Structural and Functional Characterization of the N Terminus of Schizosaccharomyces pombe Cwf10

S. Brent Livesay,a Scott E. Collier,a Danny A. Bitton,a Jürg Bähler,b Melanie D. Ohi,a

Department of Cell and Developmental Biology, Vanderbilt University Medical Center, Nashville, Tennessee, USAa; Department of Genetics, Evolution and Environment, University College London, London, United Kingdomb

The spliceosome is a dynamic macromolecular machine that catalyzes the removal of introns from pre-mRNA, yielding mature message. Schizosaccharomyces pombe Cwf10 (homolog of Saccharomyces cerevisiae Snu114 and human U5-116k), an integral member of the U5 snRNP, is a GTPase that has multiple roles within the splicing cycle. Cwf10/Snu114 family members are highly homologous to eukaryotic translation elongation factor EF2, and they contain a conserved N-terminal extension (NTE) to the EF2-like portion, predicted to be an intrinsically unfolded domain. Using S. pombe as a model system, we show that the NTE is not essential, but cells lacking this domain are defective in pre-mRNA splicing. Genetic interactions between cwf10ΔNTE and other pre-mRNA splicing mutants are consistent with a role for the NTE in spliceosome activation and second-step catalysis.

Characterization of Cwf10-NTE by various biophysical techniques shows that in solution the NTE contains regions of both structure and disorder. The first 23 highly conserved amino acids of the NTE are essential for its role in splicing but when overexpressed are not sufficient to restore pre-mRNA splicing to wild-type levels in cwf10ΔNTE cells. When the entire NTE is overexpressed in the cwf10ΔNTE background, it can complement the truncated Cwf10 protein in trans, and it immunoprecipitates a complex similar in composition to the late-stage U5 snRNP spliceosome. This data suggests that the structurally flexible NTE is capable of independently incorporating into the spliceosome and improving splicing function, possibly indicating a role for the NTE in stabilizing conformational rearrangements during a splice cycle.

Eukaryotic pre-mRNA splicing involves the precise removal of introns from pre-mRNA to form protein-coding mature messages (mRNA). This essential step in gene regulation is catalyzed by the spliceosome, a dynamic machine composed of four snRNPs (small nuclear ribonucleoproteins), U1, U2, U5, and U4/U6, and additional pre-mRNA splicing factors (1, 2). In vitro studies have led to a stepwise model of spliceosome assembly whereby the formation of an active spliceosome involves a series of regulated steps requiring the assembly and disassembly of large multiprotein complexes (1, 3). In this model, spliceosome assembly begins with the recognition of the 5′ and 3′ splice sites by the U1 snRNP and U2AF, respectively, while additional components in the U2 snRNP recognizes the branch point sequence. The subsequent engagement of the U4/U6.U5.U2/U6 spliceosome (Complexed with Cdc5) has been investigated only briefly and how Snu114 carries out these potential functions are not well understood. Schizosaccharomyces pombe Snu114 homolog Cwf10 (Complexed with Cdc5) has been investigated only briefly and noted for roles in RNA interference (RNAi)-directed centromere repeat silencing and in splicing (25).

The Snu114 family of proteins shares homology with the eukaryotic translation elongation factor EF2 (12) but is also pre-
dicted to contain regions of intrinsic disorder (26, 27). We and others have taken advantage of crystal structures of EF2 (28) to predict the EF2-like domain boundaries in the sequences of S. cerevisiae Snu14 (19) and S. pombe Cwf10 (Fig. 1A). By homology, there are six domains that define the “EF2-like” portion of Cwf10 (I, G, II, III, IVa, and V in Fig. 1A). Extensive mutagenic analysis of S. cerevisiae Snu14 has demonstrated that altering residues in all six EF2-like domains impairs protein function (19, 29).

The Snu14/Cwf10 proteins differ significantly from EF2 in that they contain a conserved N-terminal extension (NTE) (Fig. 1A). The NTE is approximately 120 amino acids (aa) long and is rich in acidic residues, with 39% of the first 56 residues being rich in acidic residues, with 39% of the first 56 residues being

![Diagram](https://example.com/diagram.png)

**FIG 1** The N-terminal extension is conserved in Snu14/Cwf10 family members. (A) Domain map of *S. pombe* Cwf10, as defined by three-dimensional modeling (Modeller [95]) of *S. pombe* Cwf10 onto the crystal structure of *S. cerevisiae* EF2 (PDB 1N0V). Domains are named as defined by homology with EF2, and the residue numbers listed above the domains refer to *S. cerevisiae* Snu114. (B) ClustalW (96) sequence alignment of the N termini of Snu14/Cwf10 homologs (from *S. pombe*, *Caenorhabditis elegans*, Drosophila melanogaster, Homo sapiens, and *S. cerevisiae*). Residues that are identical in at least three of the five homologs are in bold. Clusters of identical residues are highlighted in gray. Boundaries of the various Cwf10 truncations and the *S. cerevisiae* Snu14ΔN truncation are noted above the alignment. Symbols: *, identical; : (colon), similar. An arrow marks the transition from the NTE to EF2-like domain I, as assigned in panel A.

Functional data implicate the NTE in spliceosome activation and catalysis; however, very little is known about the structure and binding partners of the NTE. In human *in vitro* splicing assays, addition of antibodies against the U5-116K NTE partially blocks the second step of splicing (12). Removal of *S. cerevisiae* Snu14’s NTE (S. cerevisiae snu14ΔN) causes a temperature-sensitive (ts) phenotype in vivo, while *in vitro* splicing extracts prepared from *S. cerevisiae* snu14ΔN cells are unable to efficiently unwind the U4/U6 snRNAs prior to the first catalytic step of pre-mRNA splicing (14). These extracts also contain a destabilized U5 snRNP (14).

Finally, the *S. cerevisiae* snu14ΔN allele shows synthetic lethal and synthetic sick interactions with mutations in the U5 snRNA loop 1 and internal loop 1 (IL-1), as well as with U6 snRNA alleles that disrupt U2/U6 base pairing (30). These results led the authors to speculate that the *S. cerevisiae* Snu14 NTE may be involved in facilitating U5 and U6 snRNA interactions near the 5’ splice site. Interestingly, no physical interactions have been determined for the NTE in any organism, leaving open the question of how the NTE is spatially oriented within the spliceosome. Additionally, the only structural information about the NTE is a bioinformatics prediction that the NTE of human U5-116K is composed of two disordered regions of equal length (26).

In this study, we characterize the *S. pombe* Cwf10 NTE both *in vivo* and biochemically. We show that although the NTE is not essential in *S. pombe*, deleting this region leads to a general splicing defect at all temperatures. We define a small region of the NTE required for efficient splicing and demonstrate the presence of both structural order and disorder within the NTE. Finally, we show that when the NTE is overexpressed in *vivo* it stably associates with a protein complex similar to the *S. pombe* U5.52/U6 spliceosomal complex and rescues the splicing defect caused by deletion of the NTE from endogenous *cwf10* (31). Taken together, these findings suggest that the NTE is a semidimered domain that has the ability to function in *trans* to the EF2-like portion of Cwf10.

**MATERIALS AND METHODS**

**Strains, yeast methods, and molecular biology.** Strains used in this study are listed in Table S1 in the supplemental material. Yeast strains were grown in yeast extract (YE) medium or Edinburgh minimal medium with appropriate supplements. The *spp41-l* (31) and *spp42-l* (31) open reading frames (ORFs) were tagged endogenously at the 3’ end with kanMX6 for genetic analyses as previously described (32). Transformations were done as described previously (33) for all tag insertions, gene replacements, and introduction of plasmids. Integration of tags was verified using whole-cell PCR and immunoblot analysis as appropriate. Crossing of tagged and mutated loci into other strains was accomplished using standard *S. pombe* mating, sporulation, and tetrad dissection techniques. For spot assays, cells were grown to mid-log phase at 25°C and resuspended in water to achieve an optical density at 595 nm (OD595) of 0.2 (Fig. 2) or 0.6 (see Fig. 5). Tenfold serial dilutions were made, and 2.5 μl of each dilution was plated on YE. Plates were incubated at the indicated temperatures for 2 to 9 days before imaging.

For induction of the nmt3X promoter (34), the cells were first grown overnight in medium containing 5 μg/ml of thiamine and then washed...
Characterization of the S. pombe cwf10-ΔNTE allele. (A) Growth of serial dilutions of wild-type (cwf10+), and cwf10-ΔNTE. Serial dilutions of equal numbers of cells were spotted on solid YE medium and grown at 18°C, 25°C, 32°C, and 36°C. (B) Growth of cwf10+ and cwf10-ΔNTE cells in liquid YE at 25°C. Samples were collected at 2-h intervals to determine cell numbers. (C) The NTE is required for efficient pre-mRNA splicing. RT-PCR was used to characterize RNA extracted from cwf10+ and cwf10-ΔNTE cells. Total RNA was isolated from cells grown at 25°C. RT-PCR was performed with random hexamer primers. Portions of the tbp1 (containing introns) and act1 (introns) transcripts were amplified from the RT-PCR products. Mature (m) and pre-mRNA (p) forms are labeled. No bands were detected in any RT-PCR samples in which the reverse transcriptase enzyme was omitted (−RT). (D) Immunoblot demonstrating relative protein levels of pre-mRNA splicing factors in either cwf10+ or cwf10-ΔNTE backgrounds. An asterisk (*) marks the position of each protein on the blot. (E and F) Lysates from cwf10+ and cwf10-ΔNTE cells were probed with anti-PSTAIR and either anti-Cwf10 (D) or anti-Cdc5 (E) antibodies. (G) Lysates from prp3-V5 and prp3-V5 cwf10-ΔNTE cells probed with anti-V5 and anti-PSTAIR. Anti-PSTAIR detects Cdc2. Numbers to the left of each blot indicate molecular mass in kilodaltons. (I) Comparison of protein levels of each pre-mRNA splicing factor in cwf10-ΔNTE cell lysates relative to the cwf10+ value (100%), normalized to Cdc2, quantitated from Western analysis. Values were calculated using three biological replicates. Error bars indicate standard errors of the means. The P value was determined using the Student t test.

Liquid culture cell numbers were counted with a Z1 Coulter Counter (Beckman Coulter, Brea, CA) as previously described (36).

**Antibody preparation.** Hisα-Cwf10 (673–983) was expressed in Escherichia coli BL21(DE3) cells (EMD Millipore, Btymamark Park, MA), purified with HisPur Cobalt agarose (Pierce/Thermo Scientific, Rockford, IL), and used to immunize rabbits (Cocalico Biologicals, Reamstown, PA), as approved by the Vanderbilt Institutional Biosafety Committee. Cwf10-specific antibodies were affinity purified over N-hydroxysuccinimide ( NHS)-activated Sepharose Fast Flow 4 (GE Healthcare Life Sciences, Piscataway, NJ) covalently linked to Hisα-Cwf10 (673–983).

**Immunoprecipitations, immunoblotting, and sucrose gradients.** "Native" and "denatured" whole-cell lysates were prepared as previously described, with leupeptin omitted from lysis buffers (37). For immunoprecipitations, proteins were resolved by 10% SDS-PAGE (all lysates except Br2-HA), 4 to 12% Bis-Tris PAGE (Br2-HA lysates), or 8% SDS-PAGE (lysate gradients) and transferred by electroblotting to a Protran nitrocellulose membrane (Whatman, GE Healthcare, Piscataway, NJ). Primary antibodies/antisera used included anti-Cdc5 (1/5,000) (37) and anti-Asp1 (1/5,000) rabbit polyclonal antisera, anti-Cwf10 affinity purified rabbit polyclonal antibody (0.84 μg/ml), and anti-PSTAIR (detects S. pombe Cdc2)
molecules using poly(T) oligoattached magnetic beads. Following purification, the eluted poly(A) RNA was cleaved into small fragments of 120 to 210 bp using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first-strand cDNA using SuperScript II reverse transcriptase and random primers. This was followed by second-strand cDNA synthesis using DNA Polymerase I and RNase H. The cDNA fragments were then put through an end repair process, the addition of a single A base, and ligation to the Illumina multiplexing adaptors. The products were then purified and enriched with PCR to create the final cDNA sequencing library. The cDNA library underwent quality control by running on the Agilent Bioanalyzer HS DNA assay to confirm the final library size and on the Agilent Mx3005P qPCR machine using the KAPA Illumina library quantification kit to determine the concentration. A 2 nM stock was created, and samples were pooled by molarity for multiplexing. From the pool, 12 pM was loaded into each well for the flow cell on the Illumina cBot for cluster generation. The flow cell was then loaded onto the Illumina HiSeq 2500 utilizing v3 chemistry and HTA 1.8. The raw sequencing reads in BCL format were processed through CASAVA-1.8.2 for FASTQ conversion and demultiplexing. The RTA chastity filter was used, and only the PF (passfilter) reads were retained for further analysis. Raw expression data files are available from Gene Expression Omnibus (GEO accession number GSE47573; http://www.ncbi.nlm.nih.gov/geo/).

**Gene expression analyses.**

**Replicates were grown together and processed separately for all following steps.** The Vanderbilt Technologies for Advanced Genomics Core Facility (Vantage, Nashville, TN) used the TrueSeq Stranded mRNA sample preparation kit (Illumina, San Diego, CA) to convert the mRNA in 100 ng of total RNA into a library of template molecules suitable for subsequent cluster generation and sequencing on the Illumina HiSeq 2500. The input total RNA was quality checked by running an aliquot on the Agilent Bioanalyzer to confirm integrity. The Qubit RNA fluorometry assay was used to measure concentration. The input to library prep was 50 µl of 2 ng/µl DNase-treated total RNA. The total RNA underwent enrichment of the poly(A)-containing mRNA molecules using poly(T) oligoattached magnetic beads. Following purification, the eluted poly(A) RNA was cleaved into small fragments of 120 to 210 bp using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first-strand cDNA using SuperScript II reverse transcriptase and random primers. This was followed by second-strand cDNA synthesis using DNA Polymerase I and RNase H. The cDNA fragments were then put through an end repair process, the addition of a single A base, and ligation to the Illumina multiplexing adaptors. The products were then purified and enriched with PCR to create the final cDNA sequencing library. The cDNA library underwent quality control by running on the Agilent Bioanalyzer HS DNA assay to confirm the final library size and on the Agilent Mx3005P qPCR machine using the KAPA Illumina library quantification kit to determine the concentration. A 2 nM stock was created, and samples were pooled by molarity for multiplexing. From the pool, 12 pM was loaded into each well for the flow cell on the Illumina cBot for cluster generation. The flow cell was then loaded onto the Illumina HiSeq 2500 utilizing v3 chemistry and HTA 1.8. The raw sequencing reads in BCL format were processed through CASAVA-1.8.2 for FASTQ conversion and demultiplexing. The RTA chastity filter was used, and only the PF (passfilter) reads were retained for further analysis. Raw expression data files are available from Gene Expression Omnibus (GEO accession number GSE47573; http://www.ncbi.nlm.nih.gov/geo/).

**Genre level alignments and annotation.** Paired-end reads of 76-base length (each end) originating from each sample were aligned using Bowtie 0.12.7 (41) to the *S. pombe* genome sequence (Ensembl *S. pombe*, Build EF1, version 13) (42) as well as to the corresponding exon-exon junctions database (only the first part of the paired-end reads was considered). Up to 3 base pair mismatches were allowed. Reads that matched multiple loci were removed from further analysis, and the resultant alignment files were processed to generate “pile-ups” against each chromosome.

**Exon-exon junctions.** Searches were performed against the genome sequence combined with a data set of known exon-exon junctions as defined by Ensembl *S. pombe* release 13. To ensure that a 76-base read mapped to a splice junction, only the last 70 bases of the first exon and the first 70 bases of the second exon were considered (if the exon exceeded a length of 70 bases). In this way, reads that overlapped a junction by <6 nucleotides were excluded. Reads that matched to more than one junction or elsewhere in the genome were also discarded.

**Defining intronic and intergenic regions.** The known annotated set of *S. pombe* genes (7022; Ensembl version 13, as before) was used to define unambiguously antisense transcripts (i.e., those that exactly mirror known annotated genes without overlapping nearby genes), and unique accesses were assigned (“anti_xxx”; 3097). Using this augmented annotation, intergenic regions were defined (i.e., regions between known annotated and unambiguous antisense regions on each strand), and unique accesses were assigned (“int_xxx”; 8810). Unique accesses were also assigned to all known introns (“int_xxx”; 5361). Thus, in total 21,193 regions were interrogated across the 4 samples (annotated genes, introns, and intergenic regions; see Data set S1 in the supplemental material).

**Normalization, fold changes, and differential expression.** Differential expression between samples was determined using the DESeq Bioconductor package (43). A cutoff of ≥2-fold change and corrected P value of <0.05 were applied to derive a list of differentially expressed genes, introns, and intergenic regions.

**RNA-seq expression level.** For RNA sequence expression levels, normalized expression levels (E) for individual exons and introns were calculated using the following formula as described in references 44 and 45 (RPKM measure): \[ E = \log_2 \left( r / (T_i - 1) \right) \]. Briefly, the number of reads (R) detected across a given region at a given sample (i) was multiplied by a constant (C = 1 × 10^9) and divided by the total number of reads at that sample (T_i) multiplied by the region’s length (L). A small constant was added (10^-3) to all expression values to avoid taking logarithms of zero. Gene level expression values were summarized using exon data. Sample-specific expression levels for all regions included in the analysis were calculated.
terrogated in this study are provided in Data set S1 in the supplemental material.

**SE and differential splicing significance.** Splicing efficiency (SE) reflects the proportion of spliced mRNA signal relative to pre-mRNA signal. Splicing efficiency is computed by dividing junction reads (IR; also known as trans-reads) by reads that straddle an exon-intron boundary (EI; only the upstream 5' exon relative to the intron was considered) according to the following formula: \( SE = \log_2(I/R)/E \).

A Cochran-Mantel-Haenszel (CMH) chi-square test for repeated tests of independence, which accounts for biological replicates (46), was applied to identify statistically significant introns (i.e., those that display differences in their splicing efficiency between samples; mantelhaen.test was applied to identify statistically significant introns). Differences in their splicing efficiency between samples were analyzed using the program K2D2 (48) to estimate secondary structure. Near-UV sedimentation coefficients between 0 and 5S.

**Expression and purification of Cwf10(1–135).** The N-terminal sequence of amino acids 1 to 135 of Cwf10 with 6 C-terminal histidine residues \( \text{His}_{135} \) was cloned into pET15b (NcoI/BamHI; plasmid pH11020 (EMD Millipore) and transformed into E. coli Rosetta 2(DE3)pLysS cells (EMD Millipore). Cells were grown in Terrific broth (Invitrogen, Grand Island, NY) to an OD595 of 0.6, then diluted back to 2 M urea and digested overnight with trypsin (as described previously (50)). Resulting peptides (corresponding to about 95% of the TAP eluate) were analyzed by a 70-min data-dependent liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis in the Vanderbilt Mass Spectrometry core. In brief, TAP eluates were trichloroacetic acid (TCA) precipitated, resolubilized in 8 M urea–100 mM Tris (pH 8.5), reduced, alkylated, and then analyzed using the Bioconductory 2D-page/condor command in R. The false-discovery rate (q-value) was computed using the Bioconductor q-value package (47), and a cutoff for q of <0.05 was applied.

**CD.** For circular dichroism (CD) spectrometry, purified Cwf10(1–135)His4 was analyzed using a Jasco J-810 spectropolarimeter (Jasco Analytical Instruments, Easton, MD). Far-UV data were collected at a protein concentration of 0.18 mg/ml in a 1-mm quartz cuvette. Spectra were collected with an average time of 4 s for each point and a step size of 20 nm/minute from 198 to 260 nm. Far-UV spectra were collected in duplicate and background corrected against a buffer blank. Spectra were analyzed using the program K2D2 (48) to estimate secondary structure. Near-UV data were collected at a protein concentration of 2.01 mg/ml in a 1-cm quartz cuvette. Spectra were collected with an average time of 4 s for each point and a step size of 10 nm/minute from 250 to 330 nm. For both Cwf10(1–135)His6 and denatured Cwf10(1–135)His6 in 6 M guanidine- 

**NMR analysis.** Cwf10(1–135)His6 was purified as above, except that cells were grown and expressed in M9 medium supplemented with \(^{15}N\)-labeled ammonium chloride (Cambridge Isotopes, Andover, MA) as the only nitrogen source and 10% D2O was added to the final sample. During purification, one half of the sample was left in gel filtration buffer (25 mM Tris–HCl (pH 7.3), 100 mM NaCl, and 1 mM EDTA) and the other half was buffer exchanged into gel filtration buffer plus 6 M guanidine-HCl. Standard sensitivity-enhanced echo/antiecho \(^{1}H/\(^{15}N\)-heteronuclear single-quantum correlation (HSQC) nuclear magnetic resonance (NMR) data were collected at 25°C for both samples and at 50°C for the gel filtration buffer sample (no guanidine-HCl) using a 600 MHz Bruker AVIII spectrometer (Bruker, Billerica, MA) with a CPQCI probe and z-axis gradient. The spectra were processed using Topspin 3.2 (Bruker, Billerica, MA). The indirect dimension was four times zero filled to a final matrix of 2,048 × 1,024 data points, and 72 and 96 squared sine bell apodization was applied in the F2 and F1 dimensions, respectively. Spectra were further analyzed with Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

**RESULTS**

**Deleting the S. pombe Cwf10 NTE impairs splicing in vivo.** To more closely examine the in vivo role of the Cwf10 NTE in pre-mRNA splicing, we prepared an S. pombe cwf10 construct corresponding to S. cerevisiae snu114ΔN (14). Although cwf10 2–135Δ aligns with snu114ΔN (Fig. 1B), the S. pombe allele did not support growth at either 25°C or 32°C when integrated at the endogenous locus (data not shown), while the S. cerevisiae allele is viable at temperatures below 40°C (30). Upon closer examination of the alignment of Snu14 family members with EF2, it appears that Cwf10 2–135Δ lacks residues corresponding to domain I (Fig. 1B), which may affect the folding of this essential domain. Therefore, we deleted eight fewer amino acids and successfully integrated cwf10(2–127Δ) (here, cwf10-ΔNTE) at the endogenous locus. This result suggests that some of the defects seen in the S. cerevisiae snu114ΔN strain may be partly attributable to a defect in the folding of domain I in the EF2-like region of Snu114.

We further characterized cwf10-ΔNTE to examine the NTE's contribution to in vivo splicing. The cwf10-ΔNTE strain formed slightly fewer colonies on solid media at all temperatures tested (Fig. 2A) and exhibited slow growth in liquid culture at 25°C and 36°C compared with wild-type cells (Fig. 2B and data not shown).
Conversely, overexpression of the NTE (cwf10 1–135 or 1–107) under the high-strength nmt (no message in thiamine) promoter did not impair cell viability, indicating its expression is not dominant negative (data not shown). RT-PCR analysis of tpi1, an mRNA intron highly sensitive to splicing defects (53), indicated that cwf10-ΔNTE cells grown at 25°C accumulate unspliced transcript (Fig. 2C). Thus, removal of the Cwf10 NTE reduces splicing efficiency in vivo.

In our experience, yeast spliceosome mutants often have a lower steady-state mutant protein level that correlates with the splicing defect. To determine whether the deletion of the Cwf10 NTE simply decreases the amount of Cwf10 in cells, thus reducing splicing efficiency, antibodies were generated against Cwf10 amino acid residues 673 to 983 (domains IV and V). Immunoblotting of S. pombe lysate with α-Cwf10 antisemur detected a single protein at the anticipated molecular mass of ~116 kDa (Fig. 2D, cwf10+ lane). In cwf10-ΔNTE lysates, the antisemur detected a band ~20 kDa smaller than the wild-type protein, consistent with the predicted molecular mass of the deletion (Fig. 2D). Using quantitative Western blotting, we found that levels of Cwf10 were consistently higher in cwf10-ΔNTE cells than in wild-type cells when quantified against the loading control Cdc2 (CdK1) (Fig. 2I). These results show that low levels of Cwf10 do not cause the splicing defect in cwf10-ΔNTE cells. To examine whether levels of other splicing components might be affected in this background, we analyzed the protein levels of S. pombe Cdc5 (homolog of S. cerevisiae Cef1) (Fig. 2E), a core component of the Nineteen complex (NTC) (54, 55), Brr2 (S. cerevisiae Br2) (Fig. 2F), and a core U5 snRNP member (13), as well as Prp1 (S. cerevisiae Prp6) and Prp3 (S. cerevisiae Prp3) (Fig. 2G and H), both members of the U4/U6/U5 tri-snRNP and B-complex (2, 56). In the cwf10-ΔNTE background, levels of Brr2 and Prp3 were similar to levels seen in wild-type cells; however, both Cdc5 and Prp1 levels were reduced to ~70% and ~65%, respectively (Fig. 2I). Since the prp1 transcript does not contain introns, we attribute the lower level of these proteins to reduced protein stability rather than a general reduction in transcript levels.

**RNA-seq demonstrates a general splicing defect in cwf10-ΔNTE cells.** Previously it has been demonstrated that mutations in core spliceosome proteins affect splicing of introns to different degrees (57–59). That is, one mutant can have a profile of inefficiently spliced introns that is significantly different from the profile of another mutant. To examine the extent of splicing changes in cwf10-ΔNTE cells, we performed transcriptome analysis. RNA was extracted from cwf10+ and cwf10-ΔNTE cells grown at 25°C, and two replicates for each genetic background were sequenced on the Illumina HiSeq platform. Two methods were employed to evaluate the effect of NTE deletion on splicing efficiency using the RNA sequencing data. A Cochran-Mantel-Haenszel (CMH) test assessed the significance of the change in the ratio of spliced to unspliced transcripts between cwf10-ΔNTE and cwf10+ (see Materials and Methods), while DESeq (43) revealed differentially expressed introns between the two conditions. Splicing efficiency was highly reproducible between replicates, with a tight distribution along the diagonal (see Fig. S1 in the supplemental material), but displayed a nearly global shift when cwf10-ΔNTE and cwf10+ were compared; this can be seen as a shift of most points toward the right of the diagonal in Fig. 3A. This suggests an extensive splicing defect in the NTE-deleted strain that affected most introns. Using the splicing efficiency (SE) approach under a stringent threshold (q-value, <0.05; fold change, >2), we identified 1,193 introns (of 5,361 possible introns) whose splicing efficiency was significantly compromised in the mutant. Similarly, using DESeq, 883 introns were more highly expressed in cwf10-ΔNTE relative to cwf10+, indicating that their splicing was significantly less efficient in the mutant strain (adjusted P value of <0.05 and fold change of >2; Fig. 3B). There was a good correspondence between the two methods (P < 2.2e−226; hypergeometric test [Fig. 3C]), which was even stronger when no fold change cutoff was applied (P < 2.7e−249 [data not shown]). Overall, the splicing of 2,352 introns (approximately 44% of all introns) was significantly compromised by the NTE deletion, as indicated by at least one method without the fold change cutoff. The effect of the NTE deletion was not dependent on intron length, and neither was the magnitude of the reduction in splicing efficiency (Fig. 3D, top and bottom panels, respectively). Moreover, the effect of the mutation was not dependent on intron GC content, branch site sequence, 5′ intron sequence, or the position of the intron within the transcript (data not shown). For a small set of introns (total, 17), splicing efficiency was significantly improved by the deletion, but there was no agreement between the two methods under these criteria (14 were found by DESeq and 3 by SE). However, when the magnitude of the fold change was ignored, the splicing efficiency of 223 introns was significantly improved (209 by DESeq and 24 by SE, of which 10 were identified by both methods; P < 3.9e−10). We did not find any clear sequence features that could explain why a small percentage of introns were spliced with an efficiency that was either the same as or better than that in the NTE deletion background. Taken together, the RNA sequence (RNA-seq) data indicate that the NTE of Cwf10 is required for efficient splicing of most S. pombe introns.

**The lack of Cwf10-NTE changes spliceosome sedimentation patterns.** In vitro splicing assays, deletion of the S. cerevisiae Snu114 NTE inhibits U4/U6 snRNA unwinding, thus impairing activation of the spliceosome (14). Since S. pombe does not have a robust in vitro splicing extract system, likely due to the stability of the U5.U2/U6 spliceosomal complex (60, 61), we could not directly test whether S. pombe Cwf10-NTE plays a similar role in fission yeast. However, we postulated that if the S. pombe Cwf10-NTE shares a function similar to that of its S. cerevisiae ortholog, cells that lack the Cwf10-NTE might have altered sedimentation patterns of spliceosomal complexes and/or snRNAs. Thus, to test the effect of deleting Cwf10-NTE in vivo, we compared the sedimentation patterns of Cwf10, Cdc5 (S. cerevisiae Cef1), and the five snRNAs in sucrose gradients using native lysates made from either wild-type or cwf10Δ-NTE cells (Fig. 4). It has been previously shown that in lysates from asynchronously growing wild-type S. pombe cells, the majority of spliceosomes sediment as a stable ~37S U5.U2/U6 complex (10, 37, 60, 61), although less abundant U5/U6/U4/U2/U1 and U5/U6/U4/U2 complexes have also been characterized (62, 63). Unlike what is found in other organisms, S. pombe lysates lack detectable quantities of a preassembled U5/U4/U6 tri-snRNP (61).

As expected, in wild-type lysates a majority of Cwf10 sediments at an ~37S peak, corresponding to the sedimentation pattern of the U5.U2/U6 complex (Fig. 4A and B, fraction 9 [10]); however, the sedimentation pattern of Cwf10-ΔNTE in cwf10-ΔNTE lysates changes, with a portion of Cwf10-ΔNTE shifting to higher-molecular-mass fractions (Fig. 4A and B, fractions 11 and 12). The sedimentation pattern of Cdc5, a core NTC component, does not
The differences include a small but consistent shift of some of the
patterns of all the snRNAs were altered to some degree (Fig. 4E to N).
U2 snRNP at 12S (Fig. 4E and F), which most likely represents the
between fission yeast and other organisms is the abundance of the
core U2 snRNPs that contains the SF3a and SF3b subcomplexes (64–67).
However, when comparing wild-type and cwf10-ΔNTE lysates, the sedimentation patterns of all the snRNAs were altered to some degree (Fig. 4E to N).

Thus, from these analyses we conclude that deleting the S. pombe Cwf10 NTE does alter the distribution of spliceosomes in vivo, although in a complex pattern with shifts to both higher and lower molecular masses.

cwf10-ΔNTE synthetically interacts with factors predicted to be involved in spliceosome activation. As another approach to investigate whether S. pombe NTE function can be correlated to a specific step of the pre-mRNA splicing reaction or stage of spliceosomal organization, we tested for genetic interactions between cwf10-ΔNTE and eight previously characterized S. pombe pre-mRNA splicing alleles (Fig. 5). Our analysis revealed that cwf10-ΔNTE is synthetic lethal when combined with prp1-4 (68). (S.
FIG 4 Sedimentation of spliceosome components is altered in cwf10-ΔNTE cells. Lysates from cwf10+ and cwf10-ΔNTE cells grown at 25°C were run on 10 to 30% sucrose gradients. Fractions were collected, and either they were TCA precipitated for immunoblot analysis or RNA was extracted for Northern analysis. (A and C) Immunoblots of fractions collected from 10 to 30% sucrose gradients of lysates from cwf10+/H11001 prp1-myc and cwf10-ΔNTE prp1-myc using anti-Cwf10 (A) and anti-Cdc5 (C). (B and D) Intensities of the bands of the anti-Cwf10 blot (B) and anti-Cdc5 (D) were quantified and plotted as a percentage of the sum of the signal from all fractions. The bands were quantified by near-infrared detection and Odyssey software (Licor). (E, G, I, K, and M) RNA was isolated from fractions collected from 10 to 30% sucrose gradients of cwf10+/H11001 and cwf10-ΔNTE lysates. Blots were probed with 32P-labeled oligonucleotides complementary to the S. pombe U2 (E), U5 (G), U6 (I), U1 (K), and U4 (M) snRNAs. (F, H, J, L, and N) Intensities of the bands of the U2 (F), U5 (H), U6 (J), U1 (L), and U4 (N) probes were quantified and plotted as a percentage of the sum of the signal from all fractions. The bands were quantified by phosphorimaging and ImageQuant TL 8.1 software (GE Healthcare). The asterisk (*) in panels H and L highlights the peak of the U5 and U1 snRNAs found in a small-molecular-mass fraction in cwf10-ΔNTE. Column I indicates where the U5/U2/U6 spliceosomal complex sediments in the gradients. The migration of fatty acid synthase (FAS) (40S), thyroglobulin (19S), and catalase (11.3S) collected from parallel gradients is indicated.
cerevisiae PRP6), spp42-1 (31) (S. cerevisiae PRP8), and cdc5-120 (69) (S. cerevisiae CEF1). When cwf10-NTE is combined with either prp4-73 (70) (hPRPF4B) or spp41-1 (31) (S. cerevisiae BRR2) the cells are synthetically sick. Conversely, synthetic interactions are weak or nonexistent between cwf10-ΔNTE and aar2Δ (71) (S. cerevisiae AAR2), prp10-1 (68) (S. cerevisiae HSH155), or cdc28-P8 (72) (S. cerevisiae PRP2). Two of the three synthetic lethal interactions, prp1-4 and spp42-1, are with genes previously shown to be important for spliceosome activation (62, 73). The third synthetic lethal interaction, cdc5-120, is interesting because Cdc5 is likely involved in spliceosome activation as a component of the NTC (74) and in modulating the transition between first- and second-step catalysis (75). We conclude that the S. pombe Cwf10 NTE is likely important for a specific stage in spliceosome activation, and possibly also for a later stage(s) in the splicing reaction.

To further address whether the NTE may be involved in facilitating the transition from an inactive to activated spliceosome, we analyzed the protein composition of the S. pombe U5, U2/U6 spliceosomal complex purified from either a cwf10-ΔNTE or a wild-type background using one-dimensional (1D) liquid chromatography-tandem mass spectrometry (LC-MS/MS), a technique well suited for detecting peptides that are stoichiometrically present in a purification. Because the NTE is not essential and deletion of this region does not cause a temperature-sensitive phenotype (Fig. 1), we reasoned that any changes seen in the U5, U2/U6 complex purified from cwf10-ΔNTE cells would likely be subtle, and thus we did the purifications for this analysis using mild salt (75 mM NaCl) conditions. Cdc5-TAP (S. cerevisiae Cef1), which associates with the S. pombe U5, U2/U6 complex (10), was used to purify the U5, U2/U6 complex. Peptide counts for spliceosomal proteins found in each 1D LC-MS/MS run are included in Table 1 to provide a semiquantitative indication of protein amounts. From this analysis, the major differences between the two purifications were the lack of Br2 (S. cerevisiae Br2), Mug161 (human CWF19L1), and Prp43 (S. cerevisiae Prp43) peptides detected in the Cdc5-TAP from cwf10-ΔNTE cells. Br2 is an essential U5 snRNP component that forms a salt-stable complex with Prp8 (S. pombe Spp42) and Snu114 (S. pombe Cwf10) (13). Mug161 (human CWF19L1) is a protein of unknown function that has been isolated in spliceosomes purifications from both S. pombe (38) and mammalian cells (76), and Prp43 is required for spliceosome disassembly (77–79). Because this analysis was done using 1D LC-MS/MS rather than the more-sensitive MudPit (multidimensional protein identification technology) MS, the lack of detected peptides for particular proteins such as Br2 could indicate either a complete absence of the protein in the sample or that the protein is at substoichiometric levels in the purification. Overall, the differences that we detect between U5, U2/U6 complexes purified from either wild-type or cwf10-ΔNTE backgrounds suggests that NTE may play a role in stabilizing Br2’s interaction with the U5 snRNP during spliceosome activation, a model supported by the synthetic sick interaction shown between cwf10-ΔNTE and spp41-1 (S. cerevisiae Brr2) (Fig. 5A).

The Cwf10 NTE is partially folded but contains regions of disordered coiled coil. The structural characteristics of the NTE were of interest to us for several reasons. First, the only published data on NTE structure consist of a bioinformatics prediction that the region is unfolded (26). Second, the NTE does not carry strong homology to known protein domains or to primary sequences in other proteins (data not shown). To more carefully examine the structural characteristics of this domain, we used Disopred (80), a program that predicts structure disorder, to precisely map potential regions of intrinsic disorder in the Cwf10 NTE. This analysis showed that while a majority of the first N-terminal 60 amino acids are strongly predicted to not adopt any secondary structure, the C-terminal 75 amino acids are predicted to be structurally ordered (Fig. 6A), perhaps as β-strands as suggested by analysis with Psipred version 3.3 (81, 82), a program that predicts secondary structure (data not shown). To experimentally test this model, we expressed and purified recombinant Cwf10(1–135)His6, from E. coli (see Fig. S2A in the supplemental material). Analysis by sedimentation analytical ultracentrifugation (SVAU) shows that the Cwf10 NTE sediments as a monomer (s = 0.8; predicted molecular mass, ~19 kDa; root mean square deviation [rmsd] = 0.09) with a frictional ratio of 2.13 (Fig. 6B). Cwf10(1–135)His6 was then analyzed by circular dichroism (CD) spectroscopy using near-UV wavelengths, which can be used to detect tertiary structure. Analysis of this spectrum shows a strong signal between 260 and 290 nm, suggesting that there are some aromatic residues found in a folded environment (Fig. 6C). Importantly, this signal is no longer seen when the protein is completely denatured in 6 M
TABLE 1 pre-mRNA splicing factors copurifying with Cdc5-TAP in either a cwf10-ANTE or a wild-type background identified using 1D liquid chromatography-tandem mass spectrometry

| Spliceosome sub-complex | S. pombe protein | ORF number | Cdc5-TAP \(^a\) | S. cerevisiae protein | H. sapiens protein |
|------------------------|-----------------|------------|-----------------|----------------------|------------------|
|                        |                 |            | cwf10-ANTE      | Wild type            |                  |
|                        |                 |            | I               | II                   | I                | II               |
| Core snRNP             | Smnl1           | SPAC26A3.08| 4               | 4                    | 12               | 11               |
|                        | Snmd1           | SPAC27B7.07c| 4               | 3                    | 8                | 8                |
|                        | Snmd2           | SPAC2C4.03c| 5               | 6                    | 14               | 13               |
|                        | Snmd3           | SPBC19C2.14| ...             | 2                    | 5                 | 14               |
|                        | Snt1e           | SPBC11G1.11b| 2               | 1                    | 6                | 4                |
|                        | Sntfl           | SPBC3E7.14 | 5               | 8                    | 6                | 10               |
|                        | Sntgl           | SPBC4B4.05 | 7               | 3                    | 10               | 8                |
|                        | Leal            | SPBC186.01c| 11              | 11                   | 28               | 27               |
|                        | Md1l            | SPBC8D2.09c| 3               | 2                    | 2                | 6                |
|                        | Prp2/Mis11      | SPBK1.14b   | ---             | ---                  | ---              | ---              |
|                        | Sapl1           | SPBC56.09   | 3               | 2                    | 2                | 3                |
|                        | Sapl1/4         | SPAC2A12.09c| 1               | 1                    | 1                | 3                |
|                        | In1l            | SPBC2B3.02c| ...             | 1                    | 1                 | 1                |
|                        | Sapl45          | SPAC2F8.14c| ---             | 2                    | 1                | 1                |
|                        | Prp10           | SPAC2F7.11c| 3               | 5                    | 3                | 8                |
|                        | Prp12           | SPAP69G.03c| 6               | 9                    | 5                | 3                |
| U2                     | Cwf10-ANTE      | SPBC215.12 | 31              | 29                   | ---              | ---              |
|                        | Cwf10           | SPBC215.12 | ---             | ---                  | 100              | 73               |
|                        | Brn2            | SPAC9.03c   | ---             | ---                  | 41               | 11               |
|                        | Spp42           | SPAC4F8.12c| 59              | 78                   | 186              | 171              |
|                        | Cwf17/Spt58     | SPBC128.11 | 14              | 14                   | 41               | 38               |
|                        | Cdc5            | SPAC64.12   | 44              | 46                   | 92               | 71               |
|                        | Cwf2            | SPAC3A12.11c| 5               | 7                    | 37               | 38               |
|                        | Cwf7            | SPBC2B2.04c| 10              | 14                   | 34               | 13               |
|                        | Cwf15           | SPBC337.06c| 2               | 3                    | 13               | 11               |
|                        | Prp5            | SPBPC2H7.07 | 19              | 19                   | 51               | 57               |
|                        | Prp19           | SPAC29A4.03c| 25              | 24                   | 75               | 48               |
|                        | Cwf3            | SPBC211.02c| 23              | 29                   | 47               | 56               |
|                        | Cwf4            | SPBC31F10.11c| 18              | 23                   | 67               | 48               |
|                        | Cwf5/Cem2       | SPCC550.02c| 12              | 17                   | 52               | 42               |
|                        | Cwf12           | SPBC32F12.05c| 6               | 3                    | 16               | 12               |
|                        | Prp45           | SPCC18.11   | 13              | 19                   | 28               | 30               |
|                        | Prp17           | SPBC36I.10  | 19              | 17                   | 38               | 45               |
|                        | Syd2            | SPBC37.13c  | 34              | 37                   | 115              | 97               |
|                        | Cwf11           | SPBC2A6.02  | 7               | 9                    | 16               | 13               |
|                        | Cwf16           | SPAC9.13c   | 2               | 3                    | 1                | 8                |
|                        | Cwf18           | SPCC1E11.07c| 7               | 6                    | 12               | 13               |
|                        | Cwf19           | SPAC30D1.09 | 18              | 16                   | 37               | 47               |
|                        | Cwf21           | SPAC4A8.09c| 3               | 1                    | 6                | 7                |
|                        | Cwf22           | SPBC137.01 | 7               | 9                    | 15               | 18               |
|                        | Cwf25           | SPBC146.05c| ---             | ---                  | ---              | 10               |
|                        | Cwf26           | SPCC1620.10 | 1               | 1                    | 3                | 1                |
|                        | Cyp1            | SPAC37A10.03| 4               | 3                    | 7                | 7                |
|                        | Mup161          | SPAC1E3.09 | ---             | ---                  | 8                | 9                |
|                        | Cdc28           | SPBC19C2.01 | ---             | ---                  | ---              | 5                |
|                        | Prp43           | SPBC16H10.10c| ---            | ---                  | ---              | 9                |
|                        | Prp22           | SPAC2F8.02c| ---             | ---                  | 3                | 13               |
|                        | Saf4            | SPBC18H10.10c| 1               | 3                    | 11               | 11               |
|                        | Prp40/Up104     | SPACA4D7.13 | 1               | 1                    | ---              | ---              |
|                        | Usp105          | SPBC4H4.09 | 1               | 1                    | 2                | 1                |
|                        | Bbp1            | SPCC962.06c| 4               | 3                    | 3                | 11               |
|                        | Rsd1            | SPAC19G12.07c| 4               | 3                    | 3                | 11               |
|                        | Srp1            | SPAC29A1.08c| ---             | ---                  | 1                | 2                |
|                        | Dok1            | SPBC550.14c| ---             | ---                  | 5                | 5                |
|                        | Srp2            | SPAC16.02c  | 2               | 2                    | 1                | 7                |
|                        | Sum3            | SPCC1795.11 | 1               | 1                    | 3                | 1                |

\(^a\) Purifications done using 75 mM NaCl.
\(^b\) Peptide counts of identified proteins in each biological replicate (I and II).
\(^c\) —, no peptides identified.
\(^d\) UNK, an ortholog is unknown or not present.
guanidine-HCl (Fig. 6D). The NTE contains six tyrosines, two in the first 12 amino acids and the remaining four between residues 78 and 125 (Fig. 1B). Interestingly, these regions both correspond to predicted regions of order (Fig. 6A). Next, the protein was analyzed by CD spectrometry using far-UV wavelengths (Fig. 6E), which can be used to predict secondary structure. Analysis of this spectrum using the program K2D2 (48) predicts that Cwf10(1–135)His6 in solution is composed of 10% β-helix and 31% β-sheet, leaving over 50% of the NTE likely disordered (see Fig. S2B in the supplemental material).

To more carefully examine the tertiary structure of the NTE, we 15N labeled Cwf10(1–135)His6 and collected a two-dimensional 15N-1H heteronuclear single-quantum correlation (HSQC) experiment using nuclear magnetic resonance (NMR) spectroscopy. By this analysis, we found that the Cwf10 NTE contains 103 well-dispersed resonances (of a total of 141 residues) (Fig. 6F). When Cwf10(1–135)His6 is either heated to 50°C or treated with 6 M Guanidine-HCl in order to cause complete denaturation of any folded domain(s), the resonance peaks collapse and are no longer dispersed (see Fig. S2C and D in the supplemental material). Well-dispersed resonances in HSQC spectra result from the variable environment of the amines in a folded protein. Thus, the NMR analysis confirms both the computational and CD analyses indicating that the Cwf10 NTE contains regions of disorder, most

FIG 6 Biophysical characterization shows that the Cwf10 NTE contains regions of order and disorder. (A) Residues 18 to 61 of the NTE are predicted to be intrinsically disordered in solution using Disopred (80), a program that predicts structure disorder. The false-positive rate (FPR) was set at 5%. (B) Cwf10(1–135)His6 is monomeric by SVAU. The calculated s (c(s)) is plotted versus the sedimentation coefficient (s) for Cwf10(1–135)His6. The sedimentation velocity profiles were fitted to a continuous sedimentation distribution. (C) Near-UV CD spectrum of Cwf10(1–135)His6. (D) Near-UV CD spectrum of Cwf10(1–135)His6 in 6 M guanidine-HCl. (E) Far-UV CD spectrum of Cwf10(1–135)His6. (F) 15N-1H HSQC spectra of Cwf10(1–135)His6. (G) The first 23 amino acids of the Cwf10-NTE are required for efficient splicing. RNA was prepared from cwf10+/cwf10 ΔNTE, and cwf10+ cells grown at 25°C, and RT-PCR was done using random hexamer primers. (H) A portion of the tbp1 transcript was amplified from the RT-PCR products, and ratios of mRNA/pre-mRNA were calculated. Quantification represents the average of three biological replicates per genotype. Error bars indicate standard errors of the means.
likely in the first 60 amino acids (Fig. 6A), as well as regions that adopt a well-ordered secondary structure.

Interestingly, amino acids 1 to 23 of the NTE are 70% identical between S. pombe and human orthologs (Fig. 1B), and most of these residues (amino acids 1 to 17) are not predicted to be disordered (Fig. 6A). To test if this patch of conserved residues is required for NTE function, we replaced the chromosomal copy of cwf10\(^+\) with cwf10 2–23\(\Delta\) to observe whether this smaller truncation would also cause a pre-mRNA splicing defect. RNAs from wild-type, cwf10 2–23\(\Delta\), and cwf10-NTE cells were extracted, and RT-PCR was performed (Fig. 6G). Measures were taken to improve the quantitative nature of the PCR (83), including reverse transcribing similar amounts of RNA, reducing PCR cycles, and quantifying against an inherent internal control (the spliced and unspliced forms are amplified in the same reaction). The ratio of mature to premature signal for the tbp1\(_a\) intron was almost identical for the two truncations (2.1 ± 0.2 for cwf10-\(\Delta\)NTE versus 2.0 ± 0.3 for cwf10 2–23\(\Delta\)), while the ratio in the wild-type strain was 9.7 ± 1.5 (Fig. 6H). This suggests that the conserved, extreme N terminus (residues 1 to 23) is required for NTE function.

The Cwf10 NTE copurifies with splicing factors. The spliceosomal binding partners of the NTE have not been determined in any organism. We attempted to screen interactions by a yeast two-hybrid assay; however, the assay was hindered by the high self-activation of the GAL binding domain when fused to the acidic NTE. Therefore, we fused cwf10 2–135 to the N-terminal tandem affinity purification (TAP) tag in a pREP41 NTAP vector (35). We overexpressed NTAP-NTE in cwf10-\(\Delta\)NTE cells, performed a two-step purification at 150 mM NaCl, and analyzed the eluate using negative-stain electron microscopy (EM). Although the purification was clearly dilute, we did see particles that were reminiscent of negative-stain images of the S. pombe U5.U2/U6 complex (Fig. 7A) (10, 60), suggesting that the Cwf10 NTE is able to interact with spliceosomal complexes on its own. As an initial characterization of the TAP-NTE complex, we analyzed the snRNA content of the TAP-NTE purification from cwf10-\(\Delta\)NTE cells. As seen in Cdc5-TAP (10), TAP-NTE is associated with the U2, U5, and U6 snRNAs (Fig. 7B). While it is possible that a small amount of U1 snRNA is present in the purification (Fig. 7B), since no U1 snRNP components were found in the 1D LC-MS/MS analysis (Table 2) this interaction would have to be stoichiometric.

To further characterize the composition of the TAP-NTE purification, we analyzed the protein content of the eluate using 1D LC-MS/MS and compared it to Cdc5-TAP from wild-type cells under similar salt conditions. Peptide counts for spliceosomal proteins found in each LC-MS/MS run are included in Table 2 to provide a semiquantitative indication of protein amounts. After heat shock proteins (see Data set S3 in the supplemental material), the next-highest group of proteins identified in the purification comprised pre-mRNA splicing factors that are similar to the S. pombe U5.U2/U6 spliceosomal complex (10, 38) (Table 2). Highly represented in this group are the Sm proteins, components of the U5 snRNP (although Brr2 peptides were not detected), components of and related to the hPrp19/Cdc5L complex. Importantly, the EF2-like portion of Cwf10 (from the genomic cwf10-\(\Delta\)NTE allele) was also highly represented, confirming the ability of Cwf10-\(\Delta\)NTE to incorporate into higher-order complexes, as seen by sucrose gradient sedimentation (Fig. 4A). The presence of U2 snRNP proteins Lea1 and Msl1 indicates at least the partial presence of the U2 snRNP; however, the SF3 U2 snRNP components were not detected. Thus, the compositions of Cdc5-TAP and TAP-NTE are almost identical (Table 2) and show that the NTE can associate with a complex similar in composition to the late-stage U5.U2/U6 spliceosome.

To address whether we could detect Brr2 and SF3 peptides under less stringent purification conditions, the TAP-NTE purifications were repeated using 75 mM NaCl and analyzed by 1D LC-MS/MS (see Table S3 in the supplemental material). Once again, no peptides for either Brr2 or SF3 U2 snRNP components were detected (see Table S3 in the supplemental material), even though Brr2 peptides were detected in U5.U2/U6 complexes purified from wild-type cells using 75 mM NaCl (Table 1) and SF3 peptides were detected in U5.U2/U6 purifications from both wild-type and cwf10-\(\Delta\)NTE cells using 75 mM NaCl (Table 1).

In regard to the SF3 U2 snRNP components, our data suggest that while they may not be as stably associated with the U5.U2/U6 complex as Lea1 (S. cerevisiae Lea1, human U2A\(^\cdot\)), peptides for SF3 proteins were detected in Cdc5-TAP purifications from both wild-type and cwf10-\(\Delta\)NTE cells under mild salt conditions (Table 1). Conversely, no SF3 peptides were detected in the TAP-NTE purifications using either 150 or 75 mM salt (see Table S3 in the supplemental material). This shows that while the presence or absence of the NTE does not affect how SF3 components interact with the U5.U2/U6 complex, either the TAP-NTE complex does not contain SF3 proteins or they are present at substoichiometric levels.

To more closely examine whether the NTE associates with Brr2, we asked whether TAP-NTE can coimmunoprecipitate Brr2 in cwf10-\(\Delta\)NTE cells. To this end, either NTAP or NTAP-NTE was overexpressed in a cwf10-\(\Delta\)NTE brr2-3XHA strain. Immunoprecipitations were done using IgG-Sepharose beads and then immunoblotted with anti-HA antibodies to detect Brr2-HA. TAP-NTE was able to pull down Brr2-HA, while TAP alone was not (see Fig. S3 in the supplemental material), showing that Brr2 can associate with the TAP-NTE complex. However, one caveat of this experiment is that it does not replicate the two-step TAP protocol that was used for purifying the sample for mass spectrometry analysis. Therefore, it is possible that Brr2 is present after the first step of the purification but dissociates during the second step of the TAP. Combined, our analyses show that while TAP-NTE can coimmunoprecipitate Brr2 as detected by Western blot analysis, Brr2 is likely present at only substoichiometric levels in both TAP-NTE and Cdc5-TAP purifications from cwf10-\(\Delta\)NTE cells. Overall, our results demonstrate that TAP-NTE is able to bind to a surface(s) of the U5.U2/U6 core without being covalently linked to the EF2-like portion of Cwf10-NTE.

Because Cwf10-\(\Delta\)NTE copurified with spliceosomal components, we next asked whether the domain could restore pre-mRNA splicing efficiency in cwf10-\(\Delta\)NTE, i.e., function in trans. To test this, we performed semi quantitative RT-PCR on RNA extracted from cwf10-\(\Delta\)NTE cells overexpressing either NTAP or NTAP-NTE and looked at the splicing efficiency of two introns. Analysis of the tbp1\(_a\) and mrps16\(_b\) introns revealed that overexpression of NTAP-NTE in the cwf10-\(\Delta\)NTE background improved the mature/premature ratio over the NTAP control by ~25% and 30%, respectively, returning splicing efficiency close to wild-type conditions.
levels (Fig. 7C to E). Although only semiquantitative, the statistical significances of these data support a model in which the NTE can incorporate into spliceosomal complexes independent of any covalent connection with the C-terminal EF2-like portion of Cwf10, partially rescuing the splicing deficiency seen in cwf10−/H9004NTE cells.

Given the previous result that cwf10 2–127Δ (cwf10−ΔNTE) and cwf10 2–23Δ have almost identical splicing defects (Fig. 6G and H), we wondered whether overexpressing just the first 26 amino acids of the NTE would be sufficient to restore the pre-mRNA splicing efficiency of cwf10−ΔNTE cells. Therefore, we repeated the RT-PCR experiment, using RNA extracted from cwf10−ΔNTE cells overexpressing either NTAP or NTAP-cwf10 2–26. Although Western blotting confirmed the expression of the TAP-Cwf10 2–26 protein (data not shown), the smaller region was unable to complement deletion of the entire NTE (Fig. 7F to H). Thus, the
**TABLE 2** pre-RNA splicing factors copurifying with TAP-NTE in a *cwf10-ΔNTE* background and Cdc5-TAP in a wild-type background identified using 1D liquid chromatography-tandem mass spectrometry

| Spliceosome sub-complex | S. pombe protein | ORF number | TAP-NTE<sup>a</sup> | S. cerevisiae protein | H. sapiens protein |
|------------------------|-----------------|------------|----------------------|----------------------|------------------|
| **Core snRNP**         |                 |            |                      |                      |                  |
|                        | Smb1            | SPAC26A3.08| 4<sup>b</sup>        | 3                    | 16               | Smb1             | SMBB<sup>b</sup> |
|                        | Smd1            | SPAC27D7.07c| 1                    | 3                    | 12               | Smd1             | SMD1             |
|                        | Smd2            | SPAC2C4.03c| 6                    | 3                    | 15               | Smd2             | SMD2             |
|                        | Smd3            | SPAC19C2.14| 3                    | 1                    | 15               | Smd3             | SMD3             |
|                        | Sme1            | SPAC11T1.06c| 1                    | 1                    | 18               | Snc1             | SME1             |
|                        | Smt1            | SPBC3E7.14 | 5                    | 3                    | 11               | Smf1             | SMF1             |
|                        | Smg1            | SPAC4B1.05 | 6                    | 3                    | 22               | Smg2             | SMG1             |
| **U2**                 | Lea1            | SPBC1861.08c| 9                    | 3                    | 25               | Lea1             | U2A<sup>c</sup> |
|                        | Msf1            | SPBC8D2.09c| 1                    | ---                  | 10               | Msf1             | U2B<sup>d</sup> |
|                        | Prp10           | SPAC27F1.09c| ---                  | ---                  | 8                | ---              |                  |
| **U5**                 | Cwt10-ΔNTE      | SPBC215.12 | 60                   | 42                   | ---              | Snu14            | U5-116           |
|                        | Cw10            | SPBC215.12 | ---                  | ---                  | 116              | Snu14            | U5-116           |
|                        | Br2             | SPAC9.03c  | ---                  | ---                  | 4                | ---              | BR2              |
|                        | Sppt2           | SPAC4F8.12c| 57                   | 33                   | 221              | Prp8             | U5-220           |
|                        | Cwt17/Spf3     | SPBC1289.11| 9                    | 4                    | 53               | 21               | SMF1             |
| **NTC core**           | Cdc5            | SPAC644.12 | 34                   | 13                   | 99               | Cef1             | CDC5             |
|                        | Cw2             | SPAC3A12.11c| 13                  | 5                    | 52               | Cwc2             | RBBM22           |
|                        | Cw7             | SPBC28F2.04c| 10                  | 6                    | 42               | Cwc7             | SP27             |
|                        | Cwt15           | SPBC337.06c| 6                    | 1                    | 23               | Cwc15            | AD002            |
|                        | Prp5            | SPBP22H7.07 | 14                   | 8                    | 74               | Prp46            | PR1              |
|                        | Prp19           | SPAC29A4.08c| 18                   | 17                   | 93               | Prp19            | PR19             |
| **NTC-associated**     | Cwx3            | SPBC211.02c| 26                   | 9                    | 87               | Syf1             | SYF1             |
|                        | Cwx4            | SPBC31F10.11c| 21                   | 11                   | 63               | Syf1             | SYF1             |
|                        | Cwx14           | SPBC24C6.31  | 8                    | 4                    | 14                | Cwc14            | G10              |
|                        | Cwx18           | SPBP111.07c | 6                    | 5                    | 13                | ---              |                  |
|                        | Cwx19           | SPAC30D11.09 | 11                  | 3                    | 63               | ---              |                  |
|                        | Cwx14           | SPAC24A8.09c| 1                    | ---                  | 8                | 1                | Cwc21            | SMR3000          |
|                        | Cwx22           | SPAC13E7.01 | 2                    | 1                    | 20               | Cwc22            | hWC22            |
|                        | Cypl            | SPAC57A10.03 | 5                    | 1                    | 12               | ---              | PP1              |
|                        | Mug161          | SPAC1F3.09  | ---                  | ---                  | 8                | 1                | YGR093W          | CWF19L1          |
|                        | Ppr16           | SPBC1711.17 | ---                  | ---                  | ---              | ---              | PRP16            |
|                        | Pprp3           | SPBC16H5.10c| ---                  | ---                  | 3                | 3                | hPPR3            |
|                        | Pprp22          | SPAC10F6.02c| ---                  | ---                  | 6                | 6                | hPPR22           |
|                        | Sbf1            | SPBC81H10.10c| 3                   | 3                    | 13               | ---              |                  |
|                        | Sum3            | SPCC179S.15 | 7                    | 1                    | 1                | 3                | Ded1             | DDX3             |
| **Other**              | Unnamed         | SPAC20H4.09 | ---                  | ---                  | 5                | 3                | ---              |                  |

<sup>a</sup> Purifications done at 150 mM NaCl.
<sup>b</sup> Peptide counts of identified proteins in each biological replicate (I and II).
<sup>c</sup> ---, no peptides identified.
<sup>d</sup> UNK, an ortholog is unknown or not present.

Conserved region of amino acids 2 to 23 is necessary for splicing but not sufficient to complement *cwf10-ΔNTE* in *trans*.

**Discussion**

Deletion of the NTE and its effects on transcripts. In this study, we have investigated the function and structural characteristics of the Cwf10 NTE. Although this domain is not essential in fission yeast, deletion of the NTE causes a general splicing defect, a change in the sedimentation patterns of *in vivo* spliceosomal complexes, and synthetic lethal and synthetic sick interactions with mutant alleles of other pre-mRNA splicing factors.

Because *cwf10-ΔNTE* cells show robust growth at all temperatures, we were surprised to detect a pre-mRNA splicing defect and postulated that the lack of the NTE domain may affect the splicing
of a specific subset of pre-mRNAs rather than cause a global pre-
mRNA splicing defect. To test if this was the case, we used deep
sequencing analysis to comprehensively determine which tran-
scripts are affected by this mutation. Interestingly, most introns
show a tendency toward reduced splicing efficiency (all points to
the right of the diagonal in Fig. 3A), and the reduced splicing is
statistically significant for about 44% of total introns. It is not
 surprising, then, that there is no obvious type of transcript (i.e.,
specific splice site, number of introns, and/or size of introns) that
is specifically affected. Thus, the splicing defect in *cwf10*Δ*nte*
cells appears to be global in nature and widespread in scope. This
suggests that the NTE is important for a general (not transcript-
specific) process in the splicing reaction.

Because functions have not been assigned for any subregion of
the NTE, we removed just the first 23 amino acids from the NTE in
Cdc20 and found a similar splicing defect between *cwf10*Δ*nte*
and *cwf10*Δ2–23Δ. We speculate that loss of this small, highly con-
served, acidic region undermines the NTE’s ability to bridge in-
teractions or support conformational rearrangements in the spli-
cosome critical for the NTE’s function (see “Model of Cwf10
NTE interactions in the spliceosome” below).

**The NTE’s roles in the splicing cycle.** A previous study using *S.
cerevisiae Snu114Δ* demonstrated a role for the Snu114/Cwf10
NTE in U4/U6 snRNA unwinding (14). However, a study using human
splicing extracts and antibodies directed against the hu-
m humans 5-116K NTE suggested that the NTE may also be involved in the first- to second-step transition (12). This NTE function,
uncovered in the human system, is further supported by negative
genetic interactions (30) between *S. cerevisiae Snu114Δ* and both a 
substrate at snRNA U6-A59, which inhibits the second step of splicing (84), and several alleles of the U5 loop 1, which helps
position the exons for ligation in the second step (85, 86).

Our genetic findings are consistent with a role for the NTE in both spliceosome activation and second-step catalysis. *cwf10-
nte* interacts with several allelic genes known to be involved in spliceosome activation (Fig. 5), including *spp42*Δ (S. *cerevisiae
PRP8*), *prp1*Δ (S. *cerevisiae PRP6*), and *cd5*Δ120 (S. *cerevisiae
CEF1*). Additionally, the cd5Δ120 mutation, with which *cwf10-
nte* is synthetically lethal, may play an additional role in sec-
ond-step chemistry. First, *cd5*Δ120 is lethal with *s. pombe prp17A*
(38), a known second-step splicing factor in *S. cerevisiae* (87). Second, *cd5*Δ120 is a point mutation in one of the two conserved
Myb repeats (88), a region important for first-step to second-step
modulation in ortholog *S. cerevisiae Cef1* (75). The fact that the
NTE does not immunoprecipitate a preactivation spliceosome
(Table 2), but rather a complex similar to U5/U2/U6, could indi-
cate that the NTE becomes stably “locked” into the spliceosome
following activation and is positioned to act near the catalytic core
to help modulate that transition.

Our analysis is consistent with the NTE playing a role in stabi-
lizing U5 snRNP integrity during spliceosome transitions, a model
supported by the lack of Br2 peptides found in both Cdc5-CAP
and TAP-NTE purifications from the *cwf10*Δ*nte* backgrounds
(Table 1 and 2; see also Table S3 in the supplemental material)
and the increase in the amount of smaller U5 snRNPs complexes in
*cwf10*Δ*nte* cells (Fig. 4G and H). It has been proposed that in *S.
cerevisiae Snu114* acts as a transducer that signals and regulates
Br2 activity throughout the splicing cycle (16), although the mechanism(s) of how Snu114 could modulate Br2 activity is not
yet understood on a molecular level. One possibility is that the
NTE acts as a flexible scaffold stabilizing Br2’s interaction with
the U5/U2/U6 complex during the structural rearrangements that
occur during spliceosome activation. When the NTE is missing,
Br2 is not able to remain as tightly associated with the spliceo-
some during these transitions, affecting the overall integrity of the
U5 snRNP and leading to reduced pre-mRNA splicing efficiency.

**Alterations in sizes of snRNA-containing complexes.** Analysis of snRNAs in cells lacking the NTE shows that loss of this
domain changes the sedimentation patterns of spliceosomal com-
plexes. Although all the snRNA sedimentation patterns show
some degree of change in the *nte* background (Fig. 4E to N),
for both the U1 and U5 snRNAs there is a shift away from higher-
molecular-mass fractions into lower-molecular-mass fractions
(<11.3S [Fig. 4H and L]). For the U5 snRNA, this shift is too high in the gradient (<11.3S [Fig. 4H]) to contain the full complement
of U5-snRNP specific proteins, as both human and *S. cerevisiae U5
snRNPs sediment at 15 to 20S when bound to the U5-specific proteins (89, 90). Similar results with U5 snRNA sedimentation in the
*S. cerevisiae snu114Δ* background led the authors to specu-
late that the NTE is required for U5 snRNP stability (14), and our
results are consistent with that hypothesis. For the U1 snRNA, it is
unclear whether this lower-molecular-mass sedimenting fraction
represents free U1 snRNA or the snRNA complexed with at least
some of the U1 snRNP proteins.

Sucrose gradient analysis also revealed that a small fraction of
the U2, U5, U6, and U4 snRNAs, as well as Cdc20, shifted to
high-molecular-mass fractions that are more pronounced in
*cwf10*Δ*nte* (fractions 11 and 12 [Fig. 4B, F, H, J, and N]). The
presence of this peak could suggest that a preactivation spliceo-
some is stalled and accumulating. Alternatively, this peak could
represent a different multi-snRNP complex or aggregates of spliceo-
some components that are not able to organize properly.

Although the sedimentation data support a model in which the
*in vivo cwf10*Δ*nte* splicing complexes that are slow to activate or complete catalysis may be accumulating or aggregating, other pos-
sible fates for stalled spliceosomes include degradation by the pro-
teasome and/or disassembly. The lower steady-state levels of Cdc5
and Prp1-myc (Fig. 2I), which sediment mainly as part of large complexes in *cwf10*Δ*nte* (Fig. 4C and D and data not shown),
could favor the hypothesis that these slow-to-splice complexes are
degraded.

**The NTE’s structure.** Although the NTE was predicted to be
intrinsically unstructured in solution (26), our modeling and bio-
physical analysis unexpectedly showed that approximately one-
half of NTE residues are in a folded environment in solution.
Because the Disopred program (Fig. 6A) predicted disorder to
exist mostly in the N terminus (aa 18 to 61) of the NTE and the
ordered amino acids to exist at the extreme N terminus (aa 1 to 17)
and the C terminus (aa 62 to 120), it is tempting to begin thinking of
the NTE as three subregions with separate characteristics. The
extreme N terminus, which is not predicted to be disordered, is
required for function, since deleting 23 conserved residues in the
N terminus fully recapitulates the splicing defect seen when the
entire NTE is deleted (Fig. 6G and H). However, this small region
cannot complement the *cwf10* mutant lacking the NTE in *trans*,
suggesting that the C-terminal NTE regions are also important.
The region spanning amino acids 1 to 17 is physically linked to aa
18 to 61, which is predicted to be disordered and has a high con-
tent of acidic residues, as does the aa 1 to 17 region. The acidic
charge of the NTE is likely essential since we were unable to inte-
grante a cwfl0 mutant that had all negatively charged residues in aa 1 to 61 replaced with alanine residues (data not shown). Possibly, this unstructured subregion of the NTE becomes folded in the context of the dynamic spliceosome when it comes into contact with a binding partner(s). This idea has precedent in pre-mRNA splicing. A 70-amino-acid intrinsically disordered region of the NTC-associated protein human SKIP undergoes a disordered-to-ordered transition upon binding to cyclophilin PPIL1 (91). Similarly, a 31-amino-acid, predominantly random coil region of human U4/U6-60K adopts structure when bound to U4/U6-20K (another cyclophilin) (92). However, the disordered regions of human SKIP or U4/U6-60K do not share the same high acidic content of Cwf10-NTE aa 1 to 61. Next, a predicted structured portion of the NTE, amino acids 62 to 120, links the N-terminal NTE regions to the “EF2-like” portion of the protein. Further studies are needed to delineate the boundaries of order and disorder within the NTE, as well as to determine binding partners of these subregions.

Potential binding environment in the spliceosome. When overexpressing TAP-NTE (Cwf10 residues 2 to 135) in cwfl0-ΔNTE cells, we found that this fragment is able to incorporate into a spliceosomal complex that is similar in composition to the S. pombe U5.U2/U6 complex. From this observation, we conclude that the Cwf10-NTE recognizes and interacts with other spliceosomal components independently of its polypeptide linkage to the main Cwf10 “EF2-like” portion. What does the NTE bind to in the context of the late-stage spliceosome? Likely, the binding site does not involve Br2, because it is not present at stoichiometric amounts in the purification (Table 2), although other U5-specific proteins, i.e., Spp42 (S. cerevisiae Prp8), Spf38 (human U5-40), and the EF2-like portion of Cwf10, are highly represented. Therefore, the binding partner(s) is likely one of the above U5-specific proteins, the U5 Sm core, U2 snRNP protein Lea1, or one or more of the NTC and NTC-related proteins. Additionally, the specific genetic interactions of S. cerevisiae snu114ΔN with snRNA alleles (30) may also support an orientation of the NTE near U5 loop 1 and U6. Indeed, the U5 snRNA loop 1 has been shown to be a binding platform for S. cerevisiae Snu114, Br2, and Prp8 during U5 snRNP assembly (93). Importantly, NTE binding partners may change as protein-protein, protein-RNA, and RNA-RNA re-arrangements occur during the splicing reaction. The combination of structured and unstructured regions within the NTE may allow the domain to bridge transient conformational states adopted by the spliceosome as splicing occurs. However, regardless of the exact mechanism, the presence of the Cwf10 NTE in trans is sufficient to partially rescue the splicing defect seen in cwfl0-ΔNTE cells. There are other spliceosome proteins that can function when their domains are expressed separately, highlighting the multiple interactions between spliceosome components. Examples include the large, 280-kDa S. cerevisiae Prp8 (18) and first-step factor S. cerevisiae Yju2 (94).

Model of Cwf10 NTE interactions in the spliceosome. Based on our data and the work of others outlined above, we propose a model of NTE function that includes (i) stabilization of the U5 snRNP structure and (ii) facilitation of certain conformations among U5 snRNP components or between the U5 snRNP and other spliceosome proteins/RNAs. The NTE is able to independently incorporate into the spliceosome and likely makes contact with more than one splicing component through its predicted structured and unstructured regions. Although we were not able to determine specific binding partners for the Cwf10 NTE, these partners must be in the vicinity of the EF2-like portion of Cwf10. It is tempting to speculate that these contacts promote cohesion within the U5 snRNP and may transmit a signal or form a stabilizing bridge, first within the context of the preactivation B complex and then near the catalytic center of the C complex.

ACKNOWLEDGMENTS

We thank N. F. Käuffer for sharing S. pombe strains, Kathy Gould and the Gould lab for generous contributions of strains, reagents, protocols, and discussion, Liping Ren and Anna Feoktistova for valuable technical assistance, and Hayes McDonald for assistance with mass spectrometry analysis. We thank Ohl lab members and Kate Mittendorf for helpful discussions and comments on the manuscript.

This work was supported by T32 CA09582 to S.B.L., T32 GM03820 to S.E.C., and NIH DP2OD004483 to M.D.O., as well as by a Wellcome Trust Senior Investigator Award to J.B. Recombinant His6-Cwf10 (673-983) and anti-Cwf10 antibodies were produced and purified by the Vanderbilt Antibody and Protein Resource, which is supported by the Vanderbilt Institute of Chemical Biology and the Vanderbilt Ingram Cancer Center (P30 CA68485).

REFERENCES

1. Wahl MC, Will CL, Lührmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. Cell 136:701–718.
2. Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. 2009. The evolutionarily conserved core domain of the catalytic activation step of the yeast spliceosome. Mol. Cell 36:593–608.
3. van der Feltz C, Anthony K, Briolot A, Pomeranz Krummel DA. 2012. Architecture of the spliceosome. Biochemistry 51:3321–3333.
4. Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 92:315–326.
5. Will CL, Lührmann R. 2011. Spliceosome structure and function. Cold Spring Harb. Perspect. Biol. 3(7):pii:a003707. doi:10.1101/cshperspect.a003707.
6. Jurisics MS, Moore MJ. 2003. Pre-mRNA splicing: awash in a sea of proteins. Mol. Cell 12:5–14.
7. Lardelli RM, Thompson JX, Yates JR, III, Stevens SW. 2010. Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. RNA 16:516–528.
8. Agafonov DE, Deckert J, Wolf E, Odenwälder P, Bessonov S, Will CL, Urlaub H, Lührmann R. 2011. Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. Mol. Cell. Biol. 31:2667–2682.
9. Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R. 2008. Isolation of an active step I spliceosome and composition of its RNP core. Nature 452:846–850.
10. Ohi MD, Link AJ, Ren L, Jennings JL, McDonald WH, Gould KL. 2002. Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Gcd3p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. Mol. Cell. Biol. 22:2001–2024.
11. Warocki Z, Odenwälder P, Schmitzova J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio P, Lührmann R. 2009. Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. Nat. Struct. Mol. Biol. 16:1237–1243.
12. Fabrizio P, Laggerbauer B, Lauber J, Lane WS, Lührmann R. 1997. An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. EMBO J. 16:4092–4106.
13. Achsel T, Ahrens K, Brahms H, Teigelkamp S, Lührmann R. 1998. The human U5-220kD protein (Hprp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor-2, and a novel WD-40 protein. Mol. Cell. Biol. 18:6756–6766.
14. Bartels C, Klatt C, Lührmann R, Fabrizio P. 2002. The ribosomal translocase homologue Snu14p is involved in unwinding U4/U6 RNA during activation of the spliceosome. EMBO Rep. 3:875–880.
15. Brenner TJ, Guthrie C. 2006. Assembly of Snu114 into U5 snRNP requires Prp8 and a functional GTPase domain. RNA 12:862–871.
16. Small EC, Leggett SR, Winans AA, Staley JP. 2006. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brtz2p, a DEK/D/H box ATPase. Mol. Cell. 23:389–399.

17. Liu S, Rauhut R, Vornlocher HP, Lührmann R. 2006. The network of protein-protein interactions within the human U4/U6.U5 tri-snRNPs. RNA 12:1718–1727.

18. Boon KL, Norman CM, Grainger RJ, Newman AJ, Beggs JD. 2006. Prp8p dissection reveals domain structure and protein interaction networks. RNA 12:198–205.

19. Brenner TJ, Guthrie C. 2005. Genetic analysis reveals a role for the C terminus of the Saccharomyces cerevisiae GTPase Snu114p during spliceosome activation. Genetics 170:1063–1080.

20. Zhang L, Xu T, Maeder C, Bud LO, Shanks J, Nix J, Guthrie C, Pleiss JA, Zhao R. 2009. Structural evidence for conserved He1308-like modules in the spliceosomal ATPase Brtz2. Nat. Struct. Mol. Biol. 16:731–739.

21. Dix I, Russell CS, O’Keefe RT, Newman AJ, Beggs JD. 1998. Protein-RNA interactions in the U5 snRNP of Saccharomyces cerevisiae. RNA 4:1675–1686.

22. Grainger RJ, Beggs JD. 2005. Prp8 protein: at the heart of the spliceosome. RNA 11:533–557.

23. Turner IA, Norman CM, MJ, Newman AJ, Newman AJ. 2006. Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. RNA 12:375–386.

24. Maeder C, Kutach AK, Guthrie C. 2009. ATP-dependent unwinding of U4/U6 snRNAs by the Brtz2 helicase requires the C terminus of Prp8p. Nat. Struct. Mol. Biol. 16:462–48.

25. Bayne EH, Portoso M, Kaganovsky A, Kos-Braun IC, Uruno T, Ekwall K, Alves F, Rappisiller J, Allshire RC. 2008. Splicing factors facilitate RNA-directed silencing in fission yeast. Science 322:602–606.

26. Korneta I, Magnus M, Bujnicki JM. 2012. Structural bioinformatics of the human spliceosomal protein. Nucleic Acids Res. 40:7046–7065.

27. Coelho Ribeiro ML, Espinosa J, Islam S, Martínez O, Romfo CM, Lakhe-Reddy S, Wise JA. 2010. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5:621–628.

28. Agresti A. 1990. Categorical data analysis. Wiley, New York, NY.

29. Storey JD. 2002. A direct approach to false discovery rates. J. Royal Stat. Soc. Ser. B 64:479–498.

30. Perez-Iturxeta C, Andrade-Navarro MA. 2008. KD2: estimation of protein secondary structure from circular dichroism spectra. BMC Struct. Biol. 8:25, doi:10.1186/1472-6807-8-25.

31. Schaffer P. 2000. Site-directed mutagenesis analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys. J. 78:1606–1619.

32. MacCoss MJ, McDonald WH, Saraf A, Sadovy R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, Clark AJ, Yates JR, III. 2002. Shotgun identification of protein modifications from protein complexes and lens tissue. Proc. Natl. Acad. Sci. U. S. A. 99:7900–7905.

33. Pappin DJ, III, Enns KG, Scott MK. 2001. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal. Chem. 67:1426–1436.

34. Ohi M, Li Y, Cheng Y, Walz T. 2004. Negative staining and image classification—powerful tools in modern electron microscopy. Biol. Proced. Online 6:23–34.

35. Habara Y, Uruushiyama S, Shibuya T, Ohshima Y, Tani T. 2001. Mutation in the prp42 gene encoding a homolog of SAP180/SF3B1 causes differential inhibition of pre-mRNA splicing and arrest of cell-cycle progression in Saccharomyces pombe. RNA 7:597–610.

36. Tsai WY, Chow YT, Chen HR, Huang KT, Hong RI, Jan SP, Ko NU, Tsao TY, Chen CH, Cheng SC. 1999. Cel1p is a component of the Prp19-associated complex and essential for pre-mRNA splicing. J. Biol. Chem. 274:9455–9462.

37. Ohi MD, Gould KL. 2002. Characterization of interactions among the Celf1p-Prp19p-associated splicing complex. RNA 8:798–815.

38. Stevens SW, Abelson J. 1999. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proc. Natl. Acad. Sci. U. S. A. 96:7226–7231.

39. Romfo CM, Lakhe-Reddy S, Wise JA. 1999. Molecular genetic analysis of U2AF59 in Saccharomyces cerevisiae. RNA 5:471–483.

40. Hilliard RS, Jiang H, Storey RB, Wolters D, Washburn M, Weiss A, Clark AJ, Yates JR, III. 2002. Shotgun identification of protein modifications from protein complexes and lens tissue. Proc. Natl. Acad. Sci. U. S. A. 99:7900–7905.

41. Andrieux A, Whitworth GB, Bergessel M, Guthrie C. 2007. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosome components. PLoS Biol. 5:e90, doi:10.1371/journal.pbio.0050090.

42. Sridharan V, Heimiller J, Singh R. 2011. Genomic mRNA profiling reveals compensatory mechanisms for the requirement of the essential splicing factor U2AF, Mol. Cell. 31:652–661.

43. Ohi MD, Ren L, Wall JS, Gould KL, Walz T. 2007. Structural characterization of the fission yeast U12/U6 spliceosome complex. Proc. Natl. Acad. Sci. U. S. A. 104:3195–3200.

44. Huang T, Vilardell J, Query CC. 2002. Pre-spliceosome formation in S. pombe requires a stable complex of SF1-U2AF59-U2AF23. EMBO J. 21:5516–5526.

45. Lützelberger M, Bottner CA, Schwelmus W, Zock-Emmenthal S, Razanay A, Käfer NF. 2010. The N-terminus of Prp1 (Prp6/U5-102 K)
is essential for spliceosome activation in vivo. Nucleic Acids Res. 38:1610–1622.
63. Carnahan RH, Feoktistova A, Ren L, Niessen S, Yates JR, III, Gould KL. 2005. Dim I p is required for efficient splicing and export of mRNA encoding lid I p, a component of the fission yeast anaphase-promoting complex. Eukaryot. Cell 4:577–587.
64. Behrens SE, Galisson F, LeGrain P, Lührmann R. 1993. Evidence that the 60-kDa protein of U2 small nuclear ribonucleoprotein is immunologically and functionally related to the yeast PRP9 splicing factor and is required for the efficient formation of prespliceosomes. Proc. Natl. Acad. Sci. U. S. A. 90:8229–8233.
65. Behrens SE, Tyc K, Kastner B, Reichelt J, Lührmann R. 1993. Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. Mol. Cell. Biol. 13:307–319.
66. Brosi R, Hauri HP, Kramer A. 1993. Separation of splicing factor SF3 into two components and purification of SF3a activity. J. Biol. Chem. 268:17640–17646.
67. Pauling MH, McPheeters DS, Ares M, Jr. 2000. Functional Cus1p is found with Hsh155p in a multiprotein splicing factor associated with U2 snRNA. Mol. Cell. Biol. 20:2176–2185.
68. Urushiyama S, Tani T, Ohshima Y. 1996. Isolation of novel pre-mRNA splicing mutants of Schizosaccharomyces pombe. Mol. Gen. Genet. 253:118–127.
69. Nurse P, Thuriaux P, Nasmyth K. 1976. Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. 146:167–178.
70. Rosenberg GH, Alahari SK, Kaufer NF. 1991. prp4 from Schizosaccharomyces pombe, a mutant deficient in pre-mRNA splicing isolated using genetic criteria. Mol. Gen. Genet. 226:305–309.
71. Kim DU, Hayles J, Kim D, Wood V, Park HO, Won M, Yoo HS, Duhiug T, Nam M, Palmer G, Han S, Jeffery L, Baek ST, Lee H, Shim YS, Lee M, Kim L, Heo KS, Nob EJ, Lee AR, Jang YJ, Chung KS, Choi SJ, Park JY, Park Y, Kim HM, Park SK, Park HI, Kang EI, Kim HB, Kang HS, Park HM, Kim K, Song K, Song KB, Nurse P, Hoe KL. 2010. Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe. Nat. Biotechnol. 28:617–623.
72. Nasmyth K, Nurse P. 1981. Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. 182:119–124.
73. Kuhn AN, Li Z, Brow DA. 1999. Splicing factor Prp8 governs U4/U6 RNA unwinding during splicing activation. Mol. Cell. 3:65–75.
74. Chan SP, Kao DI, Tsai WY, Cheng SC. 2003. The Prp19p-associated complex in spliceosome activation. Science 302:279–282.
75. Query CC, Konarska MM. 2012. CEF1/CDC5 alleles modulate transitions between catalytic conformations of the spliceosome. RNA 18:1001–1013.
76. Rappsilber J, Ryder U, Lamond AI, Mann M. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res. 12:1231–1245.
77. Arenas JE, Abelson JN. 1997. Prp43: an RNA helicase-like factor involved in spliceosome disassembly. Proc. Natl. Acad. Sci. U. S. A. 94:11798–11802.
78. Martin A, Schneider S, Schwer B. 2002. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. J. Biol. Chem. 277:17743–17750.
79. Tsai RT, Fu RH, Yeh FL, Tseng CK, Lin YC, Huang YH, Cheng SC. 2005. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes Dev. 19:2991–3003.