Physical and genetic interactions of yeast Cwc21p, an ortholog of human SRm300/SRRM2, suggest a role at the catalytic center of the spliceosome

Citation for published version:
Grainger, RJ, Barrass, JD, Jacquier, A, Rain, J-C & Beggs, JD 2009, 'Physical and genetic interactions of yeast Cwc21p, an ortholog of human SRm300/SRRM2, suggest a role at the catalytic center of the spliceosome' RNA, vol. 15, no. 12, pp. 2161-2173. DOI: 10.1261/rna.1908309

Digital Object Identifier (DOI):
10.1261/rna.1908309

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
RNA

Publisher Rights Statement:
Freely available online through the RNA Open Access option

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Physical and genetic interactions of yeast Cwc21p, an ortholog of human SRm300/SRRM2, suggest a role at the catalytic center of the spliceosome

RICHARD J. GRAINGER,1 J. DAVID BARRASS,1 ALAIN JACQUIER,2 JEAN-CHRISTOPHE RAIN,3 and JEAN D. BEGGS1

1Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh, EH9 3JF, United Kingdom
2Institut Pasteur, CNRS URA2171, Génétique des Interactions Macromoléculaires, 75015 Paris, France
3Hybrigenics S.A., 75014 Paris, France

ABSTRACT

In Saccharomyces cerevisiae, Cwc21p is a protein of unknown function that is associated with the NineTeen Complex (NTC), a group of proteins involved in activating the spliceosome to promote the pre-mRNA splicing reaction. Here, we show that Cwc21p binds directly to two key splicing factors—namely, Prp8p and Snu114p—and becomes the first NTC-related protein known to dock directly to U5 snRNP proteins. Using a combination of proteomic techniques we show that the N-terminus of Prp8p contains an intramolecular fold that is a Snu114p and Cwc21p interacting domain (SCwid). Cwc21p also binds directly to the C-terminus of Snu114p. Complementary chemical cross-linking experiments reveal reciprocal protein footprints between the interacting Prp8 and Cwc21 proteins, identifying the conserved cwf21 domain in Cwc21p as a Prp8p binding site. Genetic and functional interactions between Cwc21p and Isy1p indicate that they have related functions at or prior to the first catalytic step of splicing, and suggest that Cwc21p functions at the catalytic center of the spliceosome, possibly in response to environmental or metabolic changes. We demonstrate that SRm300, the only SR-related protein known to be at the core of human catalytic spliceosomes, is a functional ortholog of Cwc21p, also interacting directly with Prp8p and Snu114p. Thus, the function of Cwc21p is likely conserved from yeast to humans.

Keywords: protein interaction; Prp8; cwf21; Snu114; splicing

INTRODUCTION

The removal of introns from precursor-messenger RNAs (pre-mRNAs) is performed by RNA splicing, a process that is evolutionarily conserved in eukaryotes from yeast to humans. It is carried out in the nucleus by a large and highly dynamic RNA–protein complex called the spliceosome that contains over 100 proteins and five small nuclear ribonucleoprotein particles, the U1, U2, U4, U5, and U6 snRNPs (Bessonov et al. 2008; Wahl et al. 2009).

Many conformational rearrangements take place during the assembly, activation, and disassembly of the spliceosome and between the two chemical reactions. First, U1 snRNP binds to the 5′ splice site on the pre-mRNA, then the ATP-dependent association of U2 snRNP at the intron branch site creates a precursor of the spliceosome referred to as A complex. Binding of the preformed U4/U6•U5 triple snRNP to the A complex creates the B complex, then a series of ATP-dependent structural rearrangements dissociate the U1 and U4 snRNPs, and association of the NineTeen Complex (NTC in yeast; Prp19–CDC5 complex in humans) of proteins leads to formation of the activated B* complex that is primed for catalysis. C complex is generated by the first splicing reaction and the RNA products of this step are repositioned as substrates for second-step catalysis (Bessonov et al. 2008; Konarska 2008; Wahl et al. 2009). Members of the DExD/H-box family of ATPases that utilize ATP hydrolysis to unwind RNA duplexes or to disrupt RNA–protein interactions promote these rearrangements throughout the splicing cycle (Staley and Guthrie 1998).

Prp8p is a very large and highly conserved U5 snRNP protein, and the only spliceosomal protein known to
interact (within 2 Å) with all the reactive pre-mRNA residues, at the 5′ splice site, branch site, and 3′ splice site. It is proposed to be a protein cofactor at the spliceosomal catalytic center (Collins and Guthrie 2000; Grainger and Beggs 2005; Abelson 2008; Butcher 2009; Valadkhani and Manley 2009; Wilson and Lilley 2009). Prp8p is tightly associated with two other U5 snRNP proteins, the GTPase, Snu114p, and the DEAH-box RNA helicase, Brr2p (Achsel et al. 1998). It appears that the GTP/GDP state of Snu114p directly or indirectly (possibly via Prp8p) regulates the unwindase activity of Brr2p; therefore, proteins that interact directly with Snu114p are potentially of great importance in spliceosome activation and control.

Prp19p-associated complexes containing up to 26 proteins, have been identified by different procedures, including affinity selection of Prp19p itself (NTC) (Tsai et al. 1998; Chen et al. 2002) or by association with the NTC component Cef1p (complexed with Cef1p [CWC]) (Ohi et al. 2002). Cwc21p is found associated with a large number of RNA splicing proteins, especially the NTC, the U2 snRNP, and U5 snRNP (Ohi et al. 2002; Krogan et al. 2006; Khanna et al. 2009). As an NTC or NTC-related protein, Cwc21p is in a position to function within the spliceosome B+ and C complexes during the stages of spliceosome activation and catalysis. Cwc21p is a short protein of 135 residues whose primary sequence has been aligned with the N-terminal half of SRm300 encoded by the SRRM2 gene (Blencowe et al. 2000; Luz Ambrosio et al. 2009) and they are thought to be orthologs, although this has not been confirmed experimentally. SRm300 was first isolated through its stringent copurification with SRm160 (Blencowe et al. 1999). SRm300 has a very long C-terminal extension (~2600 amino acids) containing stretches of polyserine and numerous Arg/Ser di-peptide (RS domain) sequences that interact with other proteins and pre-mRNAs involved in alternative splicing events (Hertel and Graevesky 2005; Long and Caceres 2009). The potential importance of SRm300 in the catalytic center was highlighted when treatment of the purified human spliceosomal C complex with 1 M NaCl revealed SRm300 to be the only SR-related protein to remain at the catalytic core (Bessonov et al. 2008; Konarska 2008). SRm300 can be purified most strongly in the human C complex, suggesting that SRm300 may function after the first chemical reaction (Bessonov et al. 2008). Here we confirm that human SRm300 is an ortholog of yeast Cwc21p.

Through extensive proteomic and genetic experiments, we show that Prp8p contains an intramolecular fold, and that Snu114p’s N-terminal GTPase domain interacts with this novel Prp8p structure. Cwc21p also binds directly to this fold in Prp8p and to the C terminus of Snu114p, a region thought to undergo a substantial conformational change during spliceosome activation. These, plus a genetic interaction with Isy1p, an NTC protein that has been implicated in splicing fidelity, suggest a role for Cwc21p/SRM300 as a regulatory factor in the catalytic center of the spliceosome.

RESULTS

Prp8p contains a novel intramolecular fold that interacts with Snu114p

Prp8p is unusually large (2413 amino acids in Saccharomyces cerevisiae; 2335 amino acids in humans), and the full-length protein performs poorly in yeast two-hybrid (Y2H) protein interaction screens (van Nues and Beggs 2001) as well as in Y2H assays (Liu et al. 2006). A genetic screen of random transposon insertions in PRP8 identified three positions where Prp8p can be split into pairs of polypeptides that function in trans in vivo, suggesting the existence of four functional regions, I–IV (Fig. 1A; Boon et al. 2006). We, therefore, carried out Y2H screens using functional region I (Prp8p1–427) or regions I plus II (Prp8p1–770) as bait polypeptides.

Prp8p1–427 as bait identified some previously known interactors, including Prp40p and Prp39p (Supplemental Table S1; Abovich and Rosbash 1997; van Nues and Beggs 2001) and also a number of overlapping clones of Prp8p itself that have 123 amino acids in common (Fig. 1B, amino acids 420–542). The Prp8p420–542 polypeptide was then used as bait in Y2H screen 2 (Supplemental Table S2), identifying 63 overlapping clones of Prp8p domain I, which encompass residues 342–423. These reciprocal results from screens 1 and 2 indicate an intramolecular interaction involving amino acids 342–542 of Prp8p, which span the transposon-partitioning sites between regions I and II. All of the Y2H interactions presented here are of high statistical significance as they represent multiple overlapping clones and are reciprocal.

Additionally, Prp8p420–542 selected fragments of the N-terminal half of Snu114p in screen 2. This complements a previous Y2H result, in which full-length Snu114p selected Prp8p fragments from this region (overlapping at amino acids 420–464) (Dix et al. 1998; Grainger and Beggs 2005). These residue numbers refer to the minimal regions of interaction of overlapping prey polypeptides, and larger clones produce stronger interactions. In a pairwise Y2H test, the interaction between the slightly larger Prp8p253–543 and full-length Snu114p appears to be strong, based on resistance to 200 mM 3-amino triazole (Fig. 1C, 3-AT). This is consistent with the strong interaction reported between human Prp8p and Snu114p in native complexes (Achsel et al. 1998).

Cwc21p interacts directly with Prp8p and Snu114p

In Y2H screen 3 (with the same prey library), Prp8p1–770 as bait identified some previously known interactors, including Prp40p and Lin1p/Snu40p (Supplemental Table S3).
Prp8p1–770 also selected 20 clones of the NTC-associated Cwc21p, all of which contain a central region encompassing amino acids 53–117 (Fig. 1D).

One of the larger prey fragments, Cwc21p 16–135, was then used as bait in a fourth Y2H screen, and pulled out 11 Prp8p clones, with overlap of amino acids 252–500 (Supplemental Table S4), a region that corresponds approximately with the proposed Prp8p intramolecular fold. The failure to select any of the smaller Prp8p polypeptides that were identified in screens 1 and 2, suggests that Cwc21p requires a more complex binding site for its interaction with Prp8p. This is supported by the finding that neither Prp8p1–427 nor Prp8p420–542 interacted with Cwc21p16–135 in pairwise Y2H tests (data not shown). Screen 4 also identified the C terminus of Snu114p (amino acids 693–951) as interacting with Cwc21p (Fig. 1D). Therefore, overall, the Y2H screens show a triangulation of protein–protein interactions between the intramolecular fold of Prp8p, Cwc21p, and each end of Snu114p.

Recombinant proteins fused at the N terminus to GST or hexa-histidine tags were produced in bacteria, including two Prp8p polypeptides, Prp8p16–135 and Prp8p253–500. In pull-down assays, recombinant Cwc21p16–135 pulled down Prp8p253–543 more efficiently (>5% of total) than Prp8p253–500 (Fig. 1E, <5% of total), suggesting that amino acids 500 to 543 of Prp8p contribute to binding Cwc21p.

Thus, Cwc21p interacts directly with Prp8p and with a C-terminal portion of Snu114p.

Recombinant proteins fused at the N terminus to GST or hexa-histidine tags were produced in bacteria, including two Prp8p polypeptides, Prp8p16–135 and Prp8p253–500. In pull-down assays, recombinant Cwc21p16–135 pulled down Prp8p253–543 more efficiently (>5% of total) than Prp8p253–500 (Fig. 1E, <5% of total), suggesting that amino acids 500 to 543 of Prp8p contribute to binding Cwc21p.

Thus, Cwc21p interacts directly with Prp8p and with a C-terminal portion of Snu114p.

**Reciprocal Prp8p and Cwc21p binding footprints identified using BS3 cross-linking**

BS3 is a homo-bifunctional cross-linking agent, containing two terminal amine-reactive esters linked by an 11.4 Å spacer that reacts specifically with primary amine groups (–NH2) on available lysines as well as with the free amino-terminal residue of proteins. BS3 has been commonly used to identify nearest neighbor protein contacts in multiprotein complexes (Sinz 2006; Sharon and Robinson 2007). We used BS3 cross-linking to elucidate the arrangement of Prp8p’s intramolecular fold and to map the protein interaction footprints between Prp8p and Cwc21p.

Prp8p1–770 was incubated separately or together with BS3, following which a slowly migrating cross-linked Cwc21p–BS3–Prp8p species could be identified in denaturing gels. After separate in-gel digestion of the proteins with three different proteases, mass spectrometric analysis (MALDI-TOF) was performed. The total peptide masses were compared for (1) each individual protein with no BS3 treatment (negative control), (2) each individual protein plus BS3 (showing intramolecular cross-links), and (3) the two proteins mixed together plus BS3 (to deduce intermolecular cross-links). The resulting masses were analyzed through the online software Automated Spectrum Assignment Program (ASAP) (Lee et al. 2007) and SearchXlinks (Wefing et al. 2006). Only masses that uniquely identified

Prp8p1–770 was incubated separately or together with BS3, following which a slowly migrating cross-linked Cwc21p–BS3–Prp8p species could be identified in denaturing gels. After separate in-gel digestion of the proteins with three different proteases, mass spectrometric analysis (MALDI-TOF) was performed. The total peptide masses were compared for (1) each individual protein with no BS3 treatment (negative control), (2) each individual protein plus BS3 (showing intramolecular cross-links), and (3) the two proteins mixed together plus BS3 (to deduce intermolecular cross-links). The resulting masses were analyzed through the online software Automated Spectrum Assignment Program (ASAP) (Lee et al. 2007) and SearchXlinks (Wefing et al. 2006). Only masses that uniquely identified
BS\textsuperscript{3} linked peptides were analyzed further. Some lysine residues seemed to be more reactive than others but, for simplicity, each lysine was scored as either cross-linked or uncross-linked. This allowed the identification of solvent accessible lysines in Prp8\textsuperscript{253–543} and Cwc21p that are \(~\sim\)11 Å apart within their tertiary structures. The loss of reactivity of some of these lysines when Prp8p and Cwc21p are incubated together highlights the reciprocal “lysine footprints” on the two proteins.

For Prp8p\textsuperscript{253–543} incubated alone with BS\textsuperscript{3}, all lysines except K447 and K474 became cross-linked (Fig. 2A; Supplemental Table S5) and were therefore accessible to the BS\textsuperscript{3}. These Prp8p intramolecular cross-links show that the two ends of Prp8p\textsuperscript{253–543} are in close proximity, suggesting that Prp8p\textsuperscript{253–543} forms an intramolecular fold around a solvent inaccessible core that encompasses the amino acids K447 and K474.

Upon interaction of Prp8p\textsuperscript{253–543} with Cwc21p the lysines at the protein–protein interface are likely to be inaccessible to BS\textsuperscript{3}. There are two “lysine-footprints” (amino acids 300–332 and 491–516) in Prp8p\textsuperscript{253–543} as a result of the interaction with Cwc21p, and one Prp8p footprint in the center of Cwc21p (amino acids 53–97) (Fig. 2A). An alternative explanation for the loss of reactivity of lysines in these regions is that when the two proteins interact they undergo conformational changes that make these lysines inaccessible. However, these regions are flanked by intermolecular cross-links, indicating that they

![FIGURE 2. Prp8p contains an intramolecular fold that can be cross-linked to the conserved cwf21 domain in Cwc21p. (A) Schematic diagram showing the intramolecular BS\textsuperscript{3} cross-links identified by mass spectrometry for recombinant Prp8\textsuperscript{253–543} and Cwc21\textsuperscript{16–135}, with all lysines indicated by circles. Cross-links connecting lysines or peptides containing two lysines are shown by a line. Only uniquely assigned lysine/peptide cross-links are shown. The complete data are tabulated in Supplemental Table S5. The positions of three U4cs-1 suppressor alleles are indicated in italics (L280P is also a prp28-1 suppressor) and DBF3 indicates the position of a 34 amino acids polymorphism (Shea et al. 1994). Binding Prp8p to Cwc21p resulted in a number of previously reactive lysines (black circles) now showing no reactivity to BS\textsuperscript{3} (open circles); these were plotted to show reciprocal “lysine footprints.” Inset: cartoon summarizing the interactions shown in Figures 1 and 2. The intermolecular BS\textsuperscript{3} cross-links indicate that the two interacting regions of Prp8p run antiparallel to each other in this novel structure that we refer to as SCwid (residues 253–543). The location where transposon insertions are tolerated (Boon et al. 2006) is indicated by scissors. The approximate positions of domains I–V in Snu114p are shown. (B) Alignment of the Cwc21p/3Rm300/SRRM2 family showing the highly conserved regions as calculated by the Muscle program (Edgar 2004) and highlighted using Jalview (Waterhouse et al. 2009) (numbers above refer to budding yeast Cwc21p). The Prp8p “lysine footprint” and pfam’s cwf21 domain of unknown function are indicated; their near complete overlap suggests that this highly conserved domain functions to bind directly to Prp8p. (C) An additional conserved protein known as SRRM2-like (SRRM2L) also contains the cwf21 domain near its N terminus; however, the N terminus itself is different from Cwc21p/3Rm300 and its function is unknown.](image-url)
are indeed sites of interaction of the two proteins. These results are in agreement with the Y2H data and with the recombinant protein binding studies. The two separate footprints on Prp8p can explain why Cwc21p interacts with the large Prp8p253–500 fragment, but not with smaller footprints on Prp8p can explain why Cwc21p interacts with recombinant protein binding studies. The two separate results are in agreement with the Y2H data and with the are indeed sites of interaction of the two proteins. These interacts genetically with CWC21 a partial structure is known (PDB 2e62).

A genome-wide accumulation of introns, coupled with a corresponding decrease in the mature mRNA levels, was observed with RNA from the mutant strain grown at 30°C, and the splicing defect increased after 30 min at 37°C (summarized in Fig. 3C; Supplemental Table S6). Among the top 40 most affected transcripts whose functions are known, 37 encode ribosomal proteins, translation factors (EFB1, TEF4) or a snoRNA involved in ribosome synthesis (SNR17A/B). It was previously reported that, compared with other transcripts, the splicing of ribosomal protein transcripts shows a distinct response to conditions of stress, such as amino acid starvation and exposure to ethanol (Pleiss et al. 2007). It is therefore conceivable that the preferential accumulation of unspliced ribosomal protein transcripts observed here is a consequence of an enhanced stress response in the cwc21Δ/isy1Δ strain. An examination of the transcripts that showed strong splicing defects in the cwc21Δ/isy1Δ strain showed that the 5’ splice site, branch site, and 3’ splice site

Cwc21p, SRm300, SRRM2L, and the cwf21 domain

An alignment of the Cwc21p family sequences (Fig. 2B) shows a conserved region of around 100 amino acids at the N terminus of each protein, and highlights three features: (1) a conserved N-terminal motif of about 24 amino acids, (2) a variable linker region, and (3) a conserved cwf21 domain (pIam 08312, IPRO13170) (Finn et al. 2008), named after cwf21, the Schizosaccharomyces pombe ortholog. The Prp8p binding site on Cwc21p corresponds exactly to the cwf21 domain, which can explain the evolutionary conservation of this region. Proteins of the Cwc21p/SRm300 family are usually larger in higher eukaryotes due to extended C-termini that can be 2000 or more amino acids long. These extensions are made up of poly-serine stretches and RS di-peptide repeats, the latter being known to direct protein–protein or protein–RNA interactions (Hertel and Graveley 2005). However, cDNA clones indicate production of a smaller 194–196 residue SRm300/SRRM2 protein in many tissues in higher eukaryotes (Strausberg et al. 2002). A BLAST search with the cwf21 domain identified another protein, known as SRRM2-Like (SRRM2L) protein (Fig. 2C) that is usually ~600–700 amino acids long. SRRM2L is only present in vertebrates and conceivably may have a role in development. Curiously, the U2-associated SR140 protein has a partial cwf21 domain and a partial structure is known (PDB 2e62).

CWC21 interacts genetically with ISY1

CWC21 is not required for cell viability under normal growth conditions. A sensitive and quantitative genetic interaction mapping (GIM) screen was carried out to determine if deletion of CWC21 has a synthetic genetic interaction with any of the other ~4400 nonessential genes in yeast (Decourty et al. 2008). The GIM screen with cwc21Δ identified a cluster of functionally related genes that are highlighted in Figure 3A. The five deletions that showed the greatest growth defect when combined with cwc21Δ included four genes that encode RNA splicing factors: Isy1p, Ntc20p, Syf2p, and Snu66p. Isy1p, Ntc20p, and Syf2p are NTC components and are part of a stable NTC90 subcomplex, named after the essential component Syf1p (also known as Ntc90p) (Chen et al. 2002; Dahan and Kupiec 2002). Snu66p is a component of U4/U6/U5 tri-snRNPs and deletion of SNU66 enhances the growth defects of several snu114 alleles (Brenner and Guthrie 2005). This supports a role for Cwc21p in pre-mRNA splicing and with the NTC, a role that is supported by other genetic studies (Haussmann et al. 2008; Wilmes et al. 2008; Khanna et al. 2009).

A more unexpected synthetic interaction that was observed with deletion of ERD1 may be explained by an indirect genetic link. ERD2 is one of a small number of intron-containing transcripts that are poorly spliced (<50%) in a cwc21Δ strain under normal growth conditions (Burckin et al. 2005). ERD2 is an essential, intron-containing gene that interacts with ERD1 (Miller et al. 2005), with both proteins functioning to retain KDEL proteins in the endoplasmic reticulum (ER). Thus, cwc21Δ causes reduced levels of Erd2p and the additional loss of Erd1p may explain the growth defect. In support of this, another genome-wide screen showed that Cwc21p has a significant role in ER homeostasis; deletion of CWC21 causes unregulated ER to Golgi transport (Copic et al. 2009).

Testing the double deletions of cwc21Δ with erd1Δ, ntc20Δ, or syf2Δ did not show a significant growth defect on a plate assay (less sensitive than the liquid assay used in the GIM screen) (data not shown); however, cwc21Δ/isy1Δ grew slowly at 30°C and failed to grow at 37°C (Fig. 3B). As the cwc21Δ/isy1Δ temperature-sensitive phenotype is fast acting, the effect of temperature on splicing was investigated using an oligo-based microarray that detects all of the intron-containing and spliced mRNA transcripts in S. cerevisiae. A genome-wide accumulation of introns, whose functions are known, 37 encode ribosomal proteins, translation factors (EFB1, TEF4) or a snoRNA involved in ribosome synthesis (SNR17A/B). It was previously reported that, compared with other transcripts, the splicing of ribosomal protein transcripts shows a distinct response to conditions of stress, such as amino acid starvation and exposure to ethanol (Pleiss et al. 2007). It is therefore conceivable that the preferential accumulation of unspliced ribosomal protein transcripts observed here is a consequence of an enhanced stress response in the cwc21Δ/isy1Δ strain. An examination of the transcripts that showed strong splicing defects in the cwc21Δ/isy1Δ strain showed that the 5’ splice site, branch site, and 3’ splice site
sequences do not differ significantly from the consensus sequences. *ACT1* transcripts were then analyzed by quantitative real-time PCR to detect unspliced pre-mRNA or the lariat form of the intron. A >20-fold accumulation of unspliced *ACT1* pre-mRNA was observed in the double mutant compared with wild type at the restrictive temperature, with very little or no lariat species, indicating that the *cwc21Δ/isy1Δ* strain has a block at or prior to the first catalytic step of splicing (Fig. 3D). In comparison, RNA from the *isy1Δ* and *cwc21Δ* single mutants showed only a mild or no significant splicing defect, respectively. The much greater accumulation of pre-mRNA in the double mutant compared with either single mutant can explain their synthetic genetic interaction, and suggests that Cwc21p and Isy1p may have related functions at or prior to the first step of splicing, making them partially redundant. However, whereas deletion of *isy1Δ* has been shown to rescue the cold-sensitive phenotype of *prp16-302*, this is not the case with *cwc21Δ* (data not shown), indicating that the two proteins do not have an identical function.

**Cwc21p is orthologous to human SRm300**

The temperature sensitivity caused by deletion of both *CWC21* and *ISY1* provides a test system for the identification of functionally important *cwc21* mutations as well as for complementation by *CWC21* orthologs. Thus, the N-terminal 159 amino acids of human SRm300 was shown to partially suppress the temperature-sensitive growth defect of *cwc21Δ/isy1Δ* cells at 36°C (Fig. 4A). In this experiment the Cwc21p and SRm3001–159 proteins were present at the same level, expressed from the PMET25 promoter (data not shown); therefore, the slower growth observed with SRm3001–159 probably reflects the cumulative evolutionary divergence between the two proteins that likely affects their interactions. Thus, human SRm300 functions as an ortholog of yeast Cwc21p.

Comparing the properties of the purified recombinant Cwc21 and SRm3001–159 proteins in vitro, SRm3001–159 was shown to pull down recombinant yeast Prp8p253–543, although somewhat less efficiently than Cwc21p (Fig. 4B).
combined with His 6-fusions of Prp8 and Snu114 (residues as in Cwc21p. The severity of the growth defects observed with not shown), the mutations did not completely inactivate mostly rescued by overproducing the mutant proteins (data on Prp8p extends into residues 500–543.

yeast Snu114p692–951 interacted equally well with human acids 500–543 for binding to Cwc21 proteins. In contrast, SRm3001–159 did not detectably pull down the shorter SRm300–N, indicating). The pull downs (Co-P) indicate that human SRm300 with yeast Prp8p.

The cwf21 domain in Cwc21p is important in vivo

The cwc21Δ/isy1Δ complementation assay was then used to test the effect of mutating conserved residues in the Prp8p binding site of Cwc21p, encoded on a plasmid. Compared with wild-type Cwc21p, the mutant variants H67A, R71D, and I73A resulted in slower growth at 37°C, with R71D having the most severe defect (Fig. 5A). As the temperature-sensitive growth of the cwc21Δ/isy1Δ strain could be mostly rescued by overproducing the mutant proteins (data not shown), the mutations did not completely inactivate Cwc21p. The severity of the growth defects observed with the mutant proteins correlated with the effect of the mutations on association of Cwc21p with Prp8p and Snu114p in yeast extract, with R71D having the most severe effect (Fig. 5B, cf. lanes 2, 3 and 4). Thus, mutation of three conserved residues in the cwf21 domain confirmed the importance of this motif for binding of Cwc21p to Prp8p in vivo and suggests that this interaction is important for normal cell growth in the absence of Isy1p.

The SCwid domain in Prp8p

We show that Prp8p contains a novel intramolecular fold that we have termed the Snu114/Cwc21 interacting domain or SCwid, to which Cwc21p/SRm300 binds directly, via the conserved cwf21 domain. Cwc21p also binds directly to the C-terminus of Snu114p, a region thought to undergo a conformational change when Snu114p binds GTP. The region of Snu114p that contains the GTP binding domain (spanning residues 134–510) interacts very strongly with Prp8p’s SCwid. It has been speculated that Prp8p could be a potential GTPase-activating protein (GAP) for Snu114p. Consistent with this, Luz Abrosio et al. (2009) showed that RNAi knockdown of Trypanosome U5–Cwc21 exhibited an effect on splicing prior to the first step, and Khanna et al. (2009) found yeast Cwc21p to be specifically associated with U2, U5, and U6 snRNAs, which is compatible with a role in spliceosomes after dissociation of U1 and U4.

**DISCUSSION**

We demonstrate here that Cwc21p, an NTC-associated protein is orthologous to the human splicing factor SRm300 and that both can bind directly to two key splicing factors; Prp8p, a likely cofactor in splicing catalysis, and Snu114p, the only GTPase in the spliceosome. Cwc21p interacts genetically with ISY1, a proposed fidelity factor that promotes a spliceosomal conformation favorable for first-step chemistry (Villa and Guthrie 2005; see below). We propose that Cwc21p functions at a similar stage in the splicing cycle, at or prior to the first chemical step of splicing, possibly through its interaction with Prp8p and Snu114p. Consistent with this, Luz Abrosio et al. (2009) showed that RNAi knockdown of Trypanosome U5–Cwc21 exhibited an effect on splicing prior to the first step, and Khanna et al. (2009) found yeast Cwc21p to be specifically associated with U2, U5, and U6 snRNAs, which is compatible with a role in spliceosomes after dissociation of U1 and U4.
function in trans (Boon et al. 2006). This can now be explained; the two separate polypeptides of Prp8p can be stably held together by the SCwid structure, in which the two regions interact in an antiparallel arrangement (see Fig. 2A, inset).

The existence of the SCwid structure in Prp8p may also explain how two groups of U1 snRNP proteins appear to bind to distant regions of Prp8p (Grainger and Beggs 2005, and references therein). Prp40p and Prp39p interact N-terminally, whereas Exo84p and Snp1p interact near the middle of Prp8p, and the SCwid fold could bring these sites into proximity (Fig. 6). Ten U4-cs1 mutations lie in the RRM region, and three U4-cs1 alleles plus one prp28-1 suppressor allele lie in the SCwid (Kuhn and Brow 2000; Kuhn et al. 2002; Grainger and Beggs 2005), all of which are involved in exchanging U1 and U4 snRNPs (Chan et al. 2003). These mutations do not affect the Y2H interactions described here, and may affect other interactions. As Cwc21p has been isolated along with Prp19p/Cef1p and the NTC/CWC complexes (Ohi et al. 2002), the SCwid is a potential docking site for the NTC/CWC or NTC-related proteins (Fig. 6). The NTC associates with spliceosomes after displacement of U1 and U4 snRNPs (Chan et al. 2003). It is conceivable that the SCwid and Snu114p’s C terminus are recognized by Cwc21p only after release of U1 snRNP from Prp8p and upon the exchange of GDP for GTP in Snu114p, hence Prp8’s SCwid and Cwc21p have the potential to play a major role in spliceosome activation.

Cwc21p in the regulation of splicing
Cwc21p is not essential for cell growth, and depletion of SRm300 was shown to be nonessential (under limited conditions tested) (Blencowe et al. 2000), which is compatible with function or activation of these proteins only under certain conditions. Interestingly, the Caenorhabditis elegans ortholog is essential; rsr-2 (related SR protein -2) has been shown by RNA interference to cause severe abnormalities such as sterility, embryonic lethality, larval arrest (Longman et al. 2001), multivulva and extra intestinal nuclei (Ceron et al. 2007), suggesting a possible role in development. CWC21 is one of only 125 yeast genes that are up-regulated in response to 200 mM Ca²⁺, with the level of CWC21 expression increasing 3.5-fold (Yoshimoto et al. 2002). Analysis of the promoter region upstream of CWC21 reveals a binding site for Crz1p, a factor activated step 1. These mutations do not affect the Y2H interactions described here, and may affect other interactions. As Cwc21p has been isolated along with Prp19p/Cef1p and the NTC/CWC complexes (Ohi et al. 2002), the SCwid is a potential docking site for the NTC/CWC or NTC-related proteins (Fig. 6). The NTC associates with spliceosomes after displacement of U1 and U4 snRNPs (Chan et al. 2003). It is conceivable that the SCwid and Snu114p’s C terminus are recognized by Cwc21p only after release of U1 snRNP from Prp8p and upon the exchange of GDP for GTP in Snu114p, hence Prp8’s SCwid and Cwc21p have the potential to play a major role in spliceosome activation.
by specific environmental stress conditions such as exposure to osmotic shock, mating pheromones, and elevated temperatures (Yoshimoto et al. 2002; Monteiro et al. 2008). Expression of SRm300/SRRM2 also increases by 2.6-fold upon induction by calcium (Zagranichnaya et al. 2005), suggesting a conserved mode of regulation. The link between Cwc21p, ERD2 pre-mRNA splicing, and ER homeostasis (Copic et al. 2009) discussed above may be relevant to this.

We found that the cwc21Δ/isy1Δ double-mutant strain displayed heat-sensitive growth and a first step splicing defect that was enhanced at elevated temperature. This resulted in preferential accumulation of transcripts that encode ribosomal proteins or other factors involved in translation. Similarly, in the accompanying paper Khanna et al. (2009) through a tiling array analysis of cwc21Δ, isy1Δ single and double mutants grown at the permissive temperature revealed significant enrichment of introns for genes whose functional categories include translation, post-translational modification and reproduction. Therefore, Cwc21p and/or Isy1p may regulate the splicing of these categories of transcripts in response to stresses such as high temperature or changes in other cellular conditions. Alternatively, the preferential accumulation of unspliced transcripts encoding ribosomal proteins and translation factors may reflect the fact that they are very abundant and therefore more likely to be affected by a splicing defect. Conceivably, a conformational change or post-translational modification in the SCwId or Cwc21p could immediately inhibit spliceosome activation by the NTC. This suggests a potential mechanism whereby Cwc21p might affect the splicing of some or all transcripts in response to environmental or metabolic changes (Pleiss et al. 2007).

**SRm300 is a human ortholog of Cwc21p**

The sequence similarities between Cwc21p and the N terminus of the human SRm300 have been noted previously (Blencowe et al. 2000; Luz Ambrosio et al. 2009). We confirm here that the N-terminal 159 amino acids of human SRm300 can functionally substitute for yeast Cwc21p in vivo and that both bind directly to the yeast Prp8p SCwId via a conserved region called the cwf21 domain. Mutating the cwf21 domain confirmed its function in vivo and that arginine 71 is critical for the interaction with Prp8p and Snu114p. Careful alignments show that higher eukaryotes have another potential Prp8p binding protein called SRRM2-like protein. Its function is unknown, but its cwf21 domain allows us to speculate that it might be a vertebrate-specific splicing factor. Considering that SRm300 is the only SR-related protein found at the center of the human spliceosome, and most strongly in the C complex (after step 1) (Bessonov et al. 2008; Konarska 2008) we propose that its presence at the catalytic center may be mediated by its interactions with hPrp8 and hSnu114. Conceivably, the long RS-domain of SRm300 could transmit responses from different environmental stimuli directly to the catalytic center.

**Cwc21p genetics, its role in splicing, and its relationship with Isy1p**

The GIM screen showed that Cwc21p interacts genetically with a cluster of known NTC proteins, including Isy1p, Syf2p, and Ntc20p, confirming Cwc21p’s close role with the NTC, as well as with the tri-snRNP protein Snu66p. As in the accompanying paper (Khanna et al. 2009), we also found that isy1Δ produced the greatest synthetic defect with cwc21Δ. The most useful genetic tool we found was that the cwc21Δ/isy1Δ double mutant produced a temperature-sensitive phenotype, and was shown by microarray analysis and quantitative PCR to inhibit pre-mRNA splicing at or prior to the first step of pre-mRNA splicing. Indeed, isy1Δ itself decreases the efficiency of the first step of splicing (Fig. 3D; Villa and Guthrie 2005; Khanna et al. 2009), and it was proposed that Isy1p promotes a spliceosomal conformation favorable for first-step chemistry (Villa and Guthrie 2005). The synthetic genetic interaction between ISY1 and CWC21 and the enhanced firststep splicing defect in the double mutant suggest that they may function in the same complex, or that their functions are mutually redundant. Although Cwc21p and Isy1p are functionally related, their functions seem not to be identical, as evidenced by suppression of prp16–302 by isy1Δ but not by cwc21Δ.

Prp16p is a DEAH-box RNA-stimulated ATPase that appears to interact only transiently with spliceosomes to promote a structural transition required for second-step chemistry (Schwer and Guthrie 1991; Schwer and Guthrie 1992). The cold-sensitive prp16Δ–302 allele causes the protein to stall in spliceosomes at low temperature, and pre-mRNAs with aberrant branch sites are spliced at increased frequency. The cold-sensitive growth caused by this prp16 allele is suppressed by mutations that weaken a U2/U6 snRNA structure called helix I in the catalytic center, and it was proposed that Prp16p destabilizes helix I to allow a rearrangement of the catalytic center for the second step of splicing (Meford and Staley 2009). The prp16–302 defect is also suppressed by loss of the NTC protein Isy1p that, although not essential, is required for optimal splicing of certain poorly spliced introns (Dix et al. 1999). Deletion of ISY1 appears to release the stalled prp16–302 spliceosome and restore fidelity of branch site usage (Villa and Guthrie 2005). It was proposed that Isy1p functions together with U6 snRNA to promote the spliceosomal conformation for first-step chemistry (Villa and Guthrie 2005). According to a model that proposes an equilibrium between conformational states of the spliceosome (Konarska et al. 2006; Smith et al. 2008), deletion of ISY1 may promote second step chemistry, explaining the reduced fidelity of 3‘ splice
site selection that was observed with isy1Δ in the presence of wild-type Prp16p (Villa and Guthrie 2005).

We propose that Isy1p and Cwc21p function in two distinct but related events in the catalytic center of the spliceosome. The role of Cwc21p is mediated by its direct interaction with Prp8p and Snu114p and, possibly in response to changing environmental or metabolic conditions, Cwc21p helps to stabilize the catalytic center or the position of the RNA substrate. As discussed, Cwc21p likely functions at or immediately prior to the first chemical step of splicing, however, considering the apparent equilibrium between the various spliceosomal conformations and the report that SRm300 is strongly associated with spliceosomal C complex that is formed after the first chemical reaction, we do not exclude an effect on later events. Our future studies will therefore investigate the role of Cwc21p in the regulation of splicing and in splicing fidelity under various growth conditions.

MATERIALS AND METHODS

Yeast strains, plasmids, and oligonucleotides

Yeast strains and plasmids are listed in Tables 1 and 2, respectively. Yeast manipulations were performed using standard laboratory procedures. Deletion and C-terminal tagging of CWC21 on the yeast genome was by one step PCR from a variety of templates (Longtine et al. 1998) and correct integration was confirmed by yeast colony PCR. PCR was used to amplify of templates (Longtine et al. 1998) and correct integration was on the yeast genome was by one step PCR from a variety laboratory procedures. Deletion and C-terminal tagging of

Yeast two-hybrid

Y2H screens were performed by Hybrigenics S.A. using a mating strategy (Fromont-Racine et al. 2002). The Hybrigenics’ S. cerevisiae prey library was introduced in Y187 cells and mated with L40ΔG cells that produced LexA-bait fusion proteins (www.hybrigenics.com). Tests with pairwise protein combinations involved L40ΔG cells only.

Cloning, mutagenesis, and recombinant proteins

Recombinant plasmids for protein expression were produced using the Gateway system (Invitrogen). Mutagenesis of the Cwc21–13myc plasmids was achieved using a home-made version of the “Quick-Change protocol” of Stratagene. For recombinant protein production, Escherichia coli BL21(DE3)pLysS cells (Invitrogen) grown overnight at 37°C in 50 mL LB containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol were inoculated into 1 L of LB in a 5-L flask at 37°C. At OD600nm of 0.6 (about 1 h later), 0.238g of IPTG powder (1 mM) was added and the culture shifted to 30°C for 3–4 h. Cells were pelleted, resuspended in 30 mL of lysis buffer (20 mM HEPES at pH 7.5, 200 mM NaCl, 100 μM EDTA, 10 mM β-mercaptoethanol, 1 pellet of Roche protease inhibitors), and stored at −80°C. Lysis involved three cycles of freeze–thaw and all subsequent procedures were carried out at 4°C: addition of DNase I to 5 μg/mL and rotation for 5 min followed by the addition of 1% (v/v) Triton X-100 and rotation for 15 min, the lysate was then spun at 17,000 rpm for 45 min, clear supernate was removed, and proteins were affinity-purified on either GST-Sepharose or Nickel-NTA columns and dialyzed against 20 mM HEPES at pH 7.9, 100 mM NaCl, 100 μM EDTA, 1 mM DTT, 20% glycerol. Protein yield was estimated by Bradford assay and stored at −20°C.

Recombinant protein immunoprecipitations

All immunoprecipitations were carried out at 4°C. Preswollen glutathione beads were washed three times in NTN buffer (50 mM Tris at pH 7.5, 150 mM NaCl, 0.1% [v/v] Nonidet P40) and once in X1 IP150 (6 mM HEPES at pH 7.9, 150 mM NaCl, 5 mM MgCl2), and resuspended in 100 μL of X1 IP150. The beads (30 μL per sample) were added to a premixed combination of proteins in a final volume of 300 μL, and rotated end over end for 1 h at 4°C. Beads were pelleted and washed twice in NTN then twice more in NT (no NP-40) before resuspending in protein loading buffer. Western blotting and washes were performed using PBST-milk [PBS, 0.1% (v/v) Tween 20, 5% (v/v) nonfat dried milk] (Marvel). Incubation with primary antibodies: α-Prp8 (8.6) (Grainger and Beggs 2005) at 1/4000; α-Snu114 (Patrizia Fabrizio, Max Planck Institute for Biophysical Chemistry) at 1/50,000; and α-myck-HRP (Roche) (at 1/50,000), was overnight in PBST-milk at 4°C, and with secondary antibodies, α-Rabbit-HRP (1/20,000) was in PBST-milk at room temperature.

<table>
<thead>
<tr>
<th>TABLE 1. Yeast strains used in this work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>L40ΔG</td>
</tr>
<tr>
<td>BY4741</td>
</tr>
<tr>
<td>YRG-C21-13myc</td>
</tr>
<tr>
<td>YRG-C21Δ</td>
</tr>
<tr>
<td>Y16648-isy1Δ</td>
</tr>
<tr>
<td>YRGC21Δ/Δ</td>
</tr>
</tbody>
</table>
Yeast protein immunoprecipitations

We followed the same protocol as above except that after three washes in NTN buffer, 0.5 mL of c-myc 9E11 X Trans cruz (Santa Cruz Biotechnology) were coupled to Protein A-Sepharose CL-4B beads (GE Healthcare) in 300 mL of NTN buffer for 1 h with end-over-end rotation. Following 45 min with blocking solution (containing 100 mg/mL each of tRNA, BSA, Glycogen) the beads were washed three times with NTN buffer and once in X1 IP 150, then we followed the above protocol.

BS3 cross-linking

Approximately 20 mg of each protein were mixed in a final volume of 60 mL in X1 IP150 for 1 h at 4°C. BS3/DMSO solutions from 1 mM upward were added to the proteins for a further 2 h. The reaction was stopped by addition of 1 mL of 400 mM ammonium hydrogen carbonate, and after 5 min proteins were immunoprecipitated using glutathione beads as above. Protein gels were stained with GelCode (Pierce), bands cut out, and analyzed by MALDI-TOF mass spectrometry.

Microarray analysis and quantitative PCR

The design of the microarray used in this study was as described previously (Kershaw et al. 2009), and the arrays were printed at the Division of Pathway Medicine, University of Edinburgh. The protocols used can be found at http://www.biology.ed.ac.uk/research/groups/jbeggs/microarray/cwc21_isy1_37.

For the quantitative real-time PCR, a Roche transcriptor was used to produce cDNA according to the manufacturer’s instructions. The primer design for reverse transcription of the lariat species was based on Vogel et al. (1997). The real-time PCR was performed with Invitrogen Express SybrGreen as per the manufacturer’s instructions on a Stratagene Mx3005P. The data are presented in Supplemental Table S6 and at the above website.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at http://www.rnajournal.org.

POTENTIAL CONFLICT OF INTEREST DISCLOSURE

The two-hybrid screens were performed by the company Hybrigenics S.A. as a collaboration within the RNOMICS project.

ACKNOWLEDGMENTS

We are indebted to Ben Blencowe for communicating data prior to publication, for helpful comments on this manuscript, and for generously facilitating the copublication of our related papers. We thank Laurence Decourty, Christophe Malabat, and Cosmin Saveanu (Institut Pasteur), who performed the GIM screen, and Patrizia Fabrizio for providing excellent Snu114p antibodies. This work was funded by Wellcome Trust grant 067311 and European Commission funding for the European Framework Programme V RNOMICS project and the European Framework Programme 6 Network of Excellence EURASNET.

Received August 31, 2009; accepted September 15, 2009.

REFERENCES


