BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

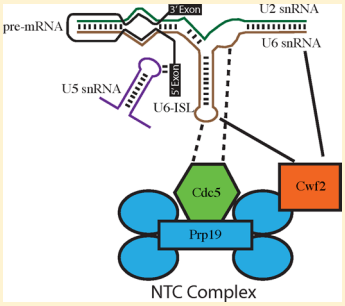
Structural and Functional Insights into the N-Terminus of *Schizosaccharomyces pombe* Cdc5

Scott E. Collier,[†] Markus Voehler,[‡] Dungeng Peng,[§] Ryoma Ohi,[†] Kathleen L. Gould,[†] Nicholas J. Reiter,[§] and Melanie D. Ohi*[,†]

[†]Department of Cell and Developmental Biology, Vanderbilt University Medical Center, Nashville, Tennessee 37232, United States
[‡]Department of Chemistry and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States
[§]Department of Biochemistry, Vanderbilt University Medical Center, Nashville, Tennessee 37235, United States

Supporting Information

ABSTRACT: The spliceosome is a dynamic macromolecular machine composed of five small nuclear ribonucleoparticles (snRNPs), the NineTeen Complex (NTC), and other proteins that catalyze the removal of introns mature to form the mature message. The NTC, named after its founding member *Saccharomyces cerevisiae* Prp19, is a conserved spliceosome subcomplex composed of at least nine proteins. During spliceosome assembly, the transition to an active spliceosome correlates with stable binding of the NTC, although the mechanism of NTC function is not understood. *Schizosaccharomyces pombe* Cdc5, a core subunit of the NTC, is an essential protein required for pre-mRNA splicing. The highly conserved Cdc5 N-terminus contains two canonical Myb (myeloblastosis) repeats (R1 and R2) and a third domain (D3) that was previously classified as a Myb-like repeat. Although the N-terminus of Cdc5 is required for its function, how R1, R2, and D3 each contribute to functionality is unclear. Using a combination of yeast genetics, structural approaches, and RNA binding assays, we show that R1, R2, and D3 are all required for the function of Cdc5 in cells. We also show that the N-terminus of Cdc5 binds RNA *in vitro*. Structural and functional analyses of Cdc5-D3 show that, while this domain does not adopt a Myb fold, Cdc5-D3 preferentially binds double-stranded RNA. Our data suggest that the Cdc5 N-terminus interacts with RNA structures proposed to be near the catalytic core of the spliceosome.



The NTC (NineTeen Complex), named after its founding member *Saccharomyces cerevisiae* (*Sc*) Prp19, is a conserved protein-only spliceosome subcomplex that has been isolated in both human and *S. cerevisiae* splicing extract systems.[1−3] Stable binding of the NTC is required for the formation of a catalytically active spliceosome that is competent to precisely remove introns from precursor mRNA (pre-mRNA) to form mature message (mRNA). Although the exact NTC composition varies among eukaryotes, at least nine proteins are conserved in yeast and are often referred to as the core NTC complex. For both *S. cerevisiae* and *Schizosaccharomyces pombe* (*Sp*), these include the proteins Prp19/Prp19 (*Sc*/*Sp*), Cef1/Cdc5, Prp46/Prp5, Clf1/Cwf4, Syf2/Cwf3, Cwc15/Cwf15, Isy1/Cwf12, Snt309/Cwf7, Cwc2/Cwf2, and *Sc* Ntc20.[1,4−9] In mammalian cells, the NTC is called the hPrp19/CDC5L complex and is composed of a similar, but not identical, set of proteins. These include hPrp19 (Prp19/Prp19 (*Sc*/*Sp*)), CDC5L (Cef1/Cdc5), PRL1 (Prp46/Prp5), AD002 (Cwc15/Cwf15), SPF27 (Snt309/Cwf7), HSP73 (Ssa4/Sks2), and CTNNBL1 (*Sp* SPAC1952.06c).[2,3]

Although it is clear that the NTC is essential for the spliceosome to transition from an inactive to active complex,[10,11] as well as ensuring both the fidelity and efficiency of the splicing reaction,[12,13] the molecular mechanism(s) of NTC function is not known. One model is that the NTC acts as a molecular scaffold, supporting and/or facilitating essential RNA−RNA, RNA−protein, and protein−protein rearrangements that are required for the formation of a catalytically active spliceosome. In support of this model, a number of NTC components contain characteristic protein−protein interaction domains that include WD40 repeats, TPR (tetratricopeptide) repeats, and HAT (half a TPR domain) repeats. Additionally, two conserved components, Cef1/Cdc5 (*Sc*/*Sp*) and Cwc2/Cwf2 (*Sc*/*Sp*), contain Myb (myeloblastosis) repeats and a RRM (RNA recognition motif), respectively, serving as potential nucleic acid binding domains.[14,15] While the zinc-finger (ZnF) and RRM in *Sc* Cwc2 and its mammalian homologue RBM22 cross-links directly to the U6 snRNA and pre-mRNA,[15,16] the biochemical function of the Cef1/Cdc5 (*Sc*/*Sp*) Myb repeats has not been determined.

S. pombe cdc5^{+} was first identified in a screen of fission yeast mutants defective for cell cycle progression,[17] and subsequent studies have shown that *S. pombe* Cdc5 is an essential member of the spliceosome and functions in pre-mRNA splicing,[18−20] a role conserved in other organisms.[2,3,5,19,21−23] In addition to its essential function in pre-mRNA splicing, Cdc5 proteins in various organisms have also been implicated in transcription,[24]

Received: July 11, 2014
Revised: September 17, 2014
Published: September 29, 2014

   dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

DNA damage response,[25−28] mitotic spindle assembly,[29] and microRNA (miRNA) biogenesis,[30] although whether these diverse cellular activities are splicing-dependent or -independent has not been fully determined.

The N-terminus of Cdc5 family members contains two canonical Myb repeats (R1 and R2) and a third conserved domain (D3) previously classified as a Myb-like repeat[31] (MLR, also referred to as Cdc5-MLR3 and -MYB3[31,32]). Myb repeats are classically considered to be DNA binding motifs, and multiple copies of these domains are often found in transcription factors (reviewed in ref 33). Structurally, the Myb domain is composed of three well-defined helices, with the second and third helices adopting a fold similar to that of the canonical helix-turn-helix motif (HTH).[34] Despite its name, a Myb repeat is not an integrated tertiary motif in a larger domain, but rather an independently folded domain that is often found in multiple copies within proteins. Although the N-terminus of Cdc5 family members has been shown to interact with DNA *in vitro*[14,35,36] and is required for cell viability,[5,14] the function of the Cdc5 N-terminus and its individual domains in pre-mRNA processing has yet to be determined. Unlike the N-terminus, the sequence of the Cdc5 C-terminus is not conserved and has no recognizable protein motifs;[31] however, it interacts directly with several NTC core components.[3,8] The presence of nucleic acid binding domains in Cdc5 and its direct interaction with other core NTC members has led us to hypothesize that Cdc5 may facilitate NTC-mediated RNA−RNA and/or RNA−protein transitions by acting as a scaffold linking NTC components and RNAs, similar to what was observed with RRM containing RBM22/Cwc2 family members in humans and *S. cerevisiae*.[15,16,37]

To further characterize how the NTC stabilizes and activates the spliceosome, we investigated the role of the N-terminus of Cdc5 in cell function and its ability to bind RNA *in vitro*. Using a combination of yeast genetics and RNA binding assays, we show that R1, R2, and D3 are all required for function and that the Cdc5 N-terminus binds RNA *in vitro*. Structural and biochemical analyses of Cdc5-D3 show that, unlike what has been predicted,[31] this domain does not adopt a canonical Myb fold and that Cdc5-D3 preferentially binds double-stranded RNA *in vitro*. Our data demonstrate that the Cdc5 N-terminus (R1, R2, and D3) can function as an RNA binding platform and can directly interact with RNA structures found near the catalytic core of the spliceosome. Our results support a model where Cdc5, by interacting with both NTC proteins and RNA, serves as an important scaffold that facilitates the conformational changes required for the formation of a catalytically active spliceosome.

MATERIALS AND METHODS

Strains, Yeast Methods, and Molecular Biology. Strains and plasmids used in this study are listed in Tables SI and SII. Yeast strains were grown in yeast extract (YE) media or Edinburgh minimal media (EMM) with appropriate supplements. A plasmid containing the *LEU2* marker (pIRT2), the ORF of $cdc5^{+}$, and at least 500 base pairs (bp) of the 5′and 3′ flanking sequence was used to generate $cdc5^{\Delta R1}$, $cdc5^{\Delta R2}$, $cdc5^{\Delta D3}$, $cdc5^{\Delta L}$, $cdc5^{\Delta MID}$, and $cdc5^{\Delta D3::Myb}$ integration plasmids using QuikChange II (Agilent Technologies, Santa Clara, CA). All vector transformations were performed as previously described.[38] Both control and *cdc5* deletion plasmids were transformed into a diploid strain of *S. pombe*, $cdc5^{+}/cdc5$::$ura4^{+}$. Transformations were grown on minimal medium lacking leucine, adenine, and uracil. Colonies were grown in EMM (−nitrogen) to induce sporulation, and haploid cells were grown on minimal media (+ adenine) to select for haploid cells that were $cdc5$::$ura4^{+}$ and carried the pIRT2 plasmid. Stable integrants were selected on the basis of resistance to 5-fluoroorotic acid (5-FOA)[39] and the acquisition of the LEU^{-} phenotype. Mutants were validated by whole-cell PCR with primers 5′ and 3′ of the *cdc5* gene. Deletions in pREP3X $cdc5^{+}$ (cDNA) vectors were generated as above and transformed to a haploid strain of *S. pombe*, *cdc5-TAP*. Transformants were grown on minimal media lacking leucine and containing thiamine (30 μM) for pREP3X repression or no thiamine for pREP3x induction. OD30 lysate western blots with anti-pSTAIR and anti-Cdc5 were performed as previously described.[39] For spot assays, cells were grown to mid log phase at 25 °C and resuspended in water to achieve an OD_{595} of 0.3. Ten-fold serial dilutions were made, and 2 μL of each dilution was plated on YE. Plates were incubated at the indicated temperatures for 3−5 days before imaging. Protparam[40] was used to calculate pI's of protein domains.

Protein Expression and Purification. Cdc5-R1-R2-$D3^{\Delta L}$ (amino acids (aa) 5−$208^{\Delta 111-146}$), Cdc5-R1-R2 (aa 5−112), Cdc5-R1 (aa 5−55), Cdc5-R2 (aa 58−111), and Cdc5-D3 (aa 155−214) (*S. pombe* Cdc5, NP_593880) were cloned into pET-15b (NdeI/BamHI) (EMD Millipore, Darmstadt, Germany) and transformed into *Escherichia coli* Rosetta 2 (DE3) pLysS cells (EMD Millipore, Darmstadt, Germany). Cells were grown in terrific broth (Invitrogen, Grand Island, NY) to an OD_{595} of ∼0.9 and cold shocked for 20 min on ice. Upon addition of 1 mM IPTG, the plasmids were overexpressed for 20 h at 15 °C. Cells were lysed in 25 mM MES (pH 6.0), 300 mM NaCl, 2.5 mM imidazole, 5% glycerol, 0.1% Triton X-100, and one SIGMAFAST protease tablet (Sigma-Aldrich, St. Louis, MO). Cdc5 constructs were purified using two 5 mL Histrap HP columns (GE Healthcare, Waukesha, WI) in 50 mM MES (pH 6.0), 500 mM NaCl, 5% glycerol, and a 2.5−1000 mM imidazole linear gradient. After the Histrap column, the protein fractions were concentrated and buffer exchanged into heparin buffer A (10 mM sodium phosphate, pH 7.0, 1 mM EDTA, and 5% glycerol) using a 3K Amicon Ultra-15 filter (Millipore, Billerica, MA). The pooled fractions were treated overnight at room temperature with RECOthrom (The Medicines Company, Parsippany, NJ) to cleave the His_6 tag. Cdc5 constructs were further purified using a heparin column (GE Healthcare, Waukesha, WI) in 10 mM sodium phosphate (pH 7.0), 1 mM EDTA, 5% glycerol, and a 0−1 M NaCl linear gradient. Gel filtration (Superdex 200, GE Healthcare, Waukesha, WI) in 25 mM MES (pH 6.0), 100 mM NaCl, and 1 mM EDTA was used for the final step of purification.

For NMR experiments, *Sp* Cdc5-D3 was purified as above except cells were grown and expressed in M9 media supplemented with the appropriate isotopic label (either ^{15}N NH_4Cl or ^{15}N NH_4Cl and ^{13}C D-glucose) (Cambridge Isotopes, Andover, MA). Ten percent D_2O was added to the final sample for all NMR experiments. For the ^{15}N Cdc5-D3 used in RNA titration experiments, S200-RNA buffer (25 mM MES (pH 6.0), 300 mM NaCl, 2 mM $MgCl_2$, and 1 mM EDTA) was used for gel filtration. For ^{15}N-leucine and ^{15}N-histidine (Cambridge Isotopes, Andover, MA) specific labeling, M9 media was supplemented with the appropriate unlabeled amino acids as well as ^{15}N L-leucine or ^{15}N L-histidine.[41]

Maltose binding protein MS2 binding protein (MBP-MS2BP) was expressed and purified using both a MBP and

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

heparin column as described.[42] After elution from the heparin column, fractions were pooled and concentrated using a 30K Amicon Ultra-15 filter (Millipore, Billerica, MA). Gel filtration (Superdex 200, GE Healthcare, Waukesha, WI) in 25 mM Tris (pH 7.4), 200 mM NaCl, 40 mM maltose, and 1 mM EDTA was used for the final step of purification. Samples were concentrated to approximately 1 mg/mL using a 30K Amicon Ultra-15 filter (Millipore, Billerica, MA) and stored at −20 °C. For the RNA pull-down experiment, samples were diluted in RNA buffer (20 mM HEPES (pH 7.4), 100 mM NaCl, 2 mM $MgCl_2$, and 5% glycerol) to a concentration of ∼0.15 mg/mL.

Analytical Ultracentrifugation. Purified *Sp* Cdc5-D3 was run in an Optima XLI ultracentrifuge (Beckman Coulter, Brea, CA) equipped with a four-hole An-60 Ti rotor at 42 000 rpm at 4 °C. Samples were loaded into double-sector cells (path length of 1.2 cm) with charcoal-filled Epon centerpieces and sapphire windows. Sedfit (version 12.0)[43] was used to analyze velocity scans using every seven scans from a total of 360 scans. Approximate size distributions were determined for a confidence level of $p = 0.95$, a resolution of $n = 300$, and sedimentation coefficients between 0 and 15 S.

NMR Spectroscopy. NMR experiments were performed at 25 °C in a 3 mm NMR tube (Wilmad Lab Glass, Vineland, NJ). Four-channel Bruker AVIII 600 and 800 NMR spectrometers (Bruker, Billerica, MA) equipped with CPCQCI and CPTCI probes, respectively, and single axis pulsed-field gradients were used. The assignment of backbone resonances for Cdc5-D3 was completed using standard 2D sensitivity enhanced echo/anti-echo $^{1}H-^{15}N$ heteronuclear single quantum coherence (HSQC)[44−47] and 3D HNCO,[48] HNCA,[48] CBCA(CO)NH,[49] HNCACB,[50] HN(CA)CO,[51] HN(CO)CA,[48] and ^{15}N-edited NOESY (τ = 120 ms)[52] experiments. The spectra were referenced to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) at 0 ppm (parts per million).

NMR Analysis, TALOS+, and Chemical Shift Rosetta. Data were processed in Topspin 3.2 (Bruker, Billerica, MA) and analyzed with Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco). Complete backbone ^{1}H and ^{15}N resonance assignments were obtained for all residues except for amino acid E188 and the GSH residues remaining from the N-terminal His tag. The chemical shifts of H, N, Cα, Cβ, and CO were analyzed with TALOS+,[53] a chemical shift index software that predicts secondary structure elements. An online version of Chemical Shift (CS)-Rosetta (https://condor.bmrb.wisc.edu/rosetta/)[54,55] was also used to generate 3000 models using the H, N, Cα, Cβ, and CO backbone chemical shift data. The flexible C-terminus of D3 (aa 201−214) was removed for CS-Rosetta modeling to allow for model convergence.

RNA and DNA Binding Assays. RNAs with a 5′ biotin were ordered from Dharmacon (GE Healthcare, Waukesha, WI) (Table SIII) and were 2′-bis(2-acetoxyethoxy)methyl (ACE) deprotected according to the manufacture's protocol. RNAs were further purified by ethanol precipitation and two 80% ethanol washes. RNA was heated at 94 °C for 10 min and then placed on ice before use. DNAs with a 5′ biotin were ordered from Operon (Eurofins Genomics, Huntsville, AL) (Table SIII). DNAs were annealed by heating at 94 °C for 10 min and then slowly cooled to room temperature. Streptavidin agarose (Life Technologies, Grand Island, NY) blocked with bovine serum albumin (BSA) (Sigma, St. Louis, MO) was incubated with 400 μL of 50 μM RNA or DNA for 45 min. The RNA or DNA-resin was washed in 10 mM HEPES (pH 7.5), 100 mM NaCl, and 10 mM $MgCl_2$ before the addition of ∼0.15 mg/mL recombinant N-terminal domains of Cdc5 for 45 min at room temperature. New RNA or DNA-resin was made for each replicate experiment. The resin was then washed in RNA buffer (20 mM HEPES (pH 7.4), 100 mM NaCl, 2 mM $MgCl_2$, and 5% glycerol). For protein−RNA or −DNA pull-down experiments, lithium dodecyl sulfate (LDS) sample buffer (Invitrogen, Carlsbad, CA) was added directly to the resin that was then boiled except in the case of Cdc5-R1-R2, due the presence of a contaminant on the resin at the same molecular weight. Cdc5-R1-R2 was eluted from the RNA-resin using RNA buffer with 1 M NaCl. Samples were then treated with LDS and boiled as above. Samples were run on 4−12% Bis-Tris PAGE gels (Invitrogen, Carlsbad, CA) and stained with Colloidal Coomassie.[56] All gels were quantified using ImageJ.[57] Average values were determined by comparing the intensity of each individual pull-down relative to the amount of protein loaded (relative to input %). Although adding excess RNA or DNA to the streptavidin agarose to saturate binding, we are unable to correct for the amount of RNA or DNA bound to the resin in these experiments. Graphs and statistics were generated using GraphPad Prism (version 5.0a).

NMR RNA Titration. RNAs (Table SIII) were ordered from Dharmacon (GE Healthcare) and were deprotected and washed per the manufacturer's instructions. RNAs were resuspended in S200-RNA buffer (25 mM MES (pH 6.0), 300 mM NaCl, 2 mM $MgCl_2$, and 1 mM EDTA). For RNA titrations, Cdc5-D3 was at ∼135 μM, and the RNA concentrations were titrated at 1:0.25, 1:0.5, 1:1, 1:2, 1:3, and 1:5 protein-to-RNA molar ratios for U6-ISL dsRNA and examined at a 1:1 protein-to-RNA molar ratio for U2 ssRNA, U6 ssRNA, and MS2 dsRNA. Chemical shift differences were calculated using the following formula

$$\Delta\delta_{\text{total}} = \sqrt{\left(\frac{\delta^{15}\text{N}_B - \delta^{15}\text{N}_{B_0}}{9.86204}\right)^2 + (\delta\text{H}_B - \delta\text{H}_{B_0})^2}$$

where B equals the chemical shift at variable RNA concentrations and B_0 is the chemical shift of protein only; 9.86204 is the absolute ratio of the gyromagnetic constants of ^{1}H and ^{15}N.[58] Changes in chemical shifts were fit to a single site-binding curve using GraphPad Prism (version 5.0a) and the following formula

$$\text{Y} = \frac{B_{\text{max}}\text{X}}{K_{\text{d}} + \text{X}}$$

where B_{max} is the maximum binding and K_{d} is the equilibrium binding constant.[59]

Circular Dichroism (CD). Purified Cdc5-D3 was analyzed using a Jasco J-810 spectropolarimeter (Jasco Analytical Instruments, Easton, MD). Far-UV (Ultraviolet) data were collected at a protein concentration of 0.15 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 50 nm/min from 198 to 260 nm. Far-UV spectra were collected in quadruplicate and background-corrected against a buffer blank. Data were converted to mean residue ellipticity $[\theta]_{\text{m}}$ (degrees cm^2 dmol^{-1}) using the following formula

$$[\theta]_{\text{m}} = \frac{\theta}{(10lcn)}$$

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

where θ is the measured ellipticity, l is the cell path length in centimeters, c is the molar concentration of protein in moles/liter, and n is the number of amino acids.

Circular dichroism data was deposited in the PCDDB (PCDDB ID: 0004551000). Backbone resonance assignments have been deposited in the BMRB (accession no. 25084).

RESULTS

R1, R2, and D3 Are Essential for *S. pombe* Cdc5 Function. Although *cdc5* and *CEF1* deletion studies in *S. pombe* and *S. cerevisiae* have been reported,[5,14] most of the deletion mutants used in these analyses did not directly correspond to secondary structural elements, making it difficult to determine the specific regions required for function. Using sequence alignments and secondary structure predictions, we identified five structural elements in the Cdc5 N-terminus, including the two Myb repeats (R1 and R2), a predicted Myb-like repeat (D3),[31] a nonstructured loop (L) between R2 and D3, and a predicted α-helical region (MID) downstream from D3 (Figure 1A). As a first step toward characterizing the role of the Cdc5 N-terminus in pre-mRNA splicing, we attempted to construct five *S. pombe* strains where the only copy of *cdc5* lacked one of the identified structural elements (Figure 1A). Although strains containing only $cdc5^{\Delta L}$ or $cdc5^{\Delta MID}$ grew normally at all temperatures (Figure 1B), we were unable to recover viable strains relying on only the expression of $cdc5^{\Delta R1}$, $cdc5^{\Delta R2}$, or $cdc5^{\Delta D3}$ (Figure 1C). Importantly, the R1, R2, and D3 deletions did not destabilize Cdc5, as seen from western blot analysis of wild-type cells overexpressing these mutants (Figure S1). From these results, we conclude that both of the canonical Myb repeats (R1 and R2) as well as Cdc5-D3 are essential for Cdc5 function, whereas the loop between R2 and D3 and the MID region are not. Although it was previously demonstrated that mutations in both canonical Myb repeats R1 and R2 affect pre-mRNA splicing,[17,31,32] this is the first evidence that Cdc5-D3 is also required for Cdc5 function.

Myb repeats are approximately 50 amino acids in length, with a tryptophan or tyrosine residue every 18−19 amino acids. The highly conserved tryptophan or tyrosine residues make up the hydrophobic core of a Myb repeat and are important for proper folding.[34,60,61] Structurally, this domain is composed of three well-defined helices, with the second and third helices adopting a fold similar to that of the canonical helix-turn-helix motif (HTH) found in many DNA binding proteins.[34,62,63] Cdc5-R1 and Cdc5-R2 adopt a canonical Myb-fold, as determined by NMR analysis (PDB: 2DIM and 2DIN). Although Cdc5-D3 has been predicted to be a Myb-like domain,[31] the lack of highly conserved tryptophan residues usually found in Myb domains (Figure 1D) suggests that Cdc5-D3 could adopt a different fold. To test whether the D3 region in Cdc5 could be replaced with a Myb fold, we replaced *cdc5-D3* with the sequence of Myb repeat 3 (R3) from *Mus musculus* (*Mm*) *c-MYB* ($cdc5^{D3::Myb}$) and asked whether this mutant could support cell function. The R3 Myb repeat from *Mm* c-Myb was chosen to replace Cdc5-D3 since it has been structurally characterized[34] and has a similar pI (isoelectric point) as Cdc5-D3 (10.4 and 10.7, respectively). As was seen with the $cdc5^{\Delta R1}$, $cdc5^{\Delta R2}$, and $cdc5^{\Delta D3}$ mutants, we were unable to recover a viable strain relying solely on the domain-swap mutant (Figure 1C). The inability of a canonical Myb repeat to substitute for Cdc5-D3 in cells suggested that the functional surfaces present on Cdc5-D3 are essential for function and cannot be replaced with a structural Myb domain.

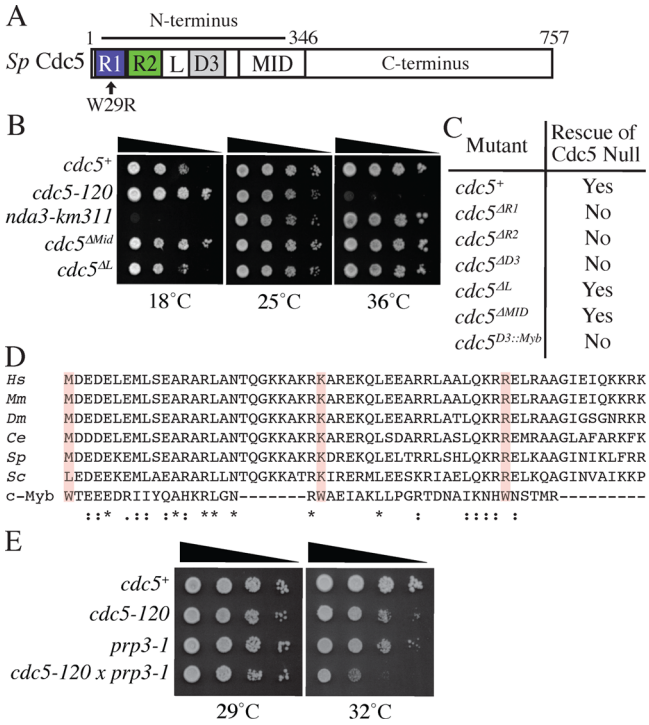


Figure 1. The canonical Myb repeats (R1, R2) and the proposed Myb-like repeat (D3) are required for function. (A) Domain architecture of *S. pombe* Cdc5. The Cdc5 N-terminus (aa 1−346) contains two canonical Myb repeats (R1, R2) and a third domain (D3) that is predicated to be a Myb-like repeat. R2 and D3 are separated by a loop region (L) predicted to be unstructured. A middle region of Cdc5 (MID) is predicted by PSI-PRED[98,99] to be α-helical. Position of domains are as follows: R1 (aa 5−55), R2 (aa 58−111), loop (aa 111−146), D3 (aa 155−214), and MID (aa 237−346). The Cdc5 C-terminus (aa 347−757) contains no predicted structural motifs. Black arrow marks the position of the mutation (W29R) found in temperature-sensitive *cdc5-120*.[31] (B) The loop and predicated α-helical regions of the Cdc5 N-terminus are not essential. Growth of serial dilutions of integrated $cdc5^{\Delta L}$ or $cdc5^{\Delta MID}$ cells grown at 18, 25, and 36 °C. *cdc5-120* is a heat-sensitive mutant,[17] whereas *nda3-km11* is a cold-sensitive mutant.[100] (C) Summary of N-terminal deletions of Cdc5 in *S. pombe*. *cdc5* deletions were determined to be essential by assaying for ability to rescue the *cdc5* null. (D) Sequence alignment of Cdc5-D3 from yeast to humans compared with *Mus musculus* c-Myb R3. *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Dm*, *Drosophila melanogaster*; *Ce*, *Caenorhabditis elegans*; *Sp*, *Schizosaccharomyces pombe*; *Sc*, *Saccharomyces cerevisiae*. The red boxes mark the location of the conserved tryptophan (W) residues in c-Myb and show that these tryptophans are not conserved in Cdc5-D3. Periods (.) represent weakly similar amino acids, colons (:) represent strongly similar amino acids, and asterisks (*) represent identical amino acids. (E) *cdc5-120* and *prp3-1* (the *S. pombe* homologue of *S. cerevisiae* cwc2[8]) interact genetically. Growth of serial dilutions of wild-type, *cdc5-120*, *prp3-1*, and *cdc5-120 prp3-1* cells grown at 29 and 32 °C.

Recently, another NTC component, *Sc* Cwc2 (*Sp* Cwf2/Prp3), was shown to interact directly with the U6 snRNA near the active site of the spliceosome.[15,16,64] Interestingly both *Sc* Cef1 (*Sp* Cdc5) and *Sc* Cwc2 (*Sp* Cwf2/Prp3), the two NTC components with nucleic acid binding domains, interact directly with Prp19,[8,65] putting them in close physical proximity to each other and suggesting they could both interact with RNA near the catalytic core of the spliceosome. To investigate if there is a functional connection between Cdc5 and Cwf2/Prp3, we tested for a genetic interaction between *cdc5-120* and

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

prp3-1. The *cdc5-120* mutation causes the amino acid substitution W29R in the R1 domain (Figure 1A),[31] whereas *prp3−1*[8] causes the amino acid substitution G123R in the ZnF domain, a domain that has been shown to cross-link to RNA in *S. cerevisiae*.[66] Our analysis shows that *cdc5−120 prp3−1* cells are synthetically sick (Figure 1E), suggesting that Cdc5 and Cwf2/Prp3 may function at similar stages of the splicing reaction.

Secondary Structure Analysis of Cdc5-D3 Shows That This Domain Is Composed of Two α-Helices and Does Not Adopt a Canonical Myb Fold. To examine the structural characteristics of Cdc5-D3, we expressed and purified recombinant Cdc5-D3 (aa 155−214) from *E. coli* (Figure 2A). Analysis by sedimentation analytical ultracentrifugation (SVAU) shows that Cdc5-D3 sediments as a monomer (S = 0.5; predicted molecular mass, ~7.2 kDa; root-mean-square deviation [RMSD] = 0.004) with a frictional ratio of 1.5 (Figure 2B). Circular dichroism (CD) analysis of Cdc5-D3 using far-UV light was done to predict secondary structure. When comparing this spectrum to known spectra,[67] Cdc5-D3 is primarily α-helical but contains random coil, as indicated by the lower signal at 222 nm (Figure S2). To further examine the secondary structure of this domain, we ^{15}N-labeled Cdc5-D3 and performed a two-dimensional (2D) ^{15}N−^{1}H HSQC experiment using nuclear magnetic resonance (NMR) spectroscopy (Figure 2C). In this spectrum, peaks were observed for ~94% of the expected amino acids in Cdc5-D3 (59 residues out of a total of 63; Figure S3), making Cdc5-D3 amenable for NMR secondary structure analysis.

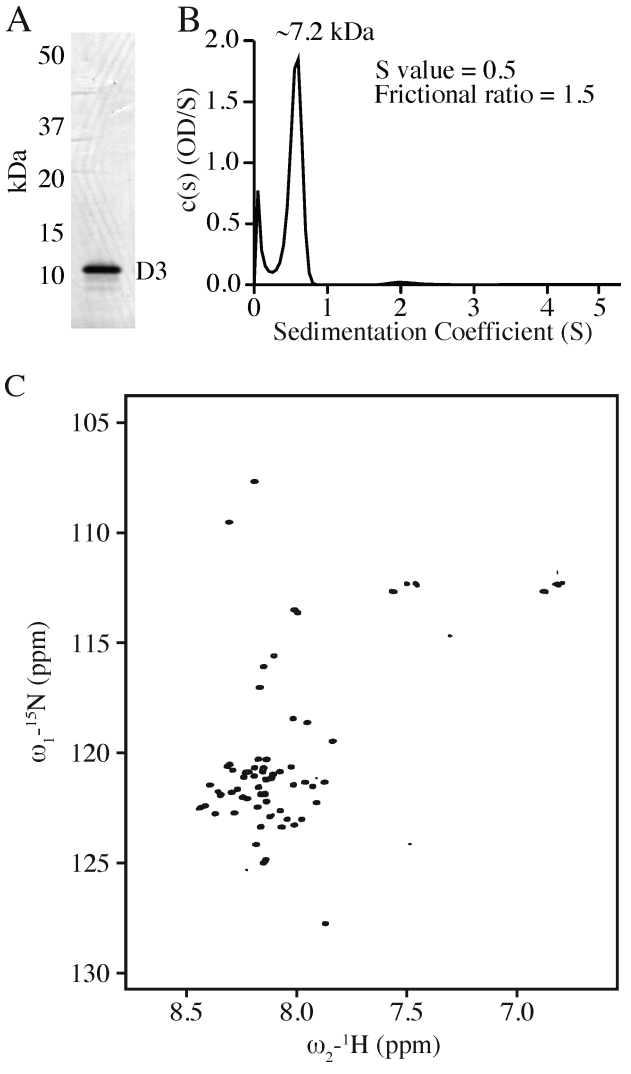


Figure 2. Biophysical characterization of Cdc5-D3. (A) Coomassie stained SDS-PAGE of Cdc5-D3 (aa 155−214). Cdc5-D3 runs at a higher molecular weight than what is predicted because of its pI. (B) SVAU analysis of Cdc5-D3. The S value, frictional ratio, and determined molecular mass are given for the main peak, which is ~93% of the sample. The RMSD is 0.004. (C) ^{15}N−^{1}H HSQC spectra of Cdc5-D3. Peaks for 59 of the 63 (94%) expected residues were identified.

To determine the secondary structure of Cdc5-D3, we used a combination of NMR experiments and computational modeling. Using ^{13}C- and ^{15}N-labeled Cdc5-D3, we ran a series of two-dimensional (2D) and three-dimensional (3D) NMR experiments to determine the backbone connectivity of Cdc5-D3. Results from a ^{15}N-NOESY-HSQC experiment showed there were two, as opposed to three, helical regions in Cdc5-D3 (Figure 3A). We then used TALOS+ to predict the secondary structure of Cdc5-D3 using the NH, Cα, Cβ, CO, and N chemical shift data. This analysis confirmed that Cdc5-D3 contains two α-helical regions (Figure 3A), rather than the three α-helices that would be expected for a Myb repeat (Figure S4A,B). To generate a 3D model of Cdc5-D3, we used the backbone chemical shift data and chemical-shift Rosetta (CS-Rosetta).[54,68] For calculations of 3D models, the chemical shift values for residues 201−214 were not included due to their predicted flexibility as calculated by TALOS+ and the lack of NOEs for these residues (Figure 3A). CS-Rosetta generated 3000 models of Cdc5-D3, and when these models were plotted with their all-atom energy versus their Cα RMSD in angstroms (Å), the models converged on the lowest energy model (Figure 3B,C). The 10 lowest energy models contain two α-helices (Figures 3D,E and S4C−L) and do not resemble a Myb fold (Figure S4B). Strikingly, all of the lowest 100 models contained two rather than three α-helices, leading us to conclude that Cdc5-D3 does not adopt a canonical Myb fold and should not be referred to as a Myb-like repeat.

The N-Terminus of *Sp* Cdc5 Binds RNA. As an essential pre-mRNA splicing factor and a core component of the NTC, Cdc5 family members in *S. pombe*, *S. cerevisiae*, and humans associate with the spliceosome starting with assembly of the B-complex and remain bound throughout the entire splicing reaction.[20,69−80] Because Cdc5 associates with the spliceosome, we hypothesize that the N-terminus may interact and bind to RNA. However, although the N-terminus of Cdc5 family members has been reported to bind DNA *in vitro*,[14,35,36] the ability of Cdc5 to interact with RNA has not been reported. Cdc5, as a core member of the NTC, is required for formation of the catalytic core,[11] suggesting that Cdc5 may interact with core regions of the spliceosome required for catalysis. This possibility is supported by the ability of mutations in the *Sc* Cef1 N-terminus to suppress first and second step splicing defects of a variety of mutants, including mutations in introns and the U6 snRNA.[32] To test the ability of the Cdc5 N-terminus to bind RNA, we used well-characterized regions of the U2 and U6 snRNAs that are predicted to be near the catalytic core of the spliceosome[81−84] (Table SIII). These RNAs include the double-stranded U6 snRNA interstem loop (U6-ISL), a single-stranded region of the U6 snRNA involved in 5′ splice-site selection (U6 ssRNA), and a single-stranded region of the U2 snRNA involved in branchpoint recognition

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

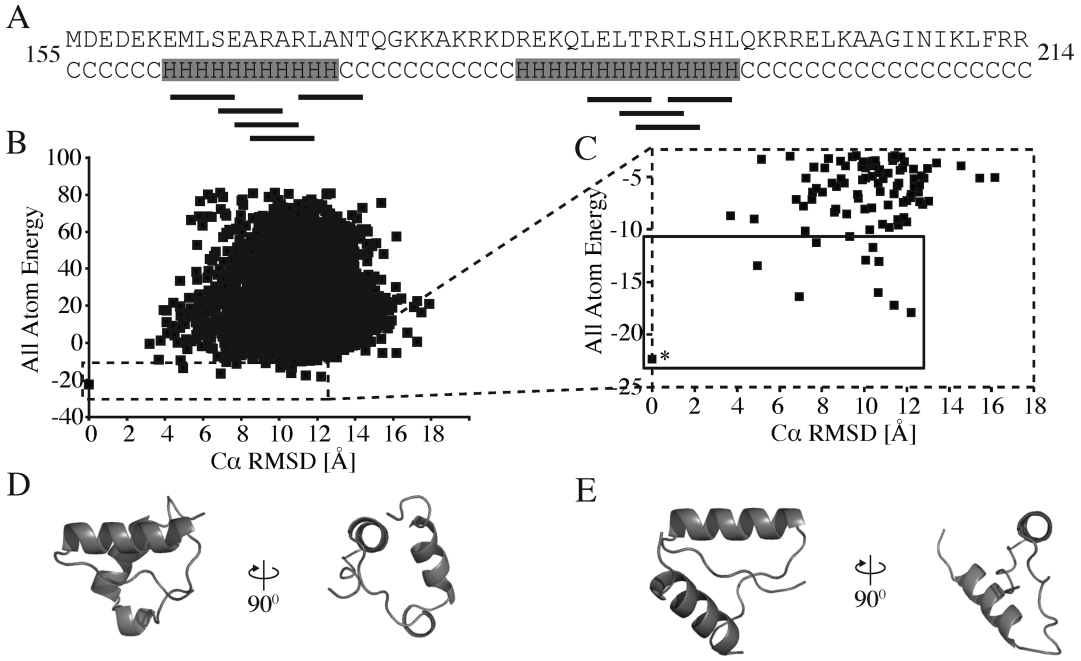


Figure 3. Secondary structure analysis and 3D modeling of Cdc5-D3 using backbone chemical shifts. (A) Secondary structure analysis of Cdc5-D3 using TALOS+.[53] Top row is the primary sequence for Cdc5-D3. The TALOS+ prediction is shown in the bottom row, where C represents random coil and H represents helical. Gray highlights the predicted helical regions. The black lines correspond to the helical regions found using ^{15}N-NOESY-HSQC. Each bar represents an i,i^{+4} or an i,i^{-4} NOE. (B) Summary of 3000 CS-Rosetta models of Cdc5-D3, comparing the all-atom energy versus the Cα RMSD relative to the lowest energy model. The black dashed box corresponds to the 100 lowest energy structures. (C) Magnification of the 100 lowest energy structures from panel B. Solid black line encompasses the 10 lowest energy structures. The asterisk (*) marks the lowest energy structure. (D, E) Two of the 10 lowest energy models calculated by CS-Rosetta shown at 0° and 90°. Panel D is the lowest energy model.

(U2 ssRNA)[83,85−87] (Figure 6A). The MS2 RNA hairpin (MS2 dsRNA) from bacteriophage was also prepared to analyze the sequence specificity of dsRNA binding.[42] To test for a direct interaction with the Cdc5 N-terminus, biotinylated RNAs bound to streptavidin agarose beads were incubated with recombinant Cdc5-R1-R2-D3$^{\Delta L}$. The nonessential loop region (Figure 1A−C) in this construct was deleted to generate a more stable protein. To detect binding, protein that remained bound to the resin after multiple washes was visualized by Coomassie staining (Figure 4A), and the percentage of protein that remained bound as compared to the original input was quantified using the results from multiple binding assays (Figure 4B). As seen in Figure 4A,B, Cdc5-R1-R2-D3$^{\Delta L}$ directly interacts with all of the tested RNAs.

The N-terminus of Cdc5 is composed of three essential domains, the two canonical Myb repeats (R1 and R2) and D3 (Figure 1A,C). To determine if the canonical Cdc5 Myb domains (R1 and R2) can individually interact with RNA, we repeated the RNA binding experiment using recombinant Cdc5-R1 and Cdc5-R2 (Figure 4C−F). In this assay, neither R1 nor R2 binds RNA (Figure 4C−F). However, since Myb repeats are often found as clustered groups,[33] we also investigated whether Cdc5-R1-R2 could bind RNA (Figure 4G−H). While Cdc5-R1-R2 does bind RNA (Figure 4G), we were surprised to find that when comparing the percentage of protein bound to the RNA (as compared to the initial input), only ∼20% of Cdc5-R1-R2 versus over 60% of Cdc5-R1-R2-D3$^{\Delta L}$ was bound to the U6-ISL (Figure 4B,H). Thus, although Cdc5-R1-R2 interacts with RNA, these two domains do not replicate the RNA binding seen with the entire Cdc5 N-terminus. Finally, since the N-terminus of Cdc5 family members have been reported to bind DNA,[14,35,36] we also tested the ability of Cdc5-R1-R2-R3$^{\Delta L}$ to interact with DNA, which it does (Figure S5A,B). Our RNA binding results suggests that the Cdc5-D3 region serves a more prominent role in facilitating RNA interactions than the canonical Myb repeats or, alternatively, all three domains are required for full RNA binding activity.

To discriminate between these possibilities, the binding experiment was repeated using Cdc5-D3 (Figure 4I−J). Cdc5-D3 alone binds directly to all of the RNAs in the assay (Figure 4I), with a significant binding preference for the U6-ISL (Figure 4J). To test whether the spliceosomal RNAs bind nonspecifically, we repeated the binding assays using an N-terminally MBP-tagged MS2 dsRNA binding protein (MBP-MS2BP) that interacts with the MS2 RNA hairpin loop.[42] While MPB-MS2BP bound to the MS2-dsRNA as expected, it did not bind to any of the spliceosomal RNAs (Figure 4K−L). In addition, like Cdc5-R1-R2-R3$^{\Delta L}$, Cdc5-D3 also binds DNA (Figure S5C,D). The results of these RNA binding assays demonstrate that the Cdc5 N-terminus binds to RNA directly and that, of the three essential domains, Cdc5-D3 binds RNA more robustly than the canonical Myb repeats alone.

Cdc5-D3 Preferentially Binds Double-Stranded RNA. To further characterize the interaction of Cdc5-D3 with RNA, we performed several NMR RNA titration experiments using the U2 ssRNA, U6 ssRNA, U6-ISL dsRNA, MS2 dsRNA, and ^{15}N Cdc5-D3 (Figures 5A and S6A−C). The preference for Cdc5-D3 binding to double-stranded RNAs detected in the *in vitro* binding assays (Figure 4I−J) was confirmed through chemical shift perturbation by examining the HSQC spectra of ^{15}N-labeled Cdc5-D3 after RNA was added at a 1:1 protein-to-RNA molar ratio (Figures 5B and S5A−C). The interaction between the U6-ISL dsRNA and Cdc5-D3 was further

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

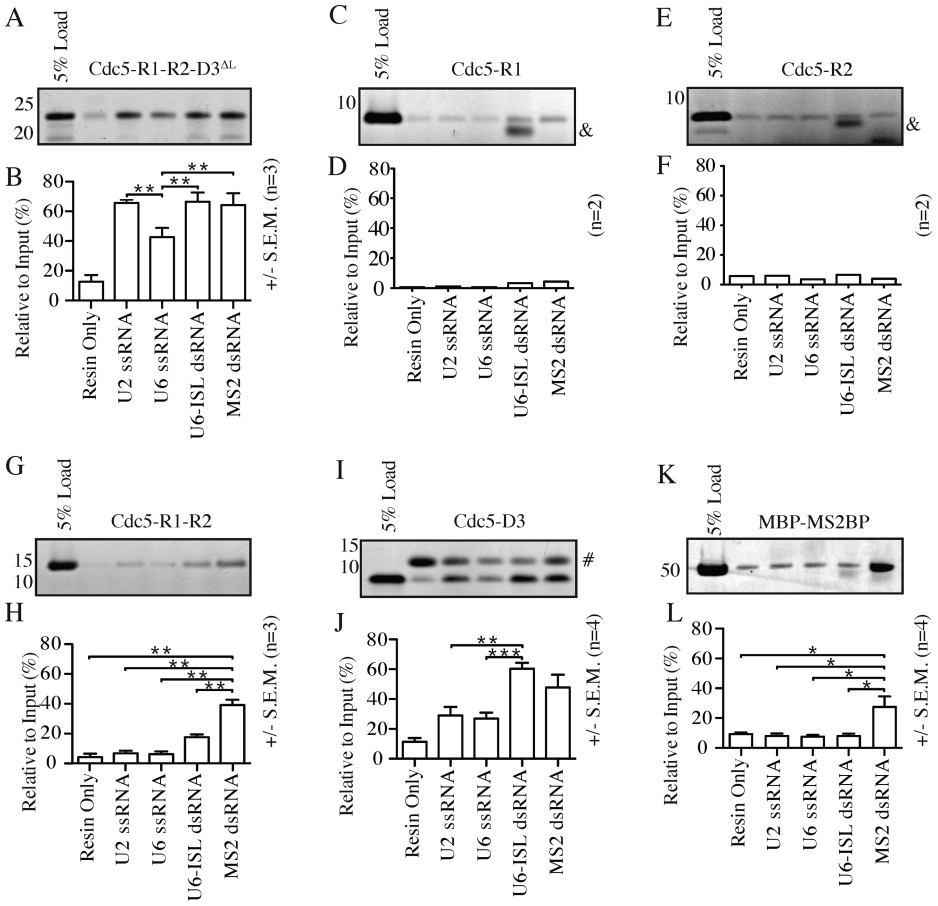


Figure 4. The N-terminus of Cdc5 binds RNA *in vitro*. Streptavidin-RNA pull-down assays using the indicated recombinant protein. Panels A, C, E, G, I, and K are Coomassie-stained SDS-PAGE gels of a representative pull-down experiment. Panels B, D, F, H, J, and L show the quantification of binding from multiple experiments. (A, B) Coomassie-stained representative SDS-PAGE gel of Cdc5 R1-R2-D3$^{\Delta L}$ bound to RNA agarose beads and quantification of three binding assays. (C, D) Coomassie-stained representative SDS-PAGE gel of Cdc5-R1 that binds to RNA agarose beads and average of two binding assays. Ampersand (&) marks the position of a contamination band released from the boiled U6 ILS dsRNA conjugated streptavidin resin. (E, F) Coomassie-stained representative SDS-PAGE gel of Cdc5-R2 bound to RNA agarose beads and average of two binding assays. Ampersand (&) marks the position of a contamination band released from the boiled U6 ILS dsRNA conjugated streptavidin resin. (G–H) Coomassie-stained representative SDS-PAGE gel of Cdc5-R1-R2 bound to RNA agarose beads and quantification of three binding assays. (I–J) Representative coomassie stained SDS-PAGE gel of Cdc5-D3 bound to RNA agarose beads and quantification of four binding assays. Pound (#) represents contamination from the streptavidin resin. (K–L) Representative coomassie stained SDS-PAGE gel of MBP tagged MS2 binding protein (MBP-MS2BP)[42] that binds to RNA agarose beads and quantification of four binding assays. SDS-PAGE gel quantification was done using ImageJ. Molecular weight markers are shown to the left of each gel. Error bars and statistics were generated using GraphPad Prism (version 5.0a), where the error bars correspond to the standard error of the mean (SEM), and the *p*-values are represented by asterisks ($p < 0.05$, *; $p < 0.01$, **; and $p < 0.001$, ***). 5′-Biotinylated RNA sequences used in the RNA binding experiment can be found in Table SIII.

characterized using a titration of RNA into ^{15}N-labeled protein ranging from substoichiometric molar ratios of protein-to-RNA (1:0.25) to saturable concentrations (1:5). This RNA titration resulted in substantial chemical shift changes in the NMR spectra for a subset of Cdc5-D3 peaks (Figure 5A), with the strongest shift differences localized to the last 16 amino acids (aa 199–214, Figure 5B,E–F), a region predicted to be unstructured in our secondary structure calculations. The last 16 amino acids also showed strong shifts with the titration of MS2 dsRNA (Figure S6C–E). Therefore, it is possible that this flexible region of Cdc5-D3 may become ordered when bound to RNA. The chemical shift changes appear to saturate at high RNA concentrations and can be fit to a 1:1 binding model (Figure 5C). Dissociation constants (K_d) from single site binding curves are ~210 μM (Figure 5C,D). The location of the Cdc5-D3 residues that interact with the U6-ISL in 2 of the 10 lowest energy models is shown in Figure 5E,F. The RNA titration experiments confirm that Cdc5-D3 preferentially binds double-stranded RNAs.

■ DISCUSSION

While the highly conserved N-terminus is essential for cell viability, the function of this portion of Cdc5 has not been characterized. The presence of Myb repeats in the N-terminus coupled with the ability of the C-terminus to directly interact with core NTC members[5,8,14] led us to hypothesize that Cdc5 may facilitate NTC-mediated RNA–RNA and/or RNA–protein transitions by interacting with both RNA and protein in the spliceosome. Using a combination of yeast genetics and *in vitro* RNA binding assays, we have shown that each of the R1, R2, and D3 domains in the Cdc5 N-terminus is essential and that Cdc5-R1-R2-D3$^{\Delta L}$ binds directly to regions within the U2 and U6 snRNAs. In addition, structural and biochemical analyses of Cdc5-D3 reveal that although this domain does not adopt a predicted Myb fold it is able to interact preferentially

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439–6451

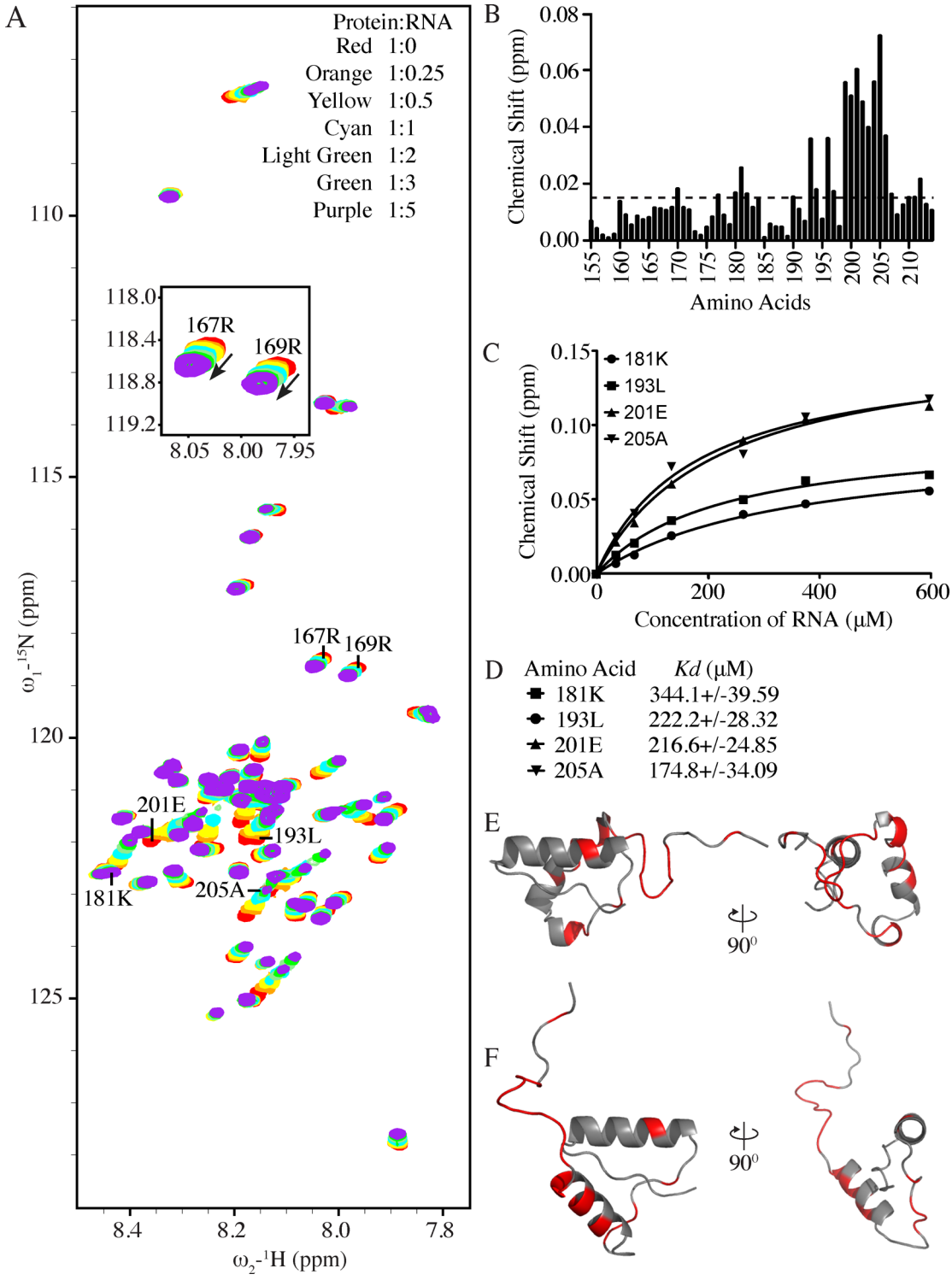


Figure 5. Cdc5-D3 RNA titration with U6-ISL dsRNA. (A) ^{15}N−^{1}H HSQC spectra of Cdc5-D3 with varying concentrations of U6-ISL dsRNA. The ^{15}N−^{1}H HSQC inset shows amino acids 167R and 169R. The black arrows indicate the direction the peaks shift upon the addition of U6-ISL dsRNA. (B) Chemical shifts (Δppm) from ^{15}N−^{1}H HSQC at a 1:1 molar ratio. The 20 strongest chemical shifts are above the dashed line. (C) Saturation curves of amino acids 181 K, 193L, 201E, and 205A are plotted and fit with a single-site binding curve. (D) Binding affinities (K_d) of 181 K, 193L, 201E, and 205A. (E, F) Chemical shifts from Cdc5-D3 and U6-ISL dsRNA NMR titration were plotted in red onto 2 of the 10 lowest energy models of Cdc5-D3. Models have an added flexible C-terminal extension that was not included in the CS-Rosetta calculations due to lack of NOEs, but do interact with RNA. Models shown at 0° and 90°.

with double-stranded RNAs, suggesting that it may be a Myb variant or a unique structural domain. Thus, the N-terminus of Cdc5 can directly interact with multiple RNAs while the C-terminus contacts other core NTC components,[3,8] suggesting that *in vivo* Cdc5 could tether NTC to RNA components within the spliceosome and provide a RNA binding platform that stabilizes the RNA−RNA and RNA−protein pre-mRNA splicing transitions.

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

Since Cdc5-D3 has a strong preference for U6-ISL and the MS2 dsRNAs despite the lack of sequence similarity between these RNAs, it seems likely that Cdc5-D3 might stabilize dsRNA conformations in a nonspecific manner. However, an alternative explanation is that Cdc5-D3 binds with high specificity to an RNA that we did not directly test. Combined with our *in vivo* results showing that the R1, R2, and D3 domains are all essential, we propose that Cdc5's N-terminus, with an overall pI of 9.4, acts as a charged RNA binding platform that binds to dsRNAs found near the catalytic core of the spliceosome.

Our finding that all three Cdc5 nucleic acid binding domains (R1, R2, and D3) are essential suggests a potential model where these domains could act synergistically *in vivo*. The interaction of multiple Myb repeats with DNA is common in transcription factors (reviewed in ref 33). For example, the transcription factor c-Myb contains three Myb repeats (R1, R2, and R3), and it is the third helix in the R2 and R3 Myb domains that directly contacts the major groove of a specific DNA sequence.[88,89] Therefore, using Myb-containing proteins as an example, we speculate that the Cdc5 R1, R2, and D3 domains may bind RNA in a similar manner as c-Myb binds DNA, with the exception that the lack of the third canonical Myb repeat may make the Cdc5 N-terminus better suited for binding nucleic acid structures (i.e., dsRNA) versus binding specific nucleic acid sequences. The ability of Cdc5 to interact with multiple RNA structures could provide a platform to stabilize the conformational changes that occur during the transition of the spliceosome from an inactive to activated complex.

The spliceosome is a dynamic macromolecular machine composed of both protein and RNA components. During the transition from the inactive spliceosomal B-complex to the activated B^{act}-complex, the U1 and U4 snRNAs dissociate, allowing the U6 snRNA to change conformation by forming the catalytic U6-ISL, and form duplexes with both the U2 snRNA and the 5′ end of the intron via the ACAGAGA box.[83] Although these RNA−RNA remodeling events are the major requirements for the formation of the spliceosome active site, additional events, including the stable association of the NTC, are required for catalysis to occur. The active site of the spliceosome is highly conserved from yeast to humans[85,90−92] and is centered on a region of the U6 snRNA, which shares three similarities with group II self-splicing introns, the AGC triad, Mg^{2+} binding, and the U6-ISL.[86,93,94] Similarities between the U6 snRNA and the group II self-splicing introns,[95] as well as the splicing-like activity demonstrated by a protein-free RNA construct containing regions of the U6 and U2 snRNAs that base pair in the spliceosome,[96,97] have led to the conclusion that the regions of U6 and U2 snRNAs make up essential components of the spliceosome active site. However, unlike the group II self-splicing introns that use additional RNA domains to support catalysis, the spliceosome has evolved to also use proteins to facilitate the splicing reaction. Two proteins, the highly conserved Prp8 (*Sp* Spp42) and *Hs* RBM22/*Sc* Cwc2 (*Sp* Cwf2/Prp3), have been shown to directly cross-link to both the U6 snRNA and the pre-mRNA within the active site of the spliceosome (Figure 6).[15,16,64] The physical interaction between *Sc* Cwc2 and U6 snRNA places the NTC at the active site of the spliceosome, and, interestingly, both *Sc* Cef1 (*Sp* Cdc5) and *Sc* Cwc2 bind directly to Prp19[8,65] (Figure 6B). Thus, the only two NTC components that contain nucleic acid binding domains are likely in close physical proximity. However, although Cwc2 family members have been shown

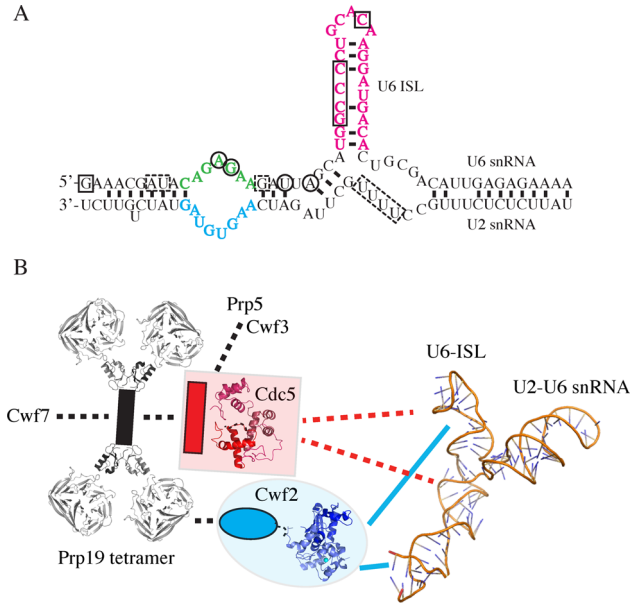


Figure 6. Model of the U2 and U6 snRNAs and NTC components in the spliceosome B^{act} complex. (A) A secondary structure model of U2 and U6 snRNA interactions in the *S. pombe* spliceosome active site, as predicted from studies in *S. cerevisiae*. (Adapted from ref 85. Copyright 2012 RNA Society). The snRNAs are highly conserved. The U6-ISL is pink, the U6 ssRNA is green, and the U2 ssRNA is blue. Dashed boxes are residues cross-linked to *Sc* Prp8,[64] and solid boxes are residues cross-linked to *Sc* Cwc2.[16] Circles mark the location of mutations in *Sc* U6 snRNA suppressed by mutations in the R1 domain of *Sc* Cef1.[32] (B) Model of a subset of NTC components and the U2−U6 snRNA. The Prp19 tetramer binds directly to NTC components Cdc5 (*Sc* Cef1), Cwf7 (*Sc* Snt309), and Cwf2 (*Sc* Cwc2) (solid lines).[8] *Sc* Cwc2 cross-links to regions of the U2 and U6 snRNAs. The C-terminus of Cdc5 binds directly to Prp5 (*Sc* Prp46) and Cwf3 (*Sc* Syf1),[8] while the N-terminus binds RNA *in vitro*. Dashed lines indicate a direct *in vitro* interaction between components, and solid lines represent *in vitro* protein−RNA cross-linking. Models are not to scale. Prp19 tetramer: light gray structures, WD40 repeats (PDB: 3LVR[65]); black and medium gray structures, U-box dimers (PDB: 2BAY[101]); and black rectangle, coiled-coil region. Cwc2: blue structures, Torus/Zn-finger and RRM domain (PDB: 3U1M[102]); blue oval, the C-terminus. Cdc5: red structures, a model of the N-terminus, generated using Modeler 9v8 and hCDC5 R1 (PDB: 2DIM), hCDC5 R2 (PDB: 2DIN), and the lowest energy CS-Rosetta structure for Cdc5-D3; red box, the C-terminus. *Sc* U2−U6 snRNA (PDB: 2LKR[85]).

to directly cross-link to the U6 snRNA *in vitro*,[15,16,66] Cdc5 family members have not. While there is no direct evidence that Cdc5 interacts with RNA *in vivo*, genetic mutations in Cef1 (*Sp* Cdc5) are able to suppress mutations in the U6 snRNA found in the active site of the spliceosome (Figure 6A),[32] and we have shown that there is a negative genetic interaction between the *cdc5* and *cwf2*/*prp3* temperature-sensitive alleles that generate point mutations in the RNA binding domains of each protein (Figure 1E). The proximity of Cwc2 to Cef1 in the NTC,[8] the capability of Cef1/Cdc5 (*Sc*/*Sp*) to suppress U6 snRNA mutations,[32] the ability of Cdc5 to bind RNA *in vitro*, and the negative genetic interaction between *cdc5-120* and *prp3-1* (*Sp* *cwf2*) make it possible that Cdc5, as a core member of the NTC, is positioned near the active site of the spliceosome and can act as a charged platform that helps to stabilize the RNA−RNA rearrangements that occur during the transition from the B- to B^{act}-spliceosome (Figure 6B).

dx.doi.org/10.1021/bi5008639 | *Biochemistry* 2014, 53, 6439−6451

Our findings suggest a model where the ability of the Cdc5 N-terminus to directly interact with a variety of RNA structures allows it to act as a binding scaffold that supports and/or facilitates the RNA−RNA and RNA−protein remodeling that occurs during the transition from the B- to B^{act}-spliceosomal complexes. Thus, we propose that Cdc5 contributes to NTC function by playing an essential role in facilitating the conformational changes that occur during spliceosome activation. Our studies indicate that Cdc5 is biochemically positioned to play a pivotal role in pre-mRNA processing through its interaction with NTC components and RNA within the active site of the spliceosome.

■ ASSOCIATED CONTENT

Supporting Information

Cdc5 deletions are stable (Figure S1), CD spectrum of Cdc5-D3 (Figure S2), labeled 2D HSQC of Cdc5-D3 (Figure S3), secondary structure analysis of Cdc5 N-terminal domains (Figure S4), the N-terminus of Cdc5 binds to DNA *in vitro* (Figure S5), and chemical shift analysis of Cdc5-D3 with ssRNA and MS2-dsRNA (Figure S6). Strains and plasmids used in this study are listed in Tables SI and SII, and sequences of RNA and DNA used in these studies in Table SIII. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 615-936-7780. E-mail: Melanie.ohi@vanderbilt.edu.

Funding

This work was supported by T32 GM08320 to S.E.C. and NIH DP2OD004483 to M.D.O.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

DSS, 4,4-dimethyl-4-silapentane-1-sulfonic acid; aa, amino acid; BSA, bovine serum albumin; CD, circular dichroism; cDNA, complementary DNA; DNA, deoxyribonucleic acid; dsRNA, double-stranded RNA; HAT, half a TPR domain; HSQC, heteronuclear single-quantum correlation; *Hs*, *Homo sapiens*; ISL, interstem loop; pI, isoelectric point; LDS, lithium dodecyl sulfate; MBP, maltose binding protein; MBP-MS2BP, maltose binding protein-MS2 binding protein; mRNA, messenger RNA; *Ms*, *Mus musculus*; myeloblastosis, Myb; NTC, NineTeen Complex; *nmt*, no message in thiamine; NMR, nuclear magnetic resonance; ORF, open reading frame; P-complex, postcatalytic complex; ppm, parts per million; pre-mRNA, precursor mRNA; RNA, ribonucleic acid; RRM, RNA recognition motif; *Sc*, *Saccharomyces cerevisiae*; *Sp*, *Schizosaccharomyces pombe*; SVAU, sedimentation velocity analytical ultracentrifugation; ssRNA, single-stranded RNA; snRNPs, small nuclear ribonucleoproteins; snRNAs, small nuclear RNAs; SEM, standard error of the mean; TPR, tetratricopeptide repeats; ACE, two-acetoxyethoxy; 2D, two-dimensional; UV, ultraviolet; ZnF, zinc finger

■ REFERENCES

(1) Tarn, W. Y., Hsu, C. H., Huang, K. T., Chen, H. R., Kao, H. Y., Lee, K. R., and Cheng, S. C. (1994) Functional association of essential splicing factor(s) with PRP19 in a protein complex. *EMBO J. 13*, 2421−2431.

(2) Ajuh, P., Kuster, B., Panov, K., Zomerdijk, J. C., Mann, M., and Lamond, A. I. (2000) Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *EMBO J. 19*, 6569−6581.

(3) Grote, M., Wolf, E., Will, C. L., Lemm, I., Agafonov, D. E., Schomburg, A., Fischle, W., Urlaub, H., and Luhrmann, R. (2010) Molecular architecture of the human Prp19/CDC5L complex. *Mol. Cell. Biol. 30*, 2105−2119.

(4) Chen, H. R., Jan, S. P., Tsao, T. Y., Sheu, Y. J., Banroques, J., and Cheng, S. C. (1998) Snt309p, a component of the Prp19p-associated complex that interacts with Prp19p and associates with the spliceosome simultaneously with or immediately after dissociation of U4 in the same manner as Prp19p. *Mol. Cell. Biol. 18*, 2196−2204.

(5) Tsai, W. Y., Chow, Y. T., Chen, H. R., Huang, K. T., Hong, R. I., Jan, S. P., Kuo, N. Y., Tsao, T. Y., Chen, C. H., and Cheng, S. C. (1999) Cef1p is a component of the Prp19p-associated complex and essential for pre-mRNA splicing. *J. Biol. Chem. 274*, 9455−9462.

(6) Chen, C. H., Tsai, W. Y., Chen, H. R., Wang, C. H., and Cheng, S. C. (2001) Identification and characterization of two novel components of the Prp19p-associated complex, Ntc30p and Ntc20p. *J. Biol. Chem. 276*, 488−494.

(7) Chen, C.-H., Yu, W.-C., Tsao, T. Y., Wang, L.-Y., Chen, H.-R., Lin, J.-Y., Tsai, W.-Y., and Cheng, S.-C. (2002) Functional and physical interactions between components of the Prp19p-associated complex. *Nucleic Acids Res. 30*, 1029−1037.

(8) Ohi, M. D., and Gould, K. L. (2002) Characterization of interactions among the Cef1p-Prp19p-associated splicing complex. *RNA 8*, 798−815.

(9) Ben-Yehuda, S., Dix, I., Russell, C. S., McGarvey, M., Beggs, J. D., and Kupiec, M. (2000) Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in Saccharomyces cerevisiae. *Genetics 156*, 1503−1517.

(10) Chan, S. P., Kao, D. I., Tsai, W. Y., and Cheng, S. C. (2003) The Prp19p-associated complex in spliceosome activation. *Science 302*, 279−282.

(11) Chan, S.-P., and Cheng, S.-C. (2005) The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J. Biol. Chem. 280*, 31190−31199.

(12) Villa, T., and Guthrie, C. (2005) The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. *Genes Dev. 19*, 1894−1904.

(13) Pleiss, J. A., Whitworth, G. B., Bergkessel, M., and Guthrie, C. (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol. 5*, e90.

(14) Ohi, R., McCollum, D., Hirani, B., Den Haese, G. J., Zhang, X., Burke, J. D., Turner, K., and Gould, K. L. (1994) The *Schizosaccharomyces pombe* cdc5+ gene encodes an essential protein with homology to c-Myb. *EMBO J. 13*, 471−483.

(15) McGrail, J. C., Krause, A., and O'Keefe, R. T. (2009) The RNA binding protein Cwc2 interacts directly with the U6 snRNA to link the nineteen complex to the spliceosome during pre-mRNA splicing. *Nucleic Acids Res. 37*, 4205−4217.

(16) Rasche, N., Dybkov, O., Schmitzova, J., Akyildiz, B., Fabrizio, P., and Luhrmann, R. (2012) Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre. *EMBO J. 31*, 1591−1604.

(17) Nurse, P., Thuriaux, P., and Nasmyth, K. (1976) Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. *Mol. Gen. Genet. 146*, 167−178.

(18) McDonald, W. H., Ohi, R., Smelkova, N., Frendewey, D., and Gould, K. L. (1999) Myb-related fission yeast cdc5p is a component of a 40S snRNP-containing complex and is essential for pre-mRNA splicing. *Mol. Cell. Biol. 19*, 5352−5362.

(19) Burns, C. G., Ohi, R., Krainer, A. R., and Gould, K. L. (1999) Evidence that Myb-related CDC5 proteins are required for pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A. 96*, 13789−13794.

(20) Ohi, M. D., Link, A. J., Ren, L., Jennings, J. L., McDonald, W. H., and Gould, K. L. (2002) Proteomics analysis reveals stable

multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. *Mol. Cell. Biol. 22*, 2011−2024.
(21) Liu, L., Gräub, R., Hlaing, M., Epting, C. L., Turck, C. W., Xu, X.-Q., Zhang, L., and Bernstein, H. S. (2003) Distinct domains of human CDC5 direct its nuclear import and association with the spliceosome. *Cell Biochem. Biophys. 39*, 119−132.
(22) Burns, C. G., Ohi, R., Mehta, S., O'Toole, E. T., Winey, M., Clark, T. A., Sugnet, C. W., Ares, M., and Gould, K. L. (2002) Removal of a single alpha-tubulin gene intron suppresses cell cycle arrest phenotypes of splicing factor mutations in Saccharomyces cerevisiae. *Mol. Cell. Biol. 22*, 801−815.
(23) Zhou, Z., Licklider, L. J., Gygi, S. P., and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature 419*, 182−185.
(24) Lin, Z., Yin, K., Zhu, D., Chen, Z., Gu, H., and Qu, L. J. (2007) AtCDC5 regulates the G2 to M transition of the cell cycle and is critical for the function of *Arabidopsis* shoot apical meristem. *Cell Res. 17*, 815−828.
(25) Zhang, N., Kaur, R., Lu, X., Shen, X., Li, L., and Legerski, R. J. (2005) The Pso4 mRNA splicing and DNA repair complex interacts with WRN for processing of DNA interstrand cross-links. *J. Biol. Chem. 280*, 40559−40567.
(26) Zhang, N., Kaur, R., Akhter, S., and Legerski, R. J. (2009) Cdc5L interacts with ATR and is required for the S-phase cell-cycle checkpoint. *EMBO Rep. 10*, 1029−1035.
(27) Maréchal, A., Li, J. M., Ji, X. Y., Wu, C. S., Yazinski, S. A., Nguyen, H. D., Liu, S., Jiménez, A. E., Jin, J., and Zou, L. (2014) PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. *Mol. Cell 53*, 235−246.
(28) Wan, L., and Huang, J. (2014) The PSO4 protein complex associates with replication protein A (RPA) and modulates the activation of ataxia telangiectasia-mutated and Rad3-related (ATR). *J. Biol. Chem. 289*, 6619−6626.
(29) Hofmann, J. C., Tegha-Dunghu, J., Dräger, S., Will, C. L., Lührmann, R., and Gruss, O. J. (2013) The Prp19 complex directly functions in mitotic spindle assembly. *PLoS One 8*, e74851.
(30) Zhang, S., Xie, M., Ren, G., and Yu, B. (2013) CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. *Proc. Natl. Acad. Sci. U.S.A. 110*, 17588−17593.
(31) Ohi, R., Feoktistova, A., McCann, S., Valentine, V., Look, A. T., Lipsick, J. S., and Gould, K. L. (1998) Myb-related Schizosaccharomyces pombe cdc5p is structurally and functionally conserved in eukaryotes. *Mol. Cell. Biol. 18*, 4097−4108.
(32) Query, C. C., and Konarska, M. M. (2012) CEF1/CDC5 alleles modulate transitions between catalytic conformations of the spliceosome. *RNA 18*, 1001−1013.
(33) Prouse, M. B., and Campbell, M. M. (2012) The interaction between MYB proteins and their target DNA binding sites. *Biochim. Biophys. Acta 1819*, 67−77.
(34) Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S., and Nishimura, Y. (1992) Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. U.S.A. 89*, 6428−6432.
(35) Hirayama, T., and Shinozaki, K. (1996) A cdc5+ homolog of a higher plant, *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A. 93*, 13371−13376.
(36) Lei, X. H., Shen, X., Xu, X. Q., and Bernstein, H. S. (2000) Human Cdc5, a regulator of mitotic entry, can act as a site-specific DNA binding protein. *J. Cell Sci. 113*, 4523−4531.
(37) Hogg, R., de Almeida, R. A., Ruckshanthi, J. P., and O'Keefe, R. T. (2014) Remodeling of U2-U6 snRNA helix I during pre-mRNA splicing by Prp16 and the NineTeen Complex protein Cwc2. *Nucleic Acids Res. 42*, 8008−8023.
(38) Keeney, J. B., and Boeke, J. D. (1994) Efficient targeted integration at leu1-32 and ura4-294 in *Schizosaccharomyces pombe*. *Genetics 136*, 849−856.
(39) Livesay, S. B., Collier, S. E., Bitton, D. A., Bähler, J., and Ohi, M. D. (2013) Structural and functional characterization of the N terminus of Schizosaccharomyces pombe Cwf10. *Eukaryotic Cell 12*, 1472−1489.
(40) Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol. 112*, 531−552.
(41) Cheng, H., Westler, W. M., Xia, B., Oh, B. H., and Markley, J. L. (1995) Protein expression, selective isotopic labeling, and analysis of hyperfine-shifted NMR signals of Anabaena 7120 vegetative [2Fe−2S]ferredoxin. *Arch. Biochem. Biophys. 316*, 619−634.
(42) Zhou, Z., Sim, J., Griffith, J., and Reed, R. (2002) Purification and electron microscopic visualization of functional human spliceosomes. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12203−12207.
(43) Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J. 78*, 1606−1619.
(44) Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sørensen, O. W., and Griesinger, C. (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol NMR 4*, 301−306.
(45) Grzesiek, S., and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc. 115*, 12593−12594.
(46) Palmer, A., Cavanagh, J., Byrd, R., and Rance, M. (1992) Sensitivity improvement in 3-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson. 96*, 416−424.
(47) Lewis, K. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc. 114*, 10663−10665.
(48) Grzesiek, S., and Bax, A. (1992) An efficient experiment for sequential backboneassignment of medium-sized isotopically enriched protein. *J. Magn. Reson. 99*, 201−207.
(49) Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J. Biomol. NMR 3*, 185−204.
(50) Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B 101*, 201−205.
(51) Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N, and $^{13}C'$ chemical shifts in $^{15}N-^{13}C$-labelled proteins. *J. Magn. Reson. 97*, 213−217.
(52) Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson. 98*, 207−216.
(53) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR 44*, 213−223.
(54) Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) De novo protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR 43*, 63−78.
(55) Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A. 105*, 4685−4690.
(56) Dyballa, N., and Metzger, S. (2009) Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *J. Visualized Exp. 30*, 1431.
(57) Rasband, W. S. (1997−2004) *ImageJ*, U. S. National Institutes of Health, Bethesda, MD, http://imagej.nih.gov/ij/.

(58) Martin-Tumasz, S., Reiter, N. J., Brow, D. A., and Butcher, S. E. (2010) Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. *RNA 16*, 792−804.
(59) Barrett, P. J., Song, Y., Van Horn, W. D., Hustedt, E. J., Schafer, J. M., Hadziselimovic, A., Beel, A. J., and Sanders, C. R. (2012) The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science 336*, 1168−1171.
(60) Saikumar, P., Murali, R., and Reddy, E. P. (1990) Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. *Proc. Natl. Acad. Sci. U.S.A. 87*, 8452−8456.
(61) Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D. N., Ogata, K., Nishimura, Y., and Ishii, S. (1990) The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the myb protooncogene product. *J. Biol. Chem. 265*, 19990−19995.
(62) Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M., and Pabo, C. O. (1982) Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature 298*, 447−451.
(63) Ohlendorf, D. H., Anderson, W. F., and Matthews, B. W. (1983) Many gene-regulatory proteins appear to have a similar alpha-helical fold that binds DNA and evolved from a common precursor. *J. Mol. Evol. 19*, 109−114.
(64) Li, X., Zhang, W., Xu, T., Ramsey, J., Zhang, L., Hill, R., Hansen, K. C., Hesselberth, J. R., and Zhao, R. (2013) Comprehensive in vivo RNA-binding site analyses reveal a role of Prp8 in spliceosomal assembly. *Nucleic Acids Res. 41*, 3805−3818.
(65) Vander Kooi, C. W., Ren, L., Xu, P., Ohi, M. D., Gould, K. L., and Chazin, W. J. (2010) The Prp19 WD40 domain contains a conserved protein interaction region essential for its function. *Structure 18*, 584−593.
(66) Schmitzová, J., Rasche, N., Dybkov, O., Kramer, K., Fabrizio, P., Urlaub, H., Lührmann, R., and Pena, V. (2012) Crystal structure of Cwc2 reveals a novel architecture of a multipartite RNA-binding protein. *EMBO J. 31*, 2222−2234.
(67) Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry 8*, 4108−4116.
(68) Shen, Y., and Bax, A. (2012) Identification of helix capping and b-turn motifs from NMR chemical shifts. *J. Biomol. NMR 52*, 211−232.
(69) Jurica, M. S., Licklider, L. J., Gygi, S. R., Grigorieff, N., and Moore, M. J. (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA 8*, 426−439.
(70) Stevens, S. W., Ryan, D. E., Ge, H. Y., Moore, R. E., Young, M. K., Lee, T. D., and Abelson, J. (2002) Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell 9*, 31−44.
(71) Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M., and Lührmann, R. (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science 298*, 2205−2208.
(72) Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C. L., Kastner, B., Stark, H., Urlaub, H., and Luhrmann, R. (2006) Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol. Cell. Biol. 26*, 5528−5543.
(73) Bessonov, S., Anokhina, M., Will, C. L., Urlaub, H., and Lührmann, R. (2008) Isolation of an active step I spliceosome and composition of its RNP core. *Nature 452*, 846−850.
(74) Warkocki, Z., Odenwalder, P., Schmitzova, J., Platzmann, F., Stark, H., Urlaub, H., Ficner, R., Fabrizio, P., and Luhrmann, R. (2009) Reconstitution of both steps of *Saccharomyces cerevisiae* splicing with purified spliceosomal components. *Nat. Struct. Mol. Biol. 16*, 1237−1243.
(75) Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009) The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol. Cell 36*, 593−608.
(76) Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M. M., Sander, B., Will, C. L., Urlaub, H., Stark, H., and Lührmann, R. (2010) Characterization of purified human B^{act} spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA 16*, 2384−2403.
(77) Ren, L., McLean, J. R., Hazbun, T. R., Fields, S., Vander Kooi, C., Ohi, M. D., and Gould, K. L. (2011) Systematic two-hybrid and comparative proteomic analyses reveal novel yeast pre-mRNA splicing factors connected to Prp19. *PLoS One 6*, e16719.
(78) Cvitkovic, I., and Jurica, M. S. (2013) Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic Acids Res. 41*, D132−141.
(79) Lardelli, R. M., Thompson, J. X., Yates, J. R., and Stevens, S. W. (2010) Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA 16*, 516−528.
(80) Fourmann, J. B., Schmitzová, J., Christian, H., Urlaub, H., Ficner, R., Boon, K. L., Fabrizio, P., and Lührmann, R. (2013) Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev. 27*, 413−428.
(81) Lesser, C. F., and Guthrie, C. (1993) Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science 262*, 1982−1988.
(82) Yean, S. L., Wuenschell, G., Termini, J., and Lin, R. J. (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature 408*, 881−884.
(83) Madhani, H. D., and Guthrie, C. (1992) A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell 71*, 803−817.
(84) Sontheimer, E. J., and Steitz, J. A. (1993) The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science 262*, 1989−1996.
(85) Burke, J. E., Sashital, D. G., Zuo, X., Wang, Y.-X., and Butcher, S. E. (2012) Structure of the yeast U2/U6 snRNA complex. *RNA 18*, 673−683.
(86) Montemayor, E. J., Curran, E. C., Liao, H. H., Andrews, K. L., Treba, C. N., Butcher, S. E., and Brow, D. A. (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nat. Struct. Mol. Biol. 21*, 544−551.
(87) Sashital, D. G., Cornilescu, G., McManus, C. J., Brow, D. A., and Butcher, S. E. (2004) U2−U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat. Struct. Mol. Biol. 11*, 1237−1242.
(88) Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell 79*, 639−648.
(89) Ogata, K., Morikawa, S., Nakamura, H., Hojo, H., Yoshimura, S., Zhang, R., Aimoto, S., Ametani, Y., Hirata, Z., and Sarai, A. (1995) Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nat. Struct. Biol. 2*, 309−320.
(90) Hausner, T. P., Giglio, L. M., and Weiner, A. M. (1990) Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev. 4*, 2146−2156.
(91) Datta, B., and Weiner, A. M. (1991) Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature 352*, 821−824.
(92) Luukkonen, B. G., and Séraphin, B. (1998) Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. *RNA 4*, 167−180.
(93) Seetharaman, M., Eldho, N. V., Padgett, R. A., and Dayie, K. T. (2006) Structure of a self-splicing group II intron catalytic effector domain 5: parallels with spliceosomal U6 RNA. *RNA 12*, 235−247.
(94) Reiter, N. J., Blad, H., Abildgaard, F., and Butcher, S. E. (2004) Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry 43*, 13739−13747.
(95) Fica, S. M., Mefford, M. A., Piccirilli, J. A., and Staley, J. P. (2014) Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nat. Struct Mol. Biol. 21*, 464−471.

(96) Valadkhan, S., Mohammadi, A., Jaladat, Y., and Geisler, S. (2009) Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc. Natl. Acad. Sci. U.S.A. 106*, 11901−11906.
(97) Valadkhan, S., and Manley, J. L. (2001) Splicing-related catalysis by protein-free snRNAs. *Nature 413*, 701−707.
(98) Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res. 41*, W349−357.
(99) Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol. 292*, 195−202.
(100) Hiraoka, Y., Toda, T., and Yanagida, M. (1984) The NDA3 gene of fission yeast encodes beta-tubulin: a cold-sensitive nda3 mutation reversibly blocks spindle formation and chromosome movement in mitosis. *Cell 39*, 349−358.
(101) Vander Kooi, C. W., Ohi, M. D., Rosenberg, J. A., Oldham, M. L., Newcomer, M. E., Gould, K. L., and Chazin, W. J. (2006) The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases. *Biochemistry 45*, 121−130.
(102) Lu, P., Lu, G., Yan, C., Wang, L., Li, W., and Yin, P. (2012) Structure of the mRNA splicing complex component Cwc2: insights into RNA recognition. *Biochem. J. 441*, 591−597.