCTTTCCCTAT |
| IMD4 | 5' exon | TF2060 | ACAGTTCCAAGGACCGTCTT |
| IMD4 | 5' exon | TF2061 | GCCATATCGCTTCAGTGACTG |
| IMD4 | 5' exon - intron junction | TF1904 | CAACCACCAGTGTGGTGAAG |
| IMD4 | 5' exon - intron junction | TF1905 | GCAGCATCTAATGAAAAGAACATAATTAC |
| intron | TF1974 | AGAATGCAGCTTTTTTAATCATCCA |
| intron | TF1975 | TCCGAAAAAAGGCAGAAT |
| intron - 3' exon junction | TF1906 | GGCTGGGCTGAAAATTTCTTCAG |
| intron - 3' exon junction | TF1907 | CACGAGAGGACTAAACCAAAC |
| 3' exon | TF1908 | TTCCATGTGTCAAGAGCGGAT |
| 3' exon | TF1909 | CAACATCCCAACCTGCTTCAG |
| MET3 | TSS1 | TF2029 | ACCAATGAGAAATAGAGGCTCTT |
| MET3 | TSS1 | TF2030 | ACAGATACCAACTCGTTAGCAG |
| mRNA | TF2031 | ACCCAGAACATCCAGCCATT |
| mRNA | TF2032 | ACACGGTCCCATTGTCTTGA |
| NRD1 | 5' end | TF2035 | CTAACTCACCAATCAAGGCTCTTC |
| NRD1 | 5' end | TF2036 | TCCTGTAATCTCGTCTTGAGG |
| NRD1 | 3' end | TF2037 | TGGTCAGTAAATGCGCAATG |
| NRD1 | 3' end | TF2038 | ATTGCTGTGTTGGAGGAT |
| RPL17A | 5' exon | TF1922 | TGCTAAATCCGCTTTCTGCTCG |
| RPL17A | 5' exon | TF1923 | AGGAATAGCCTCTTGGTGTCATA |
| RPL17A | 5' exon - intron junction | TF2048 | TCAGATAACTCGTCAAGGT |
| RPL17A | 5' exon - intron junction | TF2049 | GGAATTTCCGCGCATTTTGA |
| intron | TF1924 | TTCAAAATGCCGGAAAATCC |
| intron | TF1925 | TCAGCTTGACAGTTCCATTCATCA |
| intron - 3' exon junction | TF2050 | TGAACTGTCTAAGCTGAAAGC |
| intron - 3' exon junction | TF2051 | GCATCCAAACCTTTAGCCTGT |
| 3' exon | TF2052 | CTCAATCAGTCTAAACAGCCTCA |
| 3' exon | TF2053 | ACAGACCCTTCTTTCACAGC |
| RPL18A | 5' exon - intron junction | TF1898 | CGGATTGCCATCGTTATCGATCAC |
| RPL18A | 5' exon - intron junction | TF1899 | CCATAATTGCGTGGTTAAAAAC |
| intron | TF1976 | TCCGATCATCTACAGTTGCT |
| intron | TF1977 | AGATTCTGCTCATGTTGTCG |
| intron - 3' exon junction | TF1900 | TCCGGTCCTCAGTCAATTAGTAC |
| intron - 3' exon junction | TF1901 | GTCTTGTAGCAGACCTTCTTGC |
| Gene       | Type                    | Accession | Sequence                          |
|------------|-------------------------|-----------|-----------------------------------|
| 3' exon    | TF1902                  | GACCACTGTGTACTGCTTTTGAGA |
| 3' exon    | TF1903                  | GTGGACCCATACCGAAAGTGTC  |
| RPL23A     | 5' exon - intron junction | TF1893    | GGTAACGGTGCTCAAGGTACTAAG        |
|            | 5' exon - intron junction | TF1895    | GCATAGGAGAAAGTTCAGAGATGT        |
| intron     | TF1894                  | TCATACACGGCAAGCATATCCTAAC|
| intron     | TF1895                  | GCATAGGAGAAAGTTCAGAGATGT|
| intron - 3' exon junction | TF1894 | TCATACACGGCAAGCATATCCTAAC |
| intron - 3' exon junction | TF1897 | GGATTAGCGATGACACCACAGCAT      |
| 3' exon    | TF1896                  | GGTAAAGCCAGATTGGAAGAGAGGT  |
| 3' exon    | TF1897                  | GGATTAGCGATGACACCACAGCAT  |
| SRG1-SER3  | SRG1                    | TF1945    | AAGGGGAACAAACTTCGCTCT           |
| SRG1       | TF1946                  | TGGTCCCTCCTTCCAAACAA      |
| SER3 mRNA  | TF1947                  | TTGCCCAGGAAGAGATTTGAT    |
| SER3 mRNA  | TF1948                  | TGAGTTGGAGAAAGGCGAGT     |
| TEF4       | 5' exon                 | TF2019    | AATAGATCTCCAAGAAACTACGCT       |
|            | TF2020                  | CTTTAAGCCCTTTTGGACCCA   |
| 5' exon - intron junction | TF1918 | GCAGATTTTTGCTTCTTGTCTCCCA |
| 5' exon - intron junction | TF1919 | CGCGCACAACAATGTATGTGGAG       |
| intron     | TF1978                  | TTCGTTTATTTTGGACAGTCA   |
| intron     | TF1979                  | CCTGATGTTAGGACAGCTTT    |
| intron - 3' exon junction | TF1920 | GTCAAAGCGTGCTCCTAATCCAGG     |
| intron - 3' exon junction | TF1921 | GTGCGGAACCTAATAACGGAGCT     |
| 3' exon    | TF1931                  | GCTAATTCCGATGTCCAGACCA  |
| 3' exon    | TF1932                  | TGTTTCGAGTCAACAAAGGCT   |
| URA2       | 5' exon                 | TF1887    | GGTTCCTGGAAAGATTTTTCTCTACGG  |
|            | TF1888                  | CAATTCAGAGTTGAACGTGTGAG |
| 5' exon - intron junction | TF1880 | AGACTCACCAGGCTTCAACTTG     |
| 5' exon - intron junction | TF1881 | GCCATCGCATGGTCTGAGATAG     |
| intron     | TF2008                  | CTATTCAGACCTAGTGCCG     |
| intron     | TF1981                  | TGGTATCTTCTTACAGAGAT    |
| intron - 3' exon junction | TF1937 | CATGCTTGACGCTAACA          |
| intron - 3' exon junction | TF1938 | TGGGACAGTTGGAGGACAATA       |
| 3' exon    | TF1882                  | TATGCTCCACTGCTCCAA      |
| 3' exon    | TF1883                  | ACTAATTCCACCAGGCAGG     |
| URA8       | TSS1                    | TF1955    | CCCCTGCTTTTCTCTTTTATCT       |
|            | TF1956                  | ACCACCAGAAACAACACAGT   |
| mRNA       | TF1957       | GTTTCACTTGTGCCCGTCAT |
|-----------|-------------|---------------------|
| mRNA       | TF1958       | TGAACGTGGACCACCTTGCTC |
| YPR063C    |              |                     |
| 5' exon - intron junction | TF1889     | GAATAACGTGGCAAGGCCC |
| 5' exon - intron junction | TF1890     | GCGGTTGAACCTGTTAGTATATATTGGC |
| intron - 3' exon junction | TF1982     | GCCAATATATACCTAACGTTCAACC |
| intron - 3' exon junction | TF1983     | GTCTTCGTCGGTGTTCAAGGTAGT |
| 3' exon    | TF1891       | ACTACCTGAACCCGACGAAGAC |
| 3' exon    | TF1892       | ACGATGGTTGAGCTATTGTTGAGG |
Supplemental Figures and legends:

**Figure S1**

S1A) Context of Spt6:RNAPII interactions in the elongation complex (similar to Fig 1A).

The structure of the human RNAPII EC (5) in the left panel shows the complex transcribing towards the viewer, and in the right panel transcription is to the right. Red, blue, and green circles indicate the multiple contacts made between the Spt6 core domain and the nascent RNA, Spt5, and the Rpb4/7 knob, respectively, that were detected by cryo-EM.

S1B) Spot dilutions demonstrating phenotypes of individual and combined mutations in the linker region of *RPB1* that is bound by Spt6-tSH2.

| Strain | SC Glc 30° 2d | SC-lys Glc 30° 2d | SC-his Glc 30° 5d | SC Gal 30° 5d | SC-his Gal 30° 5d | YPA Glc 30° 2d | YPA Glc HU 30° 2d |
|--------|---------------|-------------------|-------------------|--------------|-------------------|-----------------|------------------|
| RPB1 (WT) | ![Image](image1) | ![Image](image2) | ![Image](image3) | ![Image](image4) | ![Image](image5) | ![Image](image6) | ![Image](image7) |
| rpb1T   | ![Image](image8) | ![Image](image9) | ![Image](image10) | ![Image](image11) | ![Image](image12) | ![Image](image13) | ![Image](image14) |
| rpb1Y   | ![Image](image15) | ![Image](image16) | ![Image](image17) | ![Image](image18) | ![Image](image19) | ![Image](image20) | ![Image](image21) |
| rpb1TPY | ![Image](image22) | ![Image](image23) | ![Image](image24) | ![Image](image25) | ![Image](image26) | ![Image](image27) | ![Image](image28) |
| rpb1FSP | ![Image](image29) | ![Image](image30) | ![Image](image31) | ![Image](image32) | ![Image](image33) | ![Image](image34) | ![Image](image35) |
| rpb1TFSP| ![Image](image36) | ![Image](image37) | ![Image](image38) | ![Image](image39) | ![Image](image40) | ![Image](image41) | ![Image](image42) |

Strains with the genotypes shown (Table S1) were grown to saturation in rich medium, washed in water, then serially diluted in sterile water in 10-fold steps. Aliquots were placed on the media shown and incubated as indicated in each panel. SC is synthetic medium, with all nutrients required (complete) or lacking lysine (-lys) or histidine (-his). Sugars provided were glucose (Glc) or galactose (Gal). YPA is rich medium, HU is 90 mM hydroxyurea. A similar analysis of the *SPT6* alleles was published previously (1).

The strains contain the lys2-128Δ allele whose expression (growth on media lacking lysine) indicates the Spt- phenotype (6). They also contain a Gal1p-FLO8-HIS3 reporter, whose expression (growth on media lacking histidine) indicates activation of a cryptic promoter within the *FLO8* gene (7).
Growth on SC-his Glc indicates strong activation of this reporter, as the Gal1p promoter is repressed on glucose, and growth on SC-his Gal indicates weaker activation that is dependent on transcription of the reporter gene at high frequency to assist in activating the cryptic promoter. All strains with the "FSP" mutation set (Table S1; disruption of the interaction with the C-terminal Spt6 SH2 motif) displayed both the Spt- phenotype and cryptic promoter activation on galactose, with stronger cryptic promoter activation in compound mutations that included T1471A. Sensitivity to the DNA synthesis inhibitor hydroxyurea was observed with some alleles of SPT6 (1), and is also seen with compound alleles of RPB1 that include the T1471A mutation. These results suggest that mutating the phosphorylated residue T1471 has stronger effects for some phenotypes than mutating the accessory residue Y1473, but these compound alleles produced similar changes in Spt6 occupancy as one another in a ChIP-seq analysis (1 and Fig 7C,G-I).

S1C) Principal component analysis of the three RNA-seq samples for each genotype.

RNA-seq results from each individual strain were subjected to principal component analysis, producing the clustering shown here with the two major components. Note that while the rpb1TPY allele did not alter the level of any transcripts using the criteria in Fig 1B or S1G, this genotype still separated from WT in both components of the PCA and produced significant differences relative to WT in some of the qPCR tests. In general, simple mutations affecting only one region of the interface clustered together, and compound alleles were separate from this but similar to one another.

S1D) Comparison of rlog2 (left panel) and log2FC values (right panel) for the transcript level from each gene for each genotype and the Pearson correlation coefficients for the mutants relative to WT. Global changes in transcript levels caused by the interface mutations were small, with strong effects occurring at only a small subset of loci suggesting local roles for the Spt6-tSH2:Rpb1 interface. In the right panel, the log2FC values relative to WT for each genotype compared to those from spt6R,KK, with the derived correlation values (Pearson r) shown in Figure 1D. The same genes showed similar effects with all component interface mutations, but the magnitude of the effects was larger in compound mutants. Some genes displayed additive effects of individual mutations, but others like IMD2 were affected similarly by compound and most component mutations.
S1E) Effects of the interface mutations on RNA levels correlated moderately with published values for full deletion of the Spt6-tSH2 domain.

The published log₂FC values from RNA-seq data in reference 11 for genes where an effect greater than 2-fold was observed upon deletion of the tSH2 domain of Spt6 are compared to the effects of the *spt6*<sup>R,KK</sup> allele from this study. The overall Pearson correlation was 0.53, and the full deletion had a larger activation of the heat shock response but less effect on IMD2 than the interface mutations did.

S1F) Interface mutations caused a significant growth defect, but less than that caused by full deletion of the Spt6-tSH2 domain.

6 independent clones of each genotype (Table S1) with 2 technical replicates each (the same initial cultures used to inoculate two wells at different initial dilutions) were grown in rich medium at 30°. Growth was monitored by absorbance at 600 nm using a Bioscreen C (Oy Growth Curves Ab Ltd). The doubling time was calculated using the absorbance readings between 0.25 and 0.8. Plots show the mean doubling time in minutes for each culture with the standard error of measurement. Technical replicates are indicated with different shading. Unpaired t-test values were less than 10⁻⁸ for all pairwise comparisons as indicated (** ** = P < 10⁻⁴ shown for the comparison with WT). The loss of the Spt6-tSH2:Rpb1 interaction caused a significant delay, but did not fully recapitulate the growth defect caused by full deletion of the tSH2 domain, suggesting this domain has functions in addition to providing binding sites for Rpb1.

S1G) As in Fig 1B, except using less stringent but more typical cutoffs (a 1.5-fold change in transcripts at a false discovery rate less than 5%). Gene Ontology (GO) term analysis of genes with increased transcripts in the *spt6*<sup>R,KK</sup> strain gave strongest association with the biological process category "small molecule metabolic process" (71 hits out of 264 genes found at the 2-fold increase, 1% FDR cutoff, and 152 of 707 for 1.5-fold and 5% FDR; P = 1 X 10⁻⁸ or 1 X 10⁻¹², respectively). In contrast, downregulated genes were associated with "cytoplasmic translation" (64/107 genes, P = 6 X 10⁻⁶⁶, or 143/483 genes P = 3 X 10⁻¹⁵).
S1H) Comparisons of changes in transcript levels in interface mutants with total transcripts (left panel), the synthesis and destruction rates from cDTA (8, middle panel), and stress profile values (9, right panel). Pearson r values are given for each comparison, and a subset of heat shock genes and highly affected genes are indicated for the stress profile panel.

RNA-seq measures the steady-state level of transcripts, which is determined by the balance between the frequency of transcription and the rate of turnover. If interface mutations altered either of these processes uniformly for all genes, the changes in transcript levels would be expected to correlate with the basal synthesis or destruction rates (for example, a decrease in the global rate of transcript turnover should cause a greater increase in the level of a rapidly degraded, unstable transcript than for a stable transcript). We compared log_{2}FC values for mutants to published values for destruction rates (8) but did not observe a significant correlation (for example, Pearson r = 0.03 for spt6^{R,KK}, middle panel). A similar test with either the synthesis rate (frequency of transcription, middle panel, Pearson r = -0.035) or the total transcript abundance produced a moderate negative correlation (left panel, Pearson r = -0.34). This indicates that genes that are transcribed more often (typically resulting in more abundant transcripts) were somewhat more likely to display decreased transcript abundance in the interface mutants, revealing a modest, general defect in maintaining high levels of transcription. This is consistent with a role for Spt6-tSH2:Rpb1 interaction in sustaining maximal output from frequently transcribed genes, as expected from previous results (10).

Many forms of cellular stress, including defects in transcription, produce a characteristic set of changes in transcript levels, which has been designated as the "stress profile" (9). The log_{2}FC values for interface mutants correlated moderately with this profile (Pearson r = 0.49; right panel), indicating that interface mutations caused some general stress, presumably transcriptional stress. As discussed in the main text, this is likely to at least partially explain the reduced abundance of ribosomal protein gene transcripts. However, transcripts from other components of the stress response were not affected, such as the expected increase for heat shock genes, and many of the transcripts that were most significantly increased in our mutants (for example, BIO2 and IMD2) have not previously been observed to be part of the stress profile. Disruption of the Spt6-tSH2:Rpb1 interface therefore appears to have reduced the maximal output from the most frequently transcribed genes, and to have induced some changes associated with general transcription stress, but it also increased the levels of many specific transcripts that are not in these categories. This suggests that the interface has both a general role in maintaining high-frequency transcription and more specific functions at subsets of individual genes.
Figure S2

S2A) Data compiled from the Saccharomyces Genome Database (12) for the nucleotide biosynthetic pathways. Genes whose expression was analyzed in this study are indicated in red. Mycophenolic acid (MPA) inhibits the activities of IMD3p and IMD4p but has less effect on IMD2p, and exposure to MPA leads to induction of IMD2 transcription (13). IMD1 is homologous to IMD2-4 at the DNA sequence level, but contains a null mutation and therefore does not contribute to IMDH activity but is difficult to distinguish from IMD2 in RNA-seq or qPCR data.

S2B) Primer extension at ADH1.
RNA was extracted from strains with the genotypes shown and a primer complementary to a site near the beginning of the ADH1 ORF (Table S2) was added and extended with reverse transcriptase (RT) as described in Methods. Transcription of this gene initiates from either of two sites with similar frequencies, producing 120 (TSS1) and 110 (TSS2) bp products with this primer (14). The locations and relative frequencies of usage were not affected dramatically by the interface mutants tested (of the total signal in the TSS1 and TSS2 bands shown here, TSS1 usage was 48% for WT, 52% for spt6<sup>R,KK</sup>, and 50% for rpb1<sup>TPY,FSP</sup>). We also used a qPCR assay based on the use of one forward primer that annealed to products initiated precisely at TSS1 (25 bases annealed, T<sub>m</sub> 56° for TSS1 products, 15 bases annealed, T<sub>m</sub> 42° for TSS2 products) compared with a downstream primer that annealed to products from either TSS1 or TSS2 (Table S2). This allowed quantitation of 6 independent isolates of the same strains shown here and produced average values of TSS1 usage ranging from 72% (rpb1<sup>TPY,FSP</sup>) to 78% (spt6<sup>R,KK</sup>) but these were not statistically different from the WT value of 75% in a t test (P = 0.13 and 0.18, respectively. As discussed in the main text, this does not rule out a global reduction in TSS selection efficiency which would only significantly affect transcript levels from genes with multiple TSSs used at different frequencies.
S2C-E) NNS-dependent termination efficiency at *NRD1* was not affected by interface mutations. (For all t-tests, ns = not significant, $P > 0.05$, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$).

S2C) A map of the *NRD1* locus drawn to scale shows the qPCR targets used to measure the levels of the upstream CUT320, the 5' products initiated from the single TSS for *NRD1*, and the 3' products indicating completion of an intact mRNA. Several sites that promote NNS binding and termination of transcription have been mapped to the region as shown, with additive effects leading to termination of about half of the initiated transcripts (15).

S2D) 5' and 3' signals were measured using the same samples as in Fig 2, with the same statistical analysis. Interface mutations caused increased levels of both 5' and 3' products but the completion ratio was unaffected or increased only slightly in compound mutants (calculated as 3' signal/5' signal with paired t-test results shown). The efficiency of termination therefore may have dropped slightly in the *spt6*<sup>RC, KK</sup> mutant (reduced termination efficiency gives higher completion ratios), but not enough to account for the 2-fold elevation in mRNA.

S2E) Aborted transcripts are degraded by the nuclear exosome, leading to an increase in signal in an *rrp6*-Δ mutant, as expected. However, the 3' signal also increased in cells lacking exosomes, although to a lower extent, resulting in a drop in the completion ratio to about 20% of the initiated transcripts. Loss of exosomes results in a global sequestering of NNS complexes by undegraded ncRNA (16), so it is difficult to determine how the new balance in *NRD1* mRNA production is achieved (decreased NNS availability should lead to increased completion, but a global signal to produce more NNS apparently caused increased initiation as well, resulting in an overall reduction in completion), but interface mutations did not alter the completion ratio dramatically in either *RRP6* or *rrp6*-Δ backgrounds, suggesting they did not affect NNS-dependent termination directly.
S2F) Profiles of RNA-seq reads suggest normal global initiation and termination. RNA-seq reads were aligned to annotated initiation (TSS) and termination (TTS) sites for all genes and the average RPM plotted against the distance from the feature. Normal onset and decay of signal profiles suggest normal global initiation and termination site usage.

S2G,H) To test for production of a CUT from the *IMD3* promoter, primers were designed to target the region upstream of the annotated TSS (that is also observed in the TSS-seq data in Fig S3A from reference 17), and qPCR was performed as in Fig 2 using the same samples and analysis. The signal from this region was low in normal cells but increased 132-fold in an *rrp6* strain, strongly indicating the
presence of an unannotated CUT. This promoter region is not homologous to \textit{IMD2}, the initiation site for the CUT is unknown, \textit{IMD3} is not known to be regulated by a mechanism similar to the TSS1-bypass mechanism used at \textit{IMD2}, and neither the CUT nor the \textit{IMD3} mRNA were induced by 6AU (not shown). However, as at \textit{IMD2}, the level of \textit{IMD3} CUT product increased 5-10-fold in interface mutants, this time with a smaller effect on mRNA. This supports the model that coordination of CUT degradation with transcription initiation is diminished in interface mutants, in this case without strong consequences for the ability to use the downstream TSSs. The function of the transcription of CUTs impinging on promoters is unknown but seems more common in the subset of genes affected by interface mutants than expected.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{Scale map of the \textit{IMD2} locus (same as Fig 2A).}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{qPCR measurements of the responses of strains with the genotypes indicated (Table S1) to 75 \textmu{g}/ml 6-azauracil (6AU) in synthetic medium lacking uracil for 2 hours or 5 \textmu{g}/ml mycophenolic acid (MPA) in rich medium for 1 hour using the same strains as in Fig 2E (the "no MPA" panes here are the same data as in Fig 2E, but scaled to facilitate comparison with the +MPA series performed at the same time as the results in Fig 2E but not included in that figure). Growth in synthetic medium reduced the level of TSS1-CUT (0.09 in synthetic medium vs 0.13 in rich medium here) and increased the level of mRNA (0.13 in synthetic medium vs 0.06 in rich medium here) in WT strains. The ~4-fold increases in \textit{IMD2} TSS-CUT and 10-fold increases in \textit{IMD2} mRNA seen in other experiments using rich medium were therefore 5-6 fold increases for \textit{IMD2} TSS-CUT and 2-3 fold for mRNA in panel J for strains grown in synthetic medium and 5-fold and 12-fold, respectively, for spt6\textsuperscript{R,KK} in panel K for strains grown in rich medium. 6AU caused a 2.6-fold drop in TSS1-CUT product and an 11-fold increase in mRNA, while MPA caused a 30% reduction in TSS1-CUT and a 110-fold increase in mRNA, so \textit{IMD2} induction was more sensitive to MPA than to 6AU. Unpaired t-test values are shown for comparison of mutants to WT and paired t-tests for comparisons of the effect of 6AU or MPA (ns = not significant, \textbf{P} > 0.05, \textit{*} = \textbf{P} < 0.05, \textbf{**} = \textbf{P} < 0.01, \textbf{***} = \textbf{P} < 0.001, \textbf{****} = \textbf{P} < 0.0001).}
Figure S3
S3A Browser tracks are shown for genes described in the main text. Bigwig files were displayed in IGV (18) with the scale shown for each set of tracks. RNA set 1 is from this report, using paired-end sequencing with strains from the A364a genetic background (Table S1), and RNA set 2 is from reference 4 using strains from the W303 background and single-end sequencing, which did not provide as much coverage near the ends of transcripts as paired-end sequencing, but produced strong correlations with total mRNA values for each gene (Pearson r = 0.93 for the WT strains). Spt6 occupancy is from reference 1, MNase-seq from this report, and TSS-seq tracks are from reference 17. WT is shown in gray as a separate track for RNA signals and overlaid for the Spt6 occupancy and nucleosome tracks.

**ADE12**
RNA changes were small for this locus in most mutants, but interface mutations caused a large change in the Spt6 occupancy over the 5' end, without evidence of a CUT being produced in this region, and accompanied by a shift towards use of TSS2 (Fig S3B). CUT324-CUT325 appear to impinge on this promoter but no change in TSS1 product was seen in an *rrp6*-Δ mutant so the primers used did not appear to detect these transcripts.

**BIO2**
Spt6 occupancy formed a peak over the promoter in WT cells, but this was replaced in interface mutants by a continual gradient of accumulation instead of the reductions seen at many other loci with multiple TSSs, consistent with strong upregulation of the mRNA. Interface mutants also displayed a shift towards use of TSS2. This was not associated with a change in nucleosome occupancy in the promoter or gene body, although the NDR for this gene was upstream of the -1 nucleosome, and occupancy of this site did decrease in the interface mutants.
**IMD2**
This gene provided the initial example of an accumulation of Spt6 occupancy over the promoter greater than in flanking regions, that decreased or shifted downstream in interface mutants while transcripts from the upstream TSS1-CUT region remained high. The regions losing Spt6 occupancy also lost nucleosome occupancy, although this did not extend into the gene body where transcript levels increased 10-fold.

**IMD3**
As discussed in Fig S2G, this gene does not share homology with IMD2 upstream of the TSS, or have the same regulatory mechanisms, but it displayed a peak of Spt6 occupancy over the site of unannotated CUT production, this occupancy diminished in interface mutants while the CUT level remained high, and had little effect on mRNA production. This is consistent with decreased efficiency of CUT degradation in interface mutants, possibly due to reduced coordination between termination and exosome activity in the region where Spt6 occupancy was lost, with smaller changes in CUT and mRNA production.

**IMD4**
Spt6 accumulation shifted downstream at the 5’ end in interface mutants, no change occurred over the intron, and some loss of nucleosome occupancy was seen over the promoter without a change in mRNA.
**MET3**

Two clusters of initiation sites are observed, the TSS1 product was stabilized somewhat in an exosome mutant, and interface mutants caused a shift towards use of TSS2 coincident with a decrease in the peak of accumulation of Spt6. No significant changes in nucleosome occupancy were observed.

**NRD1**

An accumulation of Spt6 is seen over the 5' end, extending into the region where several NNS-dependent termination sites are found. The accumulation diminishes and shifts downstream in interface mutants, but termination efficiency appeared to be unaffected (Fig S2C-E). The increase in mRNA therefore may result from activation of the promoter, consistent with diminished nucleosome occupancy over the +1 site in interface mutants.

**RPL17A**

Spt6 occupancy was reduced over this RP gene in interface mutants, but primarily over the promoter, 5' exon, and intron. Loss of the +1 nucleosome is evident in these mutants, but not in a FACT mutant.
**RPL18A**
This RP gene shows features similar to *RPL17A* above. In this case the loss of Spt6 accumulation over the gene body is more pronounced, but the level of *RPL18A* mRNA was relatively unaffected in these mutants (Fig 4C).

**RPL23A**
Interface mutations caused few changes in the profiles at this gene and did not affect 3' exon levels (Fig 4D). Intron retention increased, and the gradient of onset of Spt6 occupancy seen in averaged genes is pronounced at this locus with a slight further decrease over the TSS in the mutants.

**SRG1-SER3**
WT cells have a notable peak of Spt6 occupancy centered over the junction where *SRG1* transcription interferes with activation of the *SER3* promoter, and this decreased or shifted downstream in interface mutants. Nucleosome occupancy was diminished in this region in the interface mutants, but not as evidently in a FACT mutant, although both caused similar increases in the level of mRNA.

---

*RNA set 1 (Forward, 0-661)*

*RNA set 2 (Forward, 0-583)*

*MNase-Seq (6-19)*

*TSS-Seq (0-846)*

---

*RNA set 1 (Forward, 0-299)*

*RNA set 2 (Forward, 0-323)*

*MNase-Seq (6-19)*

*TSS-Seq (0-222)*

---

*RNA set 1 (Forward, 0-138)*

*RNA set 2 (Forward, 0-45)*

*MNase-Seq (6-25)*

*TSS-Seq (0-76)*

---
**TEF4**
Spt6 accumulation was diminished over this locus in interface mutants, similar to the decrease in transcript levels. The accumulation of RNA signal over the 5' side of the intron is evident in the interface mutants but not WT or other mutants.

**URA2** (promoter region)
The accumulation of Spt6 over the promoter of this gene decreased in the interface mutants, without significant changes in the mRNA level (Fig S4I). The site of an unannotated TSS2 within the intron is shown with accumulation of products that appear to be initiated at this site being evident even in the WT strain.

**URA2** (full gene)
As above, highlighting the decreased Spt6 occupancy across the gene in interface mutants without a change in mRNA level, and the reduction in the accumulation of Spt6 over the promoter in the mutants.
**URA8**

Results in an *rrp6-Δ* strain indicate an unannotated CUT over the 5’ end of this gene (Fig S3D). Two TSS clusters are evident here, as is the accumulation of Spt6 over the 5’ end in WT cells and the decrease or shift downstream in the interface mutants. This locus also illustrates the accumulation of Spt6 over the 3’ end as a result of overlap with the converging, highly transcribed gene *SOD1*, as observed in averaged gene profiles (Fig 7E). Nucleosome occupancy was only slightly decreased over the gene body, but was diminished at the -1 site for the primary initiation site (TSS2) in the interface mutants, coincident with a shift to the minor upstream site (TSS1) that is normally occluded by this nucleosome.
Genes with multiple TSS clusters respond in distinct ways to interface mutations and loss of exosome function. Strains with interface mutations and either normal RRP6 or rrp6-∆ were tested by qPCR for levels of the TSS1 and mRNA products as in Fig 2C. The increase in TSS1 signal upon deletion of RRP6 and a TSS1/mRNA ratio above 1 suggest that a significant fraction of the TSS1 signal is from unstable transcripts that are not extended to full-length mRNA. This was observed to different extents for BIO2, MET3, and URA8, but not for ADE12. Interface mutations caused an increase in TSS1 signal for BIO2 and URA8 in RRP6 cells, but not in rrp6-∆ cells, possibly indicating that the interface has a role in coordinating the activity of the exosome with transcription, but this effect is dependent on the exosome. Loss of the exosome by itself also increased the levels of the BIO2, MET3, and URA8 mRNAs, but combining these mutations either had little effect (BIO2), or blunted the increases relative to single mutants (MET3 and URA8). Although ADE12 did not appear to have an unstable product and individual mutations had little effect, combining the mutations also produced an unexplained reduction in mRNA level. The ratio of TSS1/mRNA revealed the same apparent shift towards or away from TSS2 usage noted in Fig 3 with or without RRP6. Unpaired t-test values are shown for comparisons of WT, spt6R,KK or rpb1T,FSF strains with and without RRP6 (top row, "rrp6-∆"), and interface mutants compared with the WT in the same background (RRP6 or rrp6-∆; bottom row, "Interface"). For all t-tests in all figures, ns = not significant for P > 0.05, * = P < 0.05, ** = P < 0.01, *** = P < 0.001, **** = P < 0.0001.
S3F-I: Initiation site shifts were similar in $spt6^{R, KK}$ and $rpb1^{TPY, FSP}$ strains.

The same four targets with multiple TSSs were analyzed as in Fig 3, but using independent biological replicates and an additional $spt6$ allele and two FACT mutants (Table S1). The ratio of signals initiated from TSS1 to the total mRNA was calculated for each pair of samples from the same strain as an estimate of the use of each potential start site, with decreases in the ratio suggesting a shift towards use of initiation sites downstream of TSS1. As discussed in the main text, the results show that all of the mutations generally caused altered TSS selection, but each gene had a different profile of effects consistent with dynamic, local decisions being made at each locus with distinct contributions from each factor being of variable importance in each context. Unpaired t-test values are shown for mutants compared to WT (ns = not significant for $P > 0.05$, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$).
S3J) Analysis of IMD2-IMD3 expression in \textit{spt6}^{F249K} and FACT mutants.

Samples and analysis were as in Fig S3B-E. An interface mutant had a stronger impact on the levels of TSS1-CUT (IMD2) and upstream CUT (IMD3) than the other mutations tested, but effects on mRNA were more similar to one another. The apparent stabilization of CUTs impinging on promoters/downstream TSSs therefore seems to be more prominently associated with interface mutations than with loss of Spn1 interaction or defects in histone chaperone activity.
Fig S3K) spt6-F249K produced distinct effects relative to interface mutations at some loci.

6 independent isolates of strains with the genotypes shown (Table S1) were grown to logarithmic phase (OD ~0.8) in the media indicated, then RNA was extracted and tested by qPCR using the primers indicated (Table S2). YPA and YM-1 are rich media with 2% glucose, but YPA has 10 g/L yeast extract, 20 g/L peptone, and 50 mg/L adenine, while YM-1 has half these amounts of yeast extract and peptone, 36 mg/L of adenine and uracil, 1.5 g/L yeast nitrogen base without amino acids or ammonium sulfate/L, 5 g/L of ammonium sulfate, and 10 g/L of succinic acid, adjusted to pH 5.7 with NaOH. Synthetic medium is 1.7 g/L of yeast nitrogen base without amino acids or ammonium sulfate, 5 g/L of ammonium sulfate, and 10 g/L of succinic acid, adjusted to pH 5.7 with NaOH and supplemented with 14 mg/L each of uracil, tryptophan, and histidine, 43 mg/L of lysine, and 57 mg/L of leucine (the minimal requirements for the strain background used here).

Under varying rich medium conditions, transcript levels of ADE17 (whose product catalyzes the production of the IMP that is the substrate of the IMDHs encoded by IMD2, IMD3, and IMD4 as shown in Fig S2A) were unaffected or slightly increased by interface or FACT mutations, but increased 2-3 fold in an spt6F249K strain. (Strong, uniform induction of ADE17 was observed in all strains grown in synthetic medium lacking adenine supplementation; not shown.) In contrast, spt6F249K and FACT mutations had little or no effect on NRD1 transcript levels but these increased 2-3 fold in interface mutants under different media conditions. Other loci including ARO9 and CHA1 were strongly influenced by media conditions but in YPA (with its higher levels of peptone that appear to produce higher transcript levels from these loci relative to YM-1), FACT mutations had little effect while spt6F249K and spt6R,KK had significant but opposite effects on transcript levels. Some of the effects of the interface mutations may therefore be attributed to failed recruitment of the appropriate level of the NTD of Spt6 to transcription complexes, but others appear to be independent of the multiple functions of the Spt6:Spn1 interaction that is disrupted by the spt6F249K mutation.
Figure S4

Analysis of splicing efficiency in interface mutants.

S4A) RNA-seq data were used for an initial screen of intron retention by plotting the log₂FC in intron region signals against the log₂FC values for total mRNA in interface mutants (spt6<sup>R,KK</sup> shown, similar results were obtained with rpb1<sup>TPY,FSP</sup>). Excess intron retention should be seen as an increase on the y axis relative to the change in mRNA on the x axis, so the distance above the diagonal should indicate the level of splicing defect. By this analysis, defects are widespread but affected some genes more than others, with small or no effect at RPL17A and a large effect at RPL23A. S4B) As an alternative approach, the total intron signal for each gene was divided by the total exon signal and values for all interface mutants were calculated and normalized to the value for the WT strain. Values above 1 therefore indicate increased intron retention in the mutant strain, again producing a distribution but a general increase in retention. The Pearson correlation for each distribution relative to spt6<sup>R,KK</sup> is given, indicating similar effects at the same loci for most mutants. TEF4 gave consistently high retention scores, as noted. More direct tests revealed a more global splicing defect in interface mutants, suggesting that the analysis of RNA-seq data using small numbers of reads with only three biological replicates was only partially accurate.

S4C,D) Intron-containing genes displayed decreased transcript abundance overall.

As another test of the possibility that interface mutants had a defect in processing introns, genes were parsed into four groups reflecting the presence or absence of an intron and whether or not they encode ribosomal proteins (RPs). The log₂FC in mRNA relative to WT is plotted here. The decrease in mRNA tracked primarily with RPs, but a small but significant decrease was also noted for intron-containing genes in both spt6<sup>R,KK</sup> (C) and rpb1<sup>TPY,FSP</sup> strains (D). Unpaired t-test results are shown for all pairwise comparisons, with symbols as defined in Fig 2.
S4E-G) Intron retention was calculated as in Fig 4B-D, except with genes that do not encode RP s, and whose mRNA levels were unaffected or increased in interface mutants. The RNA-seq analysis shown in S4A,B suggested high intron retention for TEF4, less for IMD4, and little effect at YPR063C, but this quantitative assay with multiple independent biological repeats revealed a more similar and widespread increase in intron retention. We conclude that the low levels of intron signal available from our RNA-seq data were inadequate for quantitative analysis of intron retention. Results of unpaired t-tests are shown, with symbols as in Fig 2.

S4H) A map of the URA2 locus showing the annotated TSS1 and TSS2 sites (19,20), whose variable usage regulates the level of the main primary transcript containing an intron within the 5’ untranslated region (UTR). The mapped TATA/TFIIB site and site of an unannotated TSS from TSS-seq data (17) are indicated as TSS3 here, along with the putative resulting mRNA. Use of TSS1 therefore leads to formation of the unstable TSS1-CUT, use of TSS2 leads to formation of the primary transcript that requires splicing to produce an mRNA, and use of the putative TSS3 leads to formation of a potentially functional mRNA that does not require splicing. TSS1 produced little signal in TSS-seq (see Fig S3A) so other figures note only the main promoter site and the putative site within the intron.

S4I) qPCR measurements from various regions of the URA2. The levels of 5’ and 3’ exon were similar to one another in the WT, but the level of 3’ exon increased in the mutants without a similar change in 5’ exon level, suggesting increased use of TSS3 (also see Fig S3A). As with other intron-containing genes, the level of retained intron also increased in the compound interface mutants (intron/5’ exon) but intron signal was constant relative to the 3’ exon, as expected if the increased mRNA level was primarily due to the use of the putative TSS3 which does not contain the intron. Paired t-test values are shown for mutants compared to WT with symbols as in Fig 2.
Figure S5
Intron retention gradient analysis in additional strains.

S5A-D) As in Figure 5, but with an independent set of biological replicates grown in rich medium to further validate the unexpected gradient of signals at IMD4 (A) and TEF4 (B), and extending the analysis to RPL17A (C) and RPL18A (D). The results at the loci with snoRNA genes within the introns were similar to one another and to the tests in Figure 5, whereas the results with RPL17A and RPL18A indicated a defect in splicing itself, with retention of 3' junction signal relative to the 5' junction that is disrupted in the first step of splicing and therefore undetected by PCR. However, the strong skewing towards the 3' junction for RPL18A may also reflect some cryptic promoter activation within the intron, as suggested by the results shown below in Fig S5E. Results of unpaired t-tests (mutants compared to WT) or paired t-tests (targets compared within a genotype) are shown, with the symbols as in Fig 5.
S5E) Browser tracks with the RNA-seq signals from intron-containing genes with WT levels in gray overlaid on the two mutant tracks, and the signal range adjusted to allow the low levels of reads from the intron regions to be displayed. As indicated by the qPCR analysis of intron retention gradients above, transcript signal is skewed towards the 5’ end of the intron for IMD4 and TEF4, to the 3’ end for RPL17A, RPL18A, and URA2, and is uniform across YPR063C. The strong skewing to the 3’ end of the introns in these cases may be due to proposed cryptic initiation within the intron of RPL18A, and previously observed (17) but unannotated initiation within the intron of URA2, at the sites noted (proposed TSS3 in Fig S4H). The arrows for the URA2 track are proportional to the amount of TSS-seq signal noted at these sites in the WT (as shown in Fig S3A).
As above, but with an additional allele of \textit{spt6} and two FACT mutants, as in Fig S3B-E, with two normalization strategies as in Fig 5, and extending the gradient analysis to \textit{URA2}. As noted above, intron retention at this locus is difficult to measure due to the large amount of signal arising from an unannotated TSS within the intron, producing high variance.

**F** Intron Signal Normalized to Exon Signal

**G** Each Intron Region Normalized to WT Avg

**H**
As in Figure 6, except 6 biological replicates of multiple strains were grown in synthetic medium without or with 75 µg/ml 6AU for 2 hours to examine the effect of combining 6AU with interface mutations on intron retention gradients at IMD4 (top panels) and TEF4 (bottom panels). Growth in synthetic medium affected the absolute signal ranges for these genes relative to cultures grown in rich medium (exons, left panels), but 6AU caused significant drops in mRNA levels for all conditions. Skewing towards 5' intron junction retention was still visible in the interface mutants alone (middle panel, normalized to the WT average for each region), and in the WT treated with 6AU (right panel, gray bars, normalized to the untreated WT average for each region), with strong additivity in the mutants treated with 6AU (right panel, maroon and green bars). Paired t-tests were performed for the exons comparing the effects of 6AU (top of the panel) unpaired tests for the effects of the mutations relative to the same culture condition (within bars). For the intron regions, comparisons were made between the 5' junction and 3' junction (above the panel), for each region relative to the same region in WT (top of the panel) and for the effect of 6AU (below the panel). Symbols are as in Fig 2.
Figure S7

Illustrations and analysis of the peak of Spt6 accumulation observed at a subset of genes.

S7A-C) Browser tracks as in Fig 7 for additional genes displaying the features associated with distinct promoters, with annotations to point out the relevant changes in each case. The rpb1^{TPY,FSP} strain recapitulated phenotypes of rpb1^{TPY,FSP} more closely than rpb1^{Y,FSP} (Fig S1B), so it is shown here.
S7D-F) Unbiased detection of the subset of genes with the “5’ hump” configuration of Spt6 accumulation.

Genes like IMD2 displayed a peak of Spt6 occupancy over the 5’ end near the TSS clusters as shown in Fig 7 and Fig S3A. To find other genes with this feature, we calculated the Spt6 occupancy over the 200 bp upstream of each gene, over the first 5% of averaged genes, and over the middle 70% of averaged genes (15% to 85%, thereby avoiding effects of neighboring genes as described for RNA-seq data in Methods). We then ranked the 5885 genes with adequate signal in all three regions by the difference between the 0-5% value and the 15-85% value, indicating higher Spt6 occupancy at the 5’ end of the gene than in the middle. The top 5% of genes by this ranking (295 genes) were then ordered by the difference between the 0-5% value and the -200 bp to TSS value, identifying those genes with higher Spt6 occupancy over the 5’ end of the gene than over the region upstream. This separated the top 95% of genes by the first criterion (5’ > middle) into 5 percentiles of 59 genes each by the second criterion (5’ > upstream).

S7D) The Spt6 occupancy profile for all genes is shown in gray as in Fig 7E, and the profiles for each of the top 4 percentile groups calculated as described above are shown in colors. The 99th percentile shows a strong peak but this is less prominent in the lower groups, suggesting they contain some spillover of occupancy from the upstream genes rather than a true peak of accumulation. We estimate that the pattern of true 5’ accumulation is therefore present in about 3% of genes.

S7E) Only the Spt6 occupancy in WT cells was used to identify the genes with 5’ peaks of Spt6 accumulation, but genes with this feature also displayed a characteristic decrease or downstream shift of this peak (Fig 7, S7A-C and S3A). Plotting the log_2 FC in Spt6 occupancy profile for the same gene sets in the spt6<sup>R,KK</sup> mutant showed that loss of accumulation tracked with the genes having 5’ occupancy peaks, roughly in proportion to the strength of the peak. The 5’ accumulation therefore appears to be particularly dependent on the integrity of the interface.

S7F) The genes indicated with red stars formed the conceptual training set for this unbiased analysis but were otherwise not used in the calculations. It was therefore striking that all of the genes except for IMD2 came out in the top percentile of genes. IMD2 is instructive because the analysis used annotated TSSs, which is TSS1 for this gene, and this is the initiation site for the regulatory CUT, not for the mRNA. Using the values for TSS2, IMD2 moves up in the ranking. It remains unclear whether this distribution of Spt6 drives changes in initiation profiles or if the changes in initiation drive the Spt6 occupancy, but this subset of genes had an unusual distribution of Spt6, this accumulation decreased in the interface mutants, and the score for the accumulation generally correlated with increased mRNA production and a shift in the use of TSSs.
Figure S8
Analysis of the effects of interface mutations on nucleosome occupancy.

S8A-B) Nucleosome occupancy was calculated in three regions, defined previously using the average occupancy profiles in WT cells (reference 4; NDR is 130 bp from -150 to -20 bp relative to the TSS, +1 is 150 bp from -20 to +130 bp, and the gene bodies are from 0-100% of the averaged genes, including a variable amount of overlap with the +1 site depending on the gene size). The \( \log_2 \text{FC} \) values relative to WT were calculated for all regions and are plotted here for each gene in the two compound interface mutants relative to one another. The high Pearson correlation coefficients indicate that the changes in the mutants at individual genes are reproducible for disruptions of either side of the interface. Least-squares fit lines pass above the origin for the NDR and +1 nucleosome due to the higher change in occupancy in these regions in the \( \text{rpb1}^{\text{TPY,FSP}} \) mutant, as shown in panel B.

S8C-D) The average nucleosome occupancy at the +1 site in the mutants correlated strongly with the values for WT, showing that in general the variation in +1 site signal was preserved in the mutants, with a global increase in the interface mutants, especially \( \text{rpb1}^{\text{TPY,FSP}} \). The exception was the RP genes, identified independently in panel D for the \( \text{rpb1}^{\text{TPY,FSP}} \) strain, showing that these genes consistently cluster below the diagonal, indicating a distinct loss of +1 nucleosome signal for this set but a gain for most other genes.
S8E-F) The average nucleosome occupancy across averaged genes is shown for multiple strains, with genes sorted into deciles by the level of transcripts and the high, middle, and low deciles plotted along with the average for all genes. In panel E, MNase-seq was performed in this study, with interface mutations causing a relative increase over 5' ends that was mainly due to genes producing moderate levels of transcripts. In panel F, data were generated using the same sample-preparation protocols but a slightly different informatics pipeline and strains described in reference 4 but analyzed in parallel with the data in panel E. The WT strains gave very similar profiles in either case, so this difference does not appear to affect this profile analysis significantly. The FACT mutant \(pob3^{Q308K}\) failed to produce the normal reduction in nucleosome occupancy seen for highly transcribed genes in both WTs and in the interface mutants (red lines), even though these genes produced the expected high levels of transcripts in these strains (4). The interface mutants also did not display the more dramatic drop in occupancy over the 3' half of genes seen in general in the FACT mutant. These results indicate that the changes in nucleosome occupancy profiles seen in the interface mutants share some overlap with a FACT mutant but are distinct from this general histone chaperone defect.
S8G) As in Fig 8D, except showing full averaged gene bodies instead of aligning by the TSS. The same general features are observed, but this shows that the loss of +1 nucleosome occupancy in RP genes is followed by a general increase in occupancy downstream of this site. The patterns seen in interface mutants are again distinct from the pattern observed in a FACT mutant.

S8H) As in Fig 8F, genes are sorted by the change in transcript level instead of the total amount of transcripts as in Fig 8D and S8G, but the whole averaged gene length is displayed. Genes with decreased transcript levels in interface mutants generally gained nucleosome occupancy over the gene body, but lost occupancy over the +1 site, with the opposite pattern for genes that gained transcripts, consistent with the correlations shown in Fig 8E.

G) Groups classified by total transcript levels:

H) Groups classified by changes in transcript levels:
The average nucleosome position near the TSS was calculated for all genes and the difference between mutant and WT strains was calculated as described in Methods or as described previously (4). Top panels: "Skinny" nucleosome distributions are shown for single-end sequencing of MNase fragments from WT, spt6\textsuperscript{R,KK} and rpb1\textsuperscript{T,FS\#} strains, and paired-end sequencing results from WT and rpb1\textsuperscript{T,FS\#} strains (for paired-end sequencing, 140-170 bp fragments were analyzed). The mutants show a progressive displacement of nucleosome positions downstream relative to the WT. Middle panels: Box-whisker plots of the displacement of nucleosomes at the position indicated relative to WT for each gene in the mutants shown. Boxes show the 25th, 50th, and 75th percentiles of the distributions, whiskers show the 10th and 90th percentiles, and a + sign indicates the mean. The -1 nucleosome position was relatively stable, and the progressive shift downstream inside genes visible in the aggregated data in the top panels is also observed in this view. Bottom panels: Genes were sorted into quintiles by the transcript level in WT strains and the average displacement for each bin is shown. All three mutants displayed a downstream shift in nucleosome positioning that was greater with higher transcript abundance, but this seemed to extend further into gene bodies for the FACT mutant while being limited to the first two nucleosomes in interface mutants. Effects were similar at the same genes, with an overall Pearson correlation coefficient of 0.70 comparing +2 nucleosome shifts for 5271 genes with both nucleosome positioning data at this site and measurable transcripts.
S8J-M Analysis of SRG1-SER3 transcripts.

S8J) A map of the SRG1-SER3 locus drawn to scale showing the location of the PCR primers used and relevant features of the locus.

S8K) qPCR was performed with the same samples and analytic methods used in Fig 2B. The ratio of SRG1 signal to SER3 signal was calculated as an indication of the linkage between SRG1 transcription and suppression of SER3 expression, showing that interface mutations have a defect in this parameter. As with other genes described here where an ncRNA impinges on the promoter/TSS of a downstream gene, the interface mutations caused increases in both the ncRNA level and the mRNA level of the affected gene.

S8L) Samples and analysis were as in Fig 2C except using primers for SRG1 and SER3. Deletion of the exosome caused the expected increase in the level of the ncRNA from SRG1, and interface mutants had little or no additional effect in this background, again suggesting the interface mutations caused a defect in exosome-mediated degradation of this transcript. Deletion of the nuclear exosome also caused an increase in SER3 mRNA by itself, but the increase caused by the interface mutants was not further impacted by loss of the exosome, similar to the effects seen with other genes described in this report that suggest some linkage between degradation of the ncRNA and the use of downstream TSSs.

S8M) Samples and analysis are as in Fig S3B-E. FACT mutations and spt6<sup>F249K</sup> caused small, but opposite changes in SRG1 levels, with FACT mutations behaving more like an interface mutation in this case. All mutants tested displayed large increases in SER3 mRNA, consistent with decreased linkage between SRG1 transcription and SER3 suppression, but with different profiles of effects for each mutation, consistent with distinct contributions by each factor to regulation of SER3 expression.
1. Sdano, M.A., Fulcher, J.M., Palani, S., Chandrasekharan, M.B., Parnell, T.J., Whitby, F.G., Formosa, T. and Hill, C.P. (2017) A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *Elife, 6*.

2. Storici, F., Lewis, L.K. and Resnick, M.A. (2001) In vivo site-directed mutagenesis using oligonucleotides. *Nat Biotechnol, 19*, 773-776.

3. Edgar, R., Domrachev, M. and Lash, A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res, 30*, 207-210.

4. McCullough, L.L., Pham, T.H., Parnell, T.J., Connell, Z., Chandrasekharan, M.B., Stillman, D.J. and Formosa, T. (2019) Establishment and Maintenance of Chromatin Architecture Are Promoted Independently of Transcription by the Histone Chaperone FACT and H3-K56 Acetylation in *Saccharomyces cerevisiae*. *Genetics, 211*, 877-892.

5. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. and Cramer, P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature, 560*, 607-612.

6. Pathak, R., Singh, P., Ananthakrishnan, S., Adamczyk, S., Schimmel, O. and Govind, C.K. (2018) Acetylation-Dependent Recruitment of the FACT Complex and Its Role in Regulating Pol II Occupancy Genome-Wide in Saccharomyces cerevisiae. *Genetics*.

7. Steinmetz, E.J., Conrad, N.K., Brow, D.A. and Corden, J.L. (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nucleic Acids Res, 40*, D700-705.

8. Villa, T., Barucco, M., Martin-Niclos, M.J., Jacquier, A. and Libri, D. (2020) Degradation of Non-coding RNAs Promotes Recycling of Termination Factors at Sites of Transcription. *Cell Rep, 32*, 107942.

9. Doris, S.M., Chuang, J., Viktorkovskaya, O., Murawska, M., Spatt, D., Churchman, L.S. and Winston, F. (2018) Spt6 is Required for the Fidelity of Promoter Selection. *Mol Cell, 72*, 687-699 e686.

10. Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform, 14*, 178-192.
19. Kwapisz, M., Wery, M., Despres, D., Ghavi-Helm, Y., Soutourina, J., Thuriaux, P. and Lacroute, F. (2008) Mutations of RNA polymerase II activate key genes of the nucleoside triphosphate biosynthetic pathways. *EMBO J*, **27**, 2411-2421.

20. Thiebaut, M., Colin, J., Neil, H., Jacquier, A., Seraphin, B., Lacroute, F. and Libri, D. (2008) Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Mol Cell*, **31**, 671-682.