Mutations in the U11/U12-65K protein associated with isolated growth hormone deficiency lead to structural destabilization and impaired binding of U12 snRNA

ANTTO J. NORPPA,[1] TUULI M. KAUPPALA,[1] HARRI A. HEIKKINEN, BHUPENDRA VERMA,[2] HIDEO IWAÏ, and MIKKO J. FRILANDER

Institute of Biotechnology, FI-00014 University of Helsinki, Finland

ABSTRACT

Mutations in the components of the minor spliceosome underlie several human diseases. A subset of patients with isolated growth hormone deficiency (IGHD) harbors mutations in the *RNPC3* gene, which encodes the minor spliceosome-specific U11/U12-65K protein. Although a previous study showed that IGHD patient cells have defects in U12-type intron recognition, the biochemical effects of these mutations on the 65K protein have not been characterized. Here, we show that a proline-to-threonine missense mutation (P474T) and a nonsense mutation (R502X) in the C-terminal RNA recognition motif (C-RRM) of the 65K protein impair the binding of 65K to U12 and U6atac snRNAs. We further show that the nonsense allele is targeted to the nonsense-mediated decay (NMD) pathway, but in an isoform-specific manner, with the nuclear-retained *65K* long-3′UTR isoform escaping the NMD pathway. In contrast, the missense P474T mutation leads, in addition to the RNA-binding defect, to a partial defect in the folding of the C-RRM and reduced stability of the full-length protein, thus reducing the formation of U11/U12 di-snRNP complexes. We propose that both the C-RRM folding defect and NMD-mediated decrease in the levels of the U11/U12-65K protein reduce formation of the U12-type intron recognition complex and missplicing of a subset of minor introns leading to pituitary hypoplasia and a subsequent defect in growth hormone secretion.

Keywords: minor spliceosome; U11/U12 di-snRNP; U11/U12-65K; RNA–protein interactions; RNA recognition motif

INTRODUCTION

Pre-mRNA splicing is carried out by a complex machinery composed of small nuclear RNAs and >170 proteins (Will and Lührmann 2011). The extensive array of factors involved in splicing and the highly complex splicing code embedded in the RNA sequence allows fine-tuned regulation of the splicing process, but makes the system vulnerable to perturbations. Indeed, defects in splicing, arising from mutations in either splicing regulatory sequences or, more rarely, in the components of the splicing machinery, are a major cause of human genetic disease (Singh and Cooper 2012).

Most introns are excised by the major, U2-dependent spliceosome, but a small subset (>0.5%) containing highly conserved 5′ splice sites (5′ss) and branch point sequences (BPS) are removed by the minor, U12-dependent spliceosome (Turunen et al. 2013a). Both spliceosomes are composed of five small nuclear RNA (snRNA) molecules, which, upon association with protein components, form small nuclear ribonucleoprotein (snRNP) particles. Additionally, numerous non-snRNP proteins associate with the spliceosome components during the splicing process (Hastings and Krainer 2001; Wahl et al. 2009; Matera and Wang 2014). The two spliceosomes share the U5 snRNP, while the remaining four snRNPs in each spliceosome are distinct but functionally analogous, with U11, U12, U4atac and U6atac of the minor spliceosome replacing the respective U1, U2, U4 and U6 in the major spliceosome (Hall and Padgett 1996; Tarn and Steitz 1996a,b; Kolossova and Padgett 1997; Yu and Steitz 1997). While the overall assembly and catalytic steps of intron removal are very similar between the two spliceosomes (Tarn and Steitz 1996a,b; Frilander and Steitz 2001), there is a significant difference in the intron recognition step, which for minor introns is carried out by a preformed U11/U12 di-snRNP complex (Wassarman and Steitz 1992; Golas et al. 2005) that cooperatively recognizes the 5′ss and BPS (Frilander and

---

Steitz 1999). In contrast, the 5′ss and BPS of the major introns are recognized independently by individual U1 and U2 snRNPs. This functional difference is reflected in the composition of the U11/U12 di-snRNP, which, in addition to the two unique snRNAs, also contains seven protein species that are unique to the minor spliceosome (Will et al. 1999, 2004), while the other proteins associated with the minor spliceosome are shared with the major spliceosome (Hastings and Krainer 2001; Schneider et al. 2002).

Several studies in the past few years have implicated minor spliceosome dysfunction in human diseases (for review, see Verma et al. 2018). Mutations in the specific components of the minor spliceosome have been found to cause four rare congenital disorders, Microcephalic Osteodysplastic Primordial Dwarfism type I/Taybi-Linder syndrome (MOPD I/TALS) (Edery et al. 2011; He et al. 2011), Roifman syndrome (RFMN) (Merico et al. 2015), early-onset cerebellar ataxia (EOCA) (Elsaid et al. 2017) and a subset of isolated growth hormone deficiency cases (IGHD) (Argente et al. 2014; Guceva et al. 2015). The first three diseases result from mutations in the snRNA components of the minor spliceosome, specifically the U4atac snRNA in MOPD I/TALS and Roifman syndrome, and U12 snRNA in cerebellar ataxia, while IGHD is caused by mutations in a gene encoding the minor spliceosome-specific U11/U12-65K protein. Furthermore, somatic mutations in the *ZRSR2* gene, encoding the Urp protein of the U11/U12 di-snRNP that is involved in 3′ splice site (3′ss) recognition (Shen et al. 2010), underlie a subset of myelodysplastic syndrome (MDS) cases (Yoshida et al. 2011; Madan et al. 2015). Finally, recent studies also suggest a role for minor spliceosome dysfunction in the motor neuron diseases spinal muscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) (Doktor et al. 2016; Reber et al. 2016).

The *RNPC3* gene mutated in IGHD patients encodes an integral 65K protein component of the U11/U12 di-snRNP (Will et al. 2004). It contains two RNA recognition motifs (RRMs), of which the carboxy-terminal RRM (C-RRM) is used to bind to the 3′ terminal stem–loop of the U12 snRNA (Benecke et al. 2005). Additionally, a recent work demonstrated that the C-RRM also interacts with the near-identical 3′ terminal stem–loop of the U6atac snRNA (Singh et al. 2016). Within the U11/U12 di-snRNP the amino-terminal part of the 65K protein in turn interacts with the U11 snRNP-specific 59K protein (Benecke et al. 2005), which itself interacts with U11-48K, a protein that participates in recognition of the U12-type 5′ splice site with U11 snRNA (Turunen et al. 2008; Tidow et al. 2009). Thus, the U11/U12-65K protein has a central role in a protein–protein interaction network that connects the individual U11 and U12 snRNPs that function in the recognition of the 5′ and 3′ ends of the U12-type introns, respectively. Furthermore, these interactions are important for the stability of the di-snRNP complex (Turunen et al. 2008; Argente et al. 2014), are conserved from plants to mammals (Park et al. 2016) and provide an important regulatory target as the levels of U11-48K and U11/U12-65K proteins are both regulated at the post-transcriptional level by evolutionarily highly conserved feedback or cross-regulation mechanism (Verbeeren et al. 2010, 2017; Turunen et al. 2013b; Niemelä et al. 2015).

Presently, two recessive patient mutations have been reported to associate with IGHD: a proline-to-threonine missense mutation (P474T) within the C-RRM and a nonsense mutation (R502X) that deletes 16 residues (502–517) from the carboxyl terminus of the protein, immediately downstream from the core C-RRM domain (Fig. 1A; Argente et al. 2014). Patients are compound heterozygotes for the two mutations and show various defects in the splicing of U12-type introns. Besides the reduced stability of the di-snRNP observed in the patient cells, detailed understanding of the molecular consequences of the mutations is lacking. Here, we have carried out biochemical and structural analyses of the IGHD-causing 65K mutations. We found that the P474T mutation leads to significant unfolding of the C-RRM and a concomitant, approximately threefold drop in its affinity to U12 and U6atac snRNAs. The mRNA carrying the R502X mutation is efficiently targeted by the nonsense-mediated decay (NMD) machinery in patient cells and the truncated protein is thus likely to be expressed at a negligible level. Furthermore, analysis of the R502X mutant protein revealed that removal of the flexible tail region of the C-RRM leads to a near-complete loss of RNA-binding activity. In the cellular context, both mutants associate with U11/U12 di-snRNPs less efficiently than the wild-type protein, and di-snRNP complexes formed by the R502X mutant are additionally unstable. In contrast, our data suggest that the defectively folded P474T C-RRM may be stabilized upon incorporation into the di-snRNP complex, allowing formation of stable di-snRNPs, thus potentially explaining the relatively mild pathological phenotype observed with the patients.

RESULTS

IGHD-causing mutations in the 65K protein lead to impaired RNA binding

To address the impact of IGHD-causing mutations within the 65K protein on binding of the U12 or U6atac snRNAs in vitro, we carried out RNA-binding assays using recombinant wild-type (WT) or mutant (P474T or R502X) 65K C-RRM and ^{32}P-labeled short RNA hairpin substrates described earlier (Fig. 1A; Benecke et al. 2005; Singh et al. 2016). As the untagged P474T-mutant RRM is stable in soluble form under RNA-binding conditions for only a short time after the tag cleavage, the in vitro binding assays were carried out using tagged proteins, while untagged proteins were used for all other experiments (Supplemental Fig. S1). Short RNA hairpins corresponding to U12 snRNA nucleotides 109–125 and U6atac nucleotides 92–108 were used as ligands and a hairpin with a sequence complementary to the U12 hairpin was used as a control for nonspecific binding (Fig. 1A).
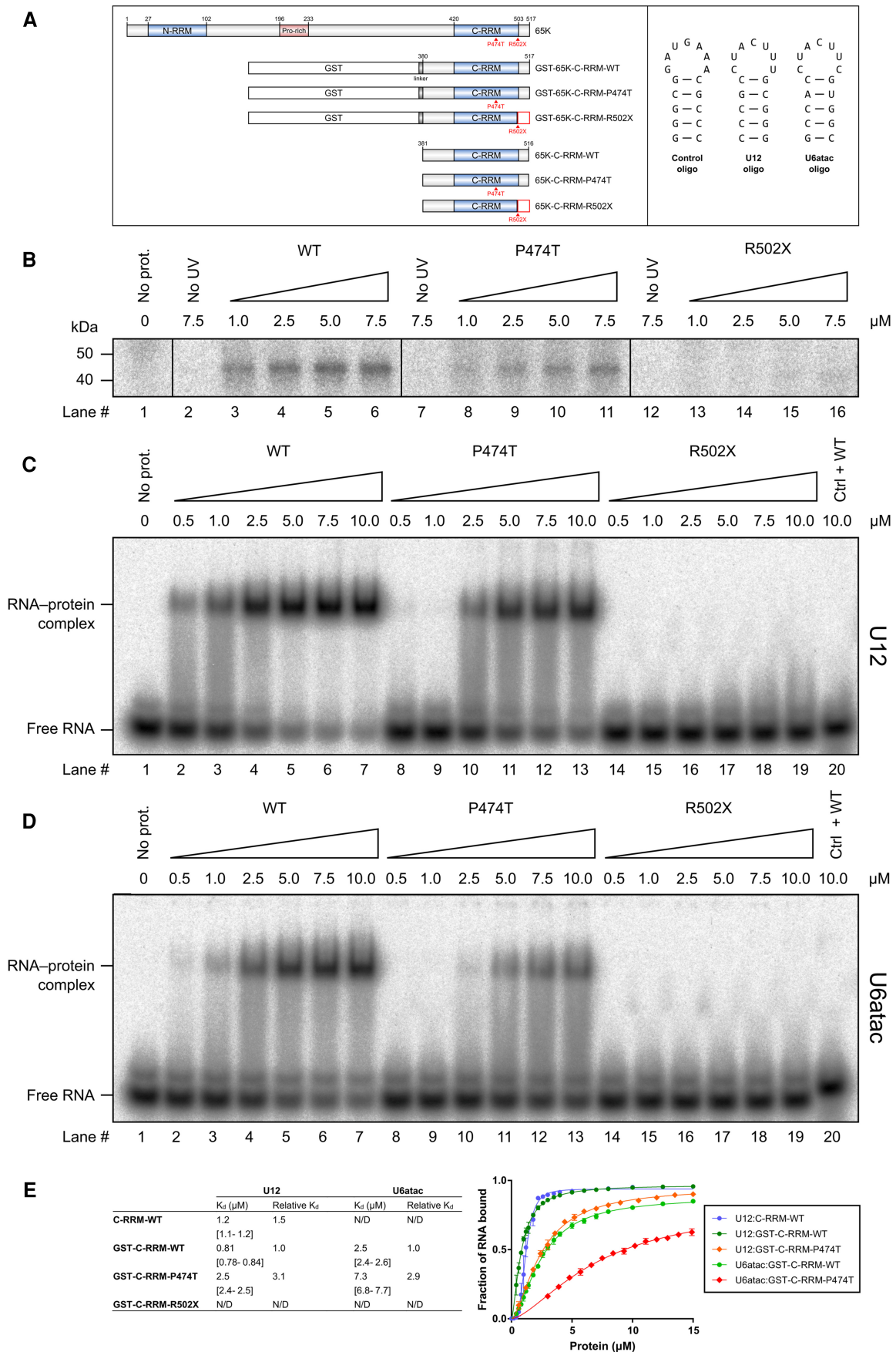
FIGURE 1. (Legend on next page)

RNA and protein samples were incubated in conditions described earlier (Benecke et al. 2005). The resulting RNA–protein complexes were either crosslinked using UV light, digested with RNase A and run on an SDS-polyacrylamide gel (UV crosslinking assay) or resolved from free RNA on a native polyacrylamide gel (EMSA).

UV treatment of increasing amounts of wild-type C-RRM in the presence of the labeled U12 oligonucleotide resulted in the appearance of a band of ~45 kDa that was not detected in the absence of protein (Fig. 1B, lane 1) or without UV treatment (lane 2). Compared to the wild-type protein, the P474T mutant showed reduced crosslink formation (lanes 3 −6 and 8−11), while no crosslinking was observed with the R502X mutant in the protein concentration range used in the assay (lanes 13−16). Next, we carried out electrophoretic mobility shift assays to quantitatively analyze the effect of the mutations on RNA binding (Fig. 1C–E; Supplemental Fig. S2). First, we tested whether the GST-tagged wild-type RRM binds RNA with similar affinity as the untagged protein and found the K_Ds to be similar (Supplemental Fig. S2A). Compared to the wild-type protein ($K_D = 0.8$ μM), the P474T mutant showed approximately threefold reduced binding to the U12 oligo ($K_D = 3.1$ μM), while the R502X mutation led to essentially complete loss of binding (Fig. 1C,E). Even when significantly higher protein concentrations were used, no specific complex formation could be observed for R502X (Supplemental Fig. S2B) under the experimental conditions.

Next, we tested the effect of 65K mutations on U6atac snRNA binding. EMSA with an U6atac-specific hairpin revealed approximately threefold weaker affinity to the WT C-RRM (Fig. 1D,E). Importantly, the relative difference between WT and P474T in binding to U12 and U6atac hairpins remained the same, as we found that the P474T mutation leads to approximately threefold reduction in U6atac binding [K_D (WT) = 2.5 μM, K_D (P474T) = 7.3 μM; Fig. 1D,E], similarly as observed with the U12 stem–loop. Furthermore, no binding of the U6atac stem–loop was observed for the R502X mutant (Fig. 1D; Supplemental Fig. S2B).

Mutated 65K proteins can bind to U11 snRNP in vivo

To ask if the mutated 65K proteins can be incorporated into the U11/U12 di-snRNP or the individual U11 or U12 snRNPs, we next expressed V5-tagged full-length 65K proteins (either WT protein or P474T or R502X mutant) in HeLa cells. After cell lysis and immunoprecipitation (IP) with anti-V5 antibody, we analyzed the association of the 65K protein with U11 and U12 snRNAs by northern blotting. We found that the WT V5-65K protein showed a robust co-IP of both U11 and U12 snRNAs, suggesting that the V5-tagged protein can be incorporated to the U11/U12 di-snRNP (Fig. 2A, lane 3). In contrast, both mutants showed a clear reduction in the levels of co-IP of both U11 and U12 snRNAs (Fig. 2A, lanes 4–5 and Fig. 2B). The difference between the two mutants was that the P474T mutation affects the association of both U11 and U12 snRNA equally, while with the R502X mutation the co-IP of U12 snRNA is very inefficient. Interestingly, comparison to western analysis of the V5-tagged proteins revealed a correlation between co-IP and 65K levels (Fig. 2B), while the mRNA levels of the V5-tagged proteins (as analyzed by RT-PCR) were equal between the three different constructs (Fig. 2A).

Together the results suggest that in the context of full-length 65K protein both mutations may reduce the stability of the protein. However, the correlation between U11 snRNA co-IP efficiency and V5-tagged protein levels suggests that the mutated proteins may be stabilized upon binding to U11 snRNP, which presumably takes place via the N-terminal RRM interacting with the U11-59K protein as previously described (Benecke et al. 2005). In contrast, the efficiency of U12 snRNA co-IP is consistent with the effect of the mutations on U12 binding in vitro (Fig. 1), and suggests that the two mutants have differential effects on U11/U12 di-snRNP stability. Destabilization, particularly by the R502X mutation, could potentially lead to a dominant negative effect that could decrease the efficiency of the splicing of U12-type introns. However, we failed to detect any increased minor intron retention (Supplemental Fig. S3A) or cryptic splice site activation (Supplemental Fig. S3B) upon expression of either V5-tagged mutated protein, while the splicing defects were readily observed after knockdown of specific U11/U12 di-snRNP components (Supplemental Fig. S3B, cf. lanes 2–4 and 7–9).

The R502X mutation leads to isoform-specific nonsense-mediated decay

The position of the 1504C>T (p. R502X) mutation ~50 bp upstream of the stop codon in the penultimate exon

FIGURE 1. P474T and R502X mutations in 65K lead to impaired U12 and U6atac binding. (*A*) Protein constructs and RNA oligonucleotides used in this study. (*Left*) Domain structures of the full-length 65K protein and the GST-tagged and untagged protein constructs are shown. Locations of the P474T and R502X mutations are also indicated. (*Right*) Secondary structures of the control, U12 and U6atac oligonucleotides used in this study. (*B*) Ultraviolet crosslinking of wild-type and mutant 65K C-RRM to a U12 oligonucleotide. Indicated concentrations of GST-tagged protein constructs (WT, P474T or R502X) shown in *A*, were incubated with a ^{32}P-labeled U12 oligo, crosslinked with UV, digested with RNase A and run on a SDS-polyacrylamide gel. "No UV" control samples were otherwise treated identically, but UV treatment was omitted. (*C*) EMSA analysis of the interaction of wild-type and mutant 65K C-RRMs with the U12 oligonucleotide. RNA–protein complexes were formed by incubating the labeled U12 oligonucleotide with varying concentrations of the indicated GST-tagged protein. Subsequently, complexes were resolved from the free unbound RNA on a native polyacrylamide gel. (*D*) EMSA analysis of the interaction of wild-type and mutant 65K C-RRM with U6atac. The experiment was carried out as in *C*, but using a U6atac oligonucleotide. (*E*) Dissociation constants and binding curves determined from EMSA experiments. Dissociation constants were determined by quantification of EMSA gels and nonlinear regression using data from three replicate EMSA experiments. Values in brackets indicate 95% confidence intervals.
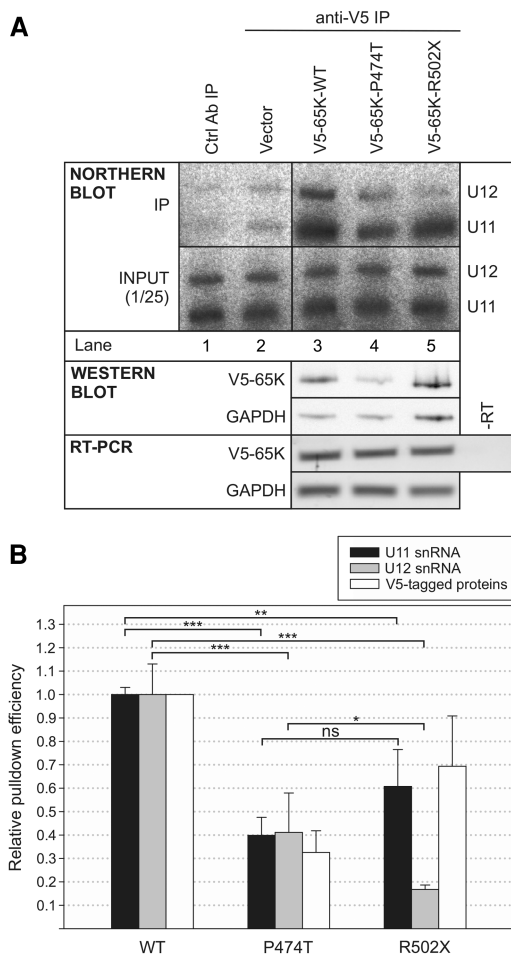
FIGURE 2. 65K mutants form di-snRNPs with reduced efficiency. (A) HeLa cells were transfected with plasmids for expressing V5-tagged 65K protein (WT, P474T, or R502X) or empty vector. After cell lysis, immunoprecipitation was carried out using either anti-V5 antibody or anti-FLAG antibody as a control for nonspecific binding and co-immunoprecipitation of U11 and U12 snRNAs was analyzed by northern blot. Protein and mRNA level measurements of exogenous 65K by western blotting and RT-PCR, respectively, were carried out from experiments run parallel with the co-IP samples. (B) Quantification of the northern (co-IP) and western blot data from A. U11 and U12 co-IP data were derived from three independent experiments which were normalized within each experiment by setting the mean U11 + U12 value of the WT co-IP reaction to one. V5-tagged protein level measurements were similarly derived from three independent experiments and normalized first using GAPDH signal, then by setting the WT value to one. Significance levels are indicated (ns: $P > 0.05$, [*] $P \leq 0.05$, [**] $P \leq 0.01$, [***] $P \leq 0.001$).

(Fig. 3A) suggests that this allele may be targeted by the cytoplasmic nonsense-mediated decay (NMD) pathway. To ask if this is indeed the case, we treated lymphoblastoid cell lines derived from two control individuals and two IGHD patients (Argente et al. 2014) either with translation blocker cycloheximide (CHX) that is expected to stabilize NMD targets, or with DMSO as a control. However, the downstream analysis is complicated by the crossregulatory pathway targeting 65K mRNA that leads to the formation of two mRNA isoforms

with different 3′UTR sequences (Verbeeren et al. 2010). Of these, only the transcript containing short 3′UTR (Fig. 2A) is cytoplasmic and accessible to translation machinery, while the transcript with long 3′UTR is retained in the nucleus (Verbeeren et al. 2017). Therefore, it was necessary to analyze the effect of the CHX treatment for the two isoforms separately, first using RT-PCR primers specific for the short-3′UTR and long-3′UTR isoforms (Fig. 3A) and then subsequently analyzing the PCR products with mismatch cleavage (Surveyor) assay (Fig. 3B) and Sanger sequencing (Fig. 3C). The cleavage pattern of the annealed PCR products in the Surveyor assay not only confirms the presence of two mismatch mutations in IGHD patient cells (Fig. 3B), but also provides evidence that one of the transcripts is targeted by the NMD machinery. Specifically, the CHX-treated short isoform reactions (Fig. 3B, lanes 5–8) show an increase of the cleaved products, suggesting that one of the transcripts is being stabilized due to block in translation. Consistently, direct sequencing of the PCR products shows an underrepresentation of the WT allele in the position of the P474T mutation (Fig. 3C, panel C1) and a similar underrepresentation of the nonsense allele in the position of the R502X mutation (Fig. 3C, panel C4). After CHX treatment, the two alleles have near-equal frequency (panels C5 and C7, respectively). In contrast, with transcripts containing long 3′UTR, no change was detected upon CHX treatment using Surveyor assay (Fig. 3B, lanes 13–16). Consistently, sequencing of the PCR products revealed that the two alleles were equally represented in the presence or absence of CHX (Fig. 3C, panels C2, C4, C6, C8). Therefore, we conclude that the nonsense 1504C>T mutation leads to isoform-specific NMD in the patient cells and affects only the cytoplasmic short-3′UTR isoform, but not the nuclear long-3′UTR isoform.

Mapping the RNA–protein contacts of the 65K C-terminal RRM using NMR

The crystal structure of the 65K C-RRM has been reported earlier (Netter et al. 2009). This study revealed that the RRM adopts a classic RRM fold with a βαββαβ topology, but with an additional N-terminal extension that stabilizes the domain. To understand the structural basis of the reduced binding of the U12 snRNA on the P474T mutant, we first solved the solution structure of the wild-type C-RRM domain (381–516), and analyzed the RNA–protein interactions by NMR spectroscopy. The NMR structure of the wild-type C-RRM revealed a canonical RRM fold with a flexible C-terminal region (Fig. 4A,B). The globular domain of C-RRM is similar to the published crystal structure (PDB ID: 3EGN) with backbone global displacements below 1 Å for the majority of the secondary structures, but shows some differences in the loop regions and around helices αII and 3$_{10}$II. The largest difference was observed for the loop between β2 and β3 strands, where two backbone nitrogen assignments are missing (Fig. 4C). The ^{15}N backbone
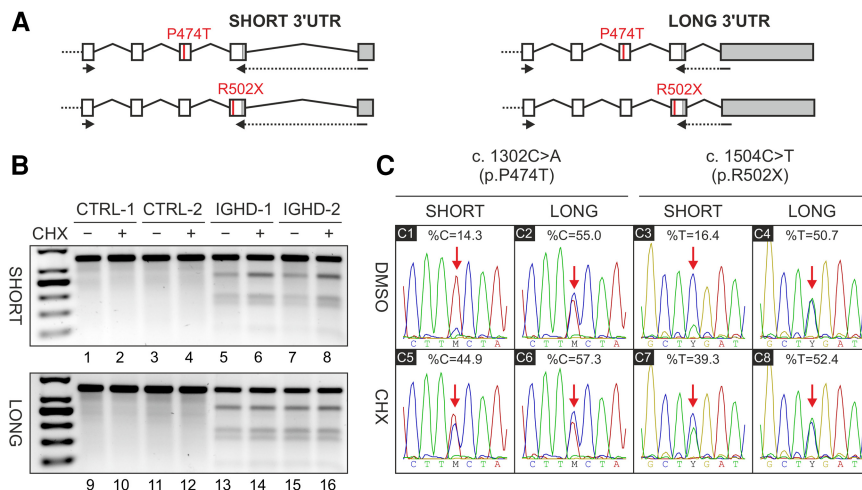
FIGURE 3. Nonsense *65K* transcripts with short 3′UTR are targeted to the NMD pathway. (*A*) Schematic presentation of the 65K long-3′UTR and short-3′UTR isoforms, with the location of the P474T and R502X mutations and the primers used in 2B and 2C indicated. (*B*) Surveyor nuclease digestion of short-3′UTR and long-3′UTR RT-PCR products from DMSO- or CHX-treated lymphoblastoid cell lines. RT-PCR was carried out using primers shown in Figure 2A, followed by reannealing of the PCR products and digestion with the mismatch-specific Surveyor nuclease. (*C*) Sanger sequencing of the short-3′UTR and long-3′UTR RT-PCR products from a patient lymphoblastoid cell line. Arrows indicate the position of the c. 1302C>A and c. 1504C>T mutations.

relaxation measurement confirms the flexibility of the C-terminal region as indicated by longer T_2 relaxation times and smaller heteronuclear NOEs (Supplemental Fig. S4). The largest difference between the crystal and NMR structures was observed at the C terminus. Whereas the crystal structure was modeled until residue 506, the C-terminal flexible region of the NMR structures starts around residue 502 as suggested by the ^{15}N relaxation data.

In the presence of the U12 RNA oligo (1:1 RNA:protein molar ratio), 15 residues showed significant chemical shift perturbation (Fig. 5A,B; Supplemental Table 1). In addition, the signals were not detectable or traceable for 16 residues in the [^{1}H, ^{15}N]-HSQC spectrum of 65K C-RRM in the presence of RNA, presumably due to the exchange broadening. Consistent with a mode of RNA binding characteristic of canonical RRMs, many of the residues on the RRM's β-sheet surface (β1–β4) responded to RNA binding, including most of the residues belonging to the highly conserved RNP2 and RNP1 motifs on β1 and β2, respectively. This was also in line with the positively charged surface of the NMR structure, which is likely to interact with the negative surface of the RNA (Fig. 5D). Furthermore, an earlier study reported that Y423A and Y423F mutations in RNP2 and a K466Q mutation in RNP1 completely abolish the binding of 65K C-RRM to RNA (Netter et al. 2009). Consistently, these residues show either significant change in chemical shifts (Fig. 5A,B: 423Y) or disappear in the presence of RNA (Fig. 5B: 466K) in our NMR experiment. Changes were also seen in the noncanonical N-terminal extension (Fig. 5B, residues 388–411), which provides stabilization and is required for enhancing

the RNA binding affinity (Netter et al. 2009). Interestingly, residues 502R and 504A, both of which are deleted by the R502X mutation, were among the residues showing notable chemical shift changes (Fig. 5B,C). As residues 507–517 are disposable for RNA binding (Netter et al. 2009), the critical residue(s) accounting for the complete loss of RNA binding seen with the R502X mutant are expected to lie within the range 502–506.

Structural effects of the P474T mutation

In the structure of the 65K C-RRM, proline 474 is located in a turn connecting β3 and αIV (Fig. 4A). Based on the position of this residue, we predicted that the P474T mutation might affect the folding of the RRM domain. Indeed, the HSQC spectrum of the P474T mutant indicates that the well-folded globular structure of the wild-type protein is no longer present, with severe line-broadening of signals, presumably due to slow conformational exchange between the partially folded and unfolded states (Fig. 6A). The mutant also slowly precipitated after cleavage of the tandem fusion protein of maltose binding protein and SUMO domain (Guerrero et al. 2015). In contrast, the HSQC spectrum of the R502X mutant revealed that this mutant retains the well-folded structure of the wild-type protein (Fig. 6B).

To confirm the NMR result with the P474T mutant, we first used differential scanning fluorimetry (DSF) to analyze the thermal stability of the untagged WT and P474T RRM in the presence and absence of RNA ligand (Fig. 6C). In the absence of RNA, we measured a melting temperature of 37°C for the wild-type RRM. When the protein was preincubated with a 10-fold molar excess of the U12 RNA oligo, the melting temperature shifted to 67°C, showing that RNA binding leads to significant thermal stabilization. Compared to WT, the melting curve of the P474T mutant assumed a very different shape, showing high fluorescence already at low temperatures and essentially lacking a clear unfolding transition, supporting the molten globule-like state of the mutant. Little change was observed upon addition of RNA ligand. The behavior of the P474T mutant in the assay suggests that this mutation has a destabilizing effect on the fold of the domain. However, the finding that P474T is still able to bind RNA, albeit with lower affinity compared to WT, suggests that the mutant retains crucial elements needed for RNA binding.

To more carefully investigate the structural effects of the P474T mutation, we recorded the NMR spectrum of the P474T mutant with and without RNA ligand (Fig. 6D). The severely broadened NMR signals of the P474T mutant
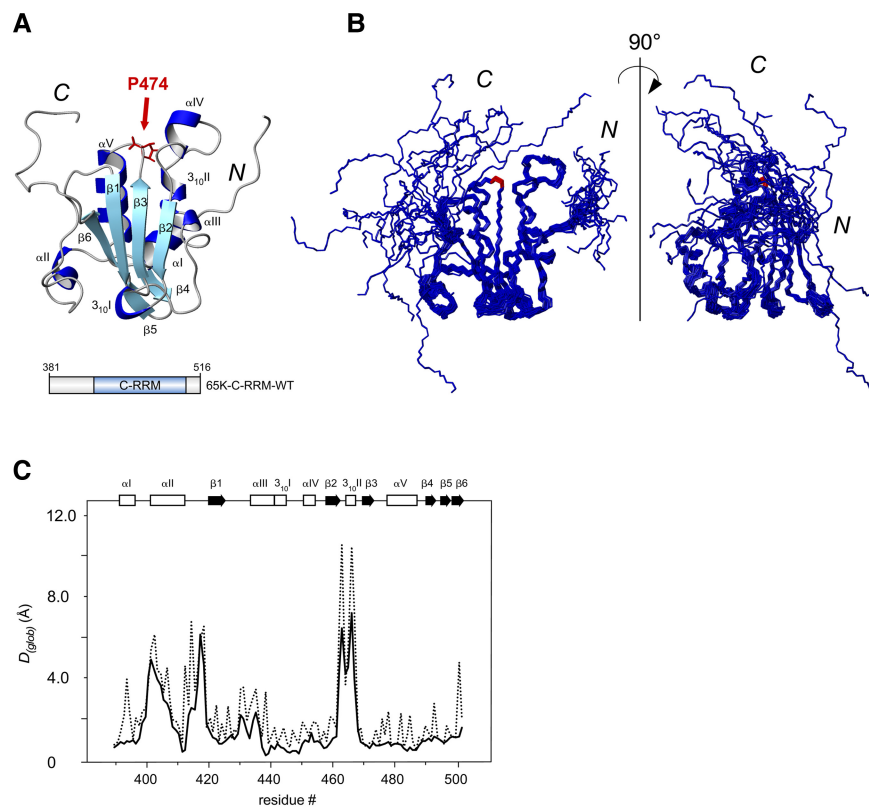
FIGURE 4. NMR structure of wild-type 65K C-RRM domain. (*A*) The solution NMR structure of wild-type C-RRM. Secondary structure elements are indicated in the figure. The side-chain of proline 474 is displayed in red by stick model. Residues within the C-RRM core region are shown. (*B*) Superposition of the 20 energy-minimized CYANA conformers representing the wild-type 65K C-RRM domain in solution. The locations of the N and C termini are indicated. The location of proline 474 is indicated in red. Structures are visualized by MOLMOL. (*C*) Plot of the global displacement ($D_{(glob)}$) per residue of backbone heavy atoms (continuous line) and all heavy atoms (dashed line) of the wild-type 65K C-RRM crystal structure versus the mean NMR structure. The backbone heavy atoms of the structured region (388–501) were initially superimposed for minimal RMSD. Secondary structure is shown *above* the plot.

due to slow conformational exchange caused some shifts of the peaks in the spectrum upon addition of the RNA ligand. This observation suggests that the P474T mutant retains some capability to bind RNA, which presumably leads to a shift in the conformational equilibrium. However, the structure of the P474T mutant seems to be still a mostly molten globule-like structure because of the severe line-broadening of the signals. Unfortunately, it was not feasible to obtain residue specific assignments of the RNA binding sites due to the low concentration and line-broadening.

DISCUSSION

The *RNPC3* gene codes for the U11/U12-65K protein that connects the individual U11 and U12 snRNPs to form the U11/U12 di-snRNP necessary for the minor intron recognition (Frilander and Steitz 1999; Benecke et al. 2005). Earlier work with *RNPC3* has described the compound heterozygous mutations P474T and R502X that are associated with

defects in minor intron splicing, pituitary hypoplasia and IGHD (Argente et al. 2014). In this work, we have carried out a detailed mechanistic investigation of the mutations, showing that they both affect the integrity of the U11/U12 di-snRNP by destabilizing *RNPC3* either at mRNA or protein level. We find that the *RNPC3* mRNA carrying R502X mutation is primarily targeted by the NMD machinery in the patient cells, and is expected to reduce the overall cellular levels of the U11/U12-65K protein (Argente et al. 2014). In contrast, the P474T mutation leads to a partial defect in the folding of the C-RRM, a subsequent threefold reduction in binding to both U12 and U6atac snRNAs, reduced formation of the U11/U12 di-snRNP complexes and reduced stability of the full-length protein. Together our results suggest that the combined effect of the reduced levels of the U11/U12-65K protein and a partial loss of the binding to the U12 and possibly U6atac snRNAs both contribute to the pathology of IGHD (Fig. 7). Furthermore, our data suggest that in compound heterozygote patients, only the U11/U12-65K protein with P474T mutation can support functional interactions between the C-RRM and the U12 or U6atac snRNAs.

Both the NMR and DSF analyses indicate that the P474T mutation leads to a partial destabilization of the C-RRM and a molten globule-like state of this domain, which then presumably reduces affinity to U12 and U6atac snRNAs. The effect on folding is consistent with the location of the phylogenetically conserved P474 residue in a turn position connecting β3 and αIV elements of the C-RRM (Fig. 4A), which is known to be sensitive to such mutations (Betts and Russell 2003; Argente et al. 2014). In contrast, NMR data indicates that the R502X C-RRM is correctly folded, suggesting that the loss of snRNA binding in our in vitro experiments is a result of a loss of a critical interaction between snRNAs and the C-terminal residues of the protein.

In the context of full-length protein, our data suggest that the folding defect associated with the P474T mutation is reversible and restricted to the C-terminal part of the protein. Specifically, comparison of the co-IP efficiencies with V5-tagged WT, P474T and R502X 65K proteins revealed that both mutations lead to reduced overall association with U11 and U12 snRNPs. Significantly, with the P474T mutation the U11/U12 ratio was equal to that of the WT protein, while the R502X showed substantially reduced U12 snRNP
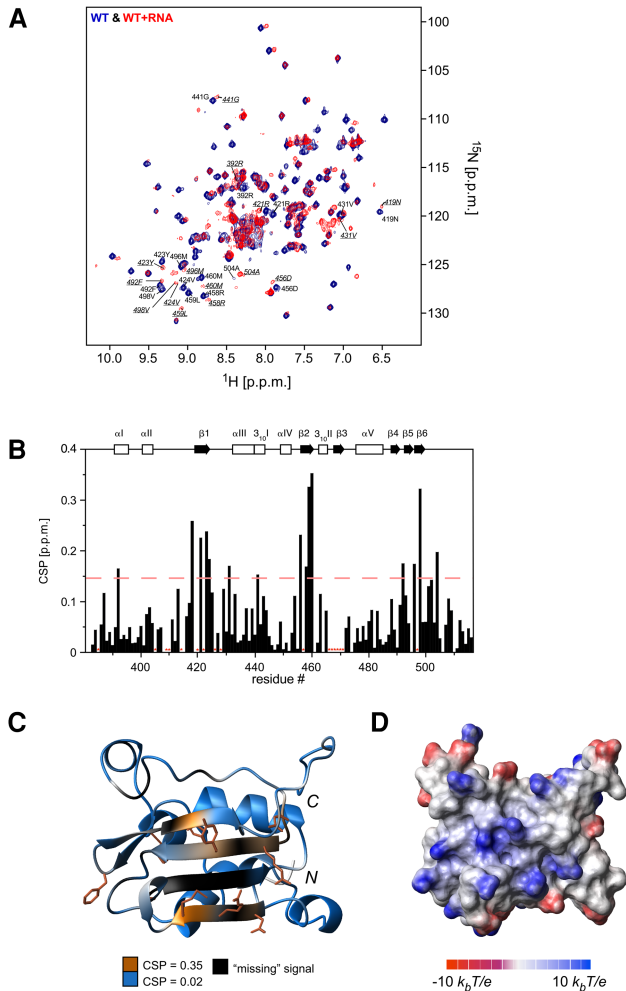
FIGURE 5. Interactions between wild-type 65K C-RRM domain and U12 RNA. (*A*) An overlay of the [¹H, ¹⁵N]-HSQC spectra of wild-type 65K C-RRM in the absence (blue) and presence of U12 RNA (red). The residues with significant chemical shift perturbations (>0.15) are labeled in both spectra. (*B*) A plot of chemical shift perturbation versus the residue number. The dashed line represents the threshold used as the criteria for significant chemical shift perturbations. The residues that could not be unambiguously identified in the [¹H, ¹⁵N]-HSQC spectrum of the C-RRM/RNA complex are indicated with asterisks in red. The regions of the secondary structures are indicated *above* the plot. (*C*) Chemical shift perturbations are mapped on the wild-type 65K C-RRM structure, where blue color indicates smallest and orange biggest CSP value. Locations with the missing signals in [¹H, ¹⁵N]-spectrum of the wild-type 65K C-RRM/RNA complex are colored black. Prolines and other unassigned residues are colored in white. Side-chains of the residues with larger chemical shift perturbations (>0.15) are displayed by stick models. (*D*) Surface electrostatic potential of wild-type 65K C-RRM is shown as a colored gradient between $-10 \, k_bT/e$ and $10 \, k_bT/e$, where k_b is Boltzmann's constant, T temperature, and e elementary charge. The structures are visualized by MOLMOL.

co-IP levels. Together these results suggests that neither one of the two mutations interferes with the interactions between 65K N-terminal RRM and the U11-59K protein, but only the P474T supports stable interactions with the U12 snRNA. Additionally, the identical U11/U12 ratio observed between

WT and P474T mutant co-IP experiments (Fig. 2B) further suggests that a successful incorporation of the P474T mutant within the di-snRNP leads to stabilization of the C-RRM, possibly as a consequence of U12 snRNA binding to the P474T C-RRM. This interpretation is consistent with the observations of improved folding of the P474T C-RRM (Fig. 6D) and the stabilization of the WT C-RRM (Fig. 6C) in the presence of U12 RNA hairpin.

Using patient cell lines, we further demonstrate that the R502X mutation leads to isoform-specific NMD of the *RNPC3* mRNA. Specifically, due to the underlying crossregulation pathway (Verbeeren et al. 2010, 2017), two *RNPC3* mRNA isoforms are produced, of which the long-3′UTR isoform is retained in the nucleus and thus escapes NMD; however, the short-3′UTR isoform is efficiently exported to the cytoplasm and is accessible to the NMD machinery. As a consequence of NMD, the translational output of the R502X allele is likely to be low or negligible. Given that the P474T mutation apparently also leads to a faster turnover of the 65K protein (Fig. 2A), it is possible that the 65K protein with P474T mutation is stable only within the U11/U12 di-snRNP. Therefore, the disproportionately strong reduction in the 65K protein levels in addition to the reduced U11/U12 di-snRNP levels observed earlier in the patient cells (Argente et al. 2014), could be explained through a concerted effect of both mutations in reducing the cellular 65K protein levels.

Even though the primary effect of the R502X mutation is mRNA decay by the NMD machinery, this mutation can potentially also delete the C-terminal tail of the 65K protein in hypothetical cases where subsets of the transcripts escape the decay pathway. Interestingly, our in vitro binding experiments with the R502X mutant provided the somewhat unexpected result that the loss of the C-terminal tail leads to complete loss of binding between the C-RRM and U12 or U6atac snRNAs. The position 502 is located at the beginning of the flexible C-terminal tail of the domain, which is disordered beyond residue 502 and 506 in the NMR and crystal structures, respectively. Our NMR data on R502X C-RRM show that this mutation does not have major structural consequences. Rather, our results suggest that residues in the C-terminal tail provide additional crucial contacts for U12 and U6atac snRNA binding. Consistently, in our NMR experiment with the WT C-RRM and U12 snRNA stem–loop, we observed chemical shift changes in residues 502 and 504 (Fig. 5A,B) both of which are evolutionarily conserved (Argente et al. 2014), suggesting that they are likely candidates contributing to the recognition of the snRNAs.

The model described above does not account for the potential role of the disrupted 65K–U6atac interaction in IGHD. Consistent with a recent report documenting the interaction between 65K C-RRM and the distal stem–loop of U6atac snRNA (Singh et al. 2016), our biochemical analysis confirmed that the 65K C-RRM can indeed bind to U6atac, but with approximately threefold reduced affinity compared
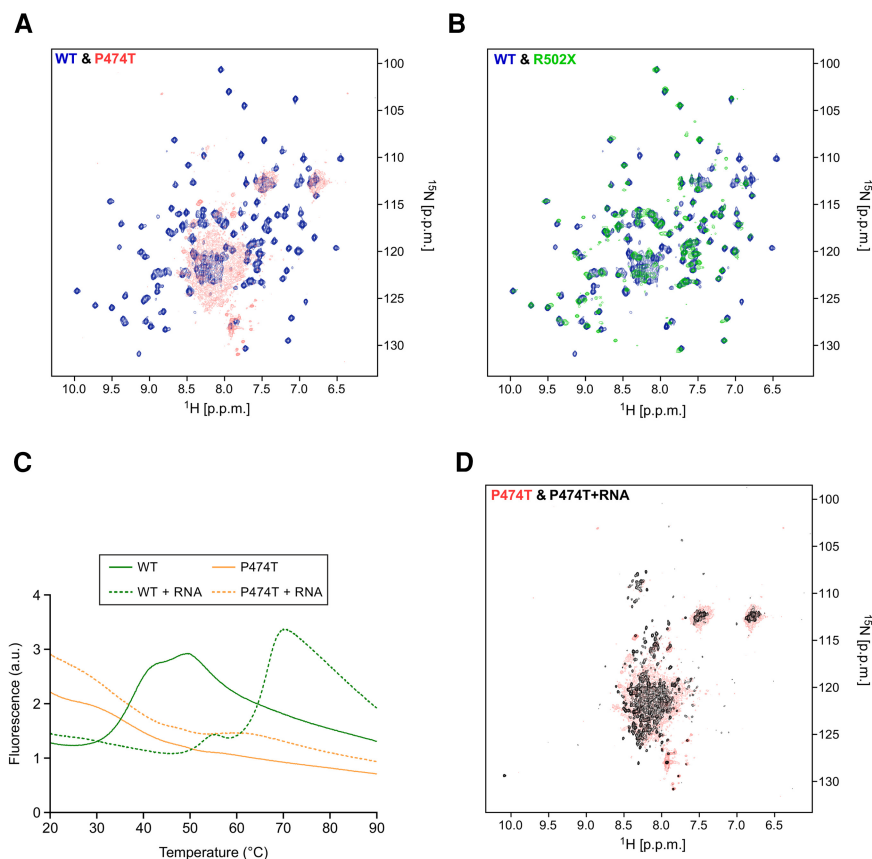
FIGURE 6. Stability and folding properties of 65K C-RRM and its P474T and 502X variants. (A) An overlay of the [^{1}H, ^{15}N]-HSQC spectra of the wild-type 65K C-RRM (blue) and P474T 65K C-RRM (pink). (B) An overlay of the [^{1}H, ^{15}N]-HSQC spectra of the wild-type 65K C-RRM (blue) and R502X 65K C-RRM (green). (C) Differential scanning fluorimetry analysis of wild-type and P474T-mutant RRM domain. Melting curves of untagged wild-type and P474T-mutant RRM were obtained in the absence or presence of U12 RNA oligonucleotide by heating the proteins or protein–RNA complexes in the presence of SYPRO Orange and recording the fluorescence of the dye as a function of temperature. (D) An overlay of the [^{1}H, ^{15}N]-HSQC spectra of 65K C-RRM (P474T) in the absence (pink) and presence of RNA (black).

to U12 snRNA. Furthermore, we documented a similar additional threefold reduction in the U6atac snRNA binding with the P474T mutant and a complete loss of binding with R502X. Currently, little is known about the functional relevance of the 65K–U6atac interaction. The 3′ end of U6atac, which includes the 65K-binding stem–loop, has been demonstrated to play a role in guiding U6atac snRNA to the minor spliceosome (Dietrich et al. 2009). Although U6atac snRNA can be pulled down with 65K from HeLa cell lysate (Singh et al. 2016), it is not known if the 65K protein associates with the U4atac/U6atac.U5 tri-snRNP or whether, in addition to the A complex, it is present in later spliceosomal complexes. Interestingly, analysis of a zebrafish 65K mutant revealed an accumulation of a slow-migrating complex containing U12, U5, and U6atac snRNAs, leading the authors to suggest an additional role for 65K in the later stages of the splicing process, such as spliceosome disassembly or recycling (Markmiller et al. 2014). Whether such a

function would be conserved in human cells remains to be determined. The 65K-binding U6atac stem–loop appears to be conserved between zebrafish and human; however, it is unknown if the observed accumulation of a late-stage splicing complex is related to the 65K–U6atac interaction. Nevertheless, it is possible that IGHD mutations may have an additional effect during the later stages of splicing in addition to intron recognition. A hypothetical exchange in interactions from U12 to U6atac snRNA could expose the P474T C-RRM to an additional round of unfolding and folding, which could in turn have downstream consequences in the kinetics of spliceosome assembly and/or the stability of the protein.

Ultimately, the reduced formation and/or stability of the di-snRNP leads to defective recognition of U12-type introns, resulting in defects in U12-type intron splicing. Indeed, RNA-seq analysis of patient blood cells revealed widespread splicing defects (Argente et al. 2014). Significantly, in addition to the expected intron retention events, a substantial number of cryptic splice site activation in genes containing U12-type introns were detected. This suggests that impaired recognition of U12-type introns is one of the outcomes of the IGHD mutations. However, it does not rule out that later steps of the spliceosome assembly process, due to defects in the putative 65K–U6atac interaction, could also be affected.

Similar molecular defects as those reported here are thought to underlie another congenital human disease, MOPD1, which is caused by mutations in U4atac snRNA (Edery et al. 2011; He et al. 2011). Akin to the impairment of the U11/U12-65K–RNA interaction in IGHD, many of the disease-causing alleles of U4atac snRNA were shown to disrupt a crucial RNA–protein interaction between the tri-snRNP specific 15.5K protein (Snu13) and the 5′ stem–loop of U4atac snRNA (Jafarifar et al. 2014). This binding defect and its downstream effects on the binding of other proteins lead to a further defect in the association of U5 snRNA to the U4atac/U6atac di-snRNP and ultimately the formation of the U4atac/U6atac.U5 tri-snRNP complex. Thus, both diseases compromise the integrity of a minor spliceosome-specific snRNP complex, either U11/U12 di-snRNP or U4atac/U6atac.U5 tri-snRNP, as a consequence of an impaired RNA–protein interaction. The molecular level difference between the two diseases is that, whereas IGHD
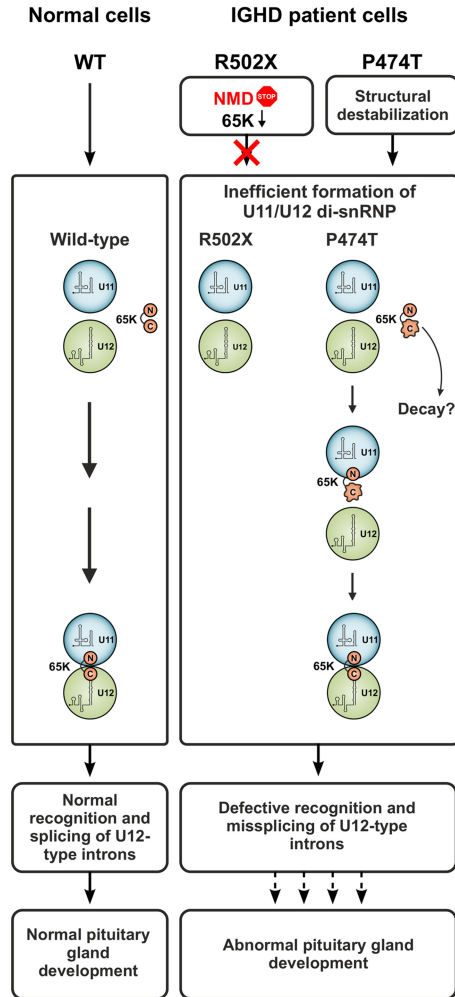
developing pituitary gland. Understanding the precise pathways involved in both diseases would require detailed investigations using animal models.

MATERIALS AND METHODS

Protein expression and purification of GST-tagged proteins

GST-tagged proteins used for in vitro binding assays were expressed in *E. coli* Rosetta (DE3)pLysS cells. After 3 h induction with 0.5 mM IPTG at 37°C, cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris–HCl, 150 mM NaCl, 5 mM DTT, pH 8.0 + 1× Roche cOmplete protease inhibitor cocktail) and lysed by sonication. After centrifugation to pellet cell debris, the GST-tagged fusion proteins were purified from the lysate using glutathione agarose (Thermo Scientific) according to the manufacturer's instructions. Purified protein was concentrated and buffer-exchanged to 20 mM HEPES–KOH, 100 mM KCl, 1.5 mM $MgCl_2$, 5% glycerol, pH 8.0.

Cloning, protein expression, and purification for NMR studies

The coding sequence amplified from plasmid pHAT2-65K-C-RRM using the two oligonucleotides, I693: 5′-TGAGGATCCGATGAAAT GCCTTCAGAATG and I710: 5′-GTGGTACCTTACTTTCTTTTT CCTTCCTTAG, was cloned between BamHI and KpnI of pHYRSF53 vector (Addgene #64696), resulting in pBHRSF109 bearing H_6-SMT3-C-$RRM_{381-516}$. P474T variant was constructed with the same oligonucleotide but using the plasmid containing the mutant as the template and cloned into pLJSRSF7 (Addgene #64693), resulting in pBHRSF118 coding the fusion protein of His_6-MBP-SUMO-P474 mutant (Guerrero et al. 2015). The R502X variant was produced from the plasmid pADHRSF12 bearing H_6-SMT3-502X, which was constructed using pHYRSF53, and two oligonucleotides, I693 and J475: 5′-TGCGGTACCTTAA GCAAACTGAACCACCATGGG. For NMR studies, doubly [^{13}C, ^{15}N]-labeled U11/U12-65K C-RRM (381–516) was expressed using *E. coli* strain ER2566 transformed with pBHRSF109 in the M9 medium, supplemented with 25 μg/mL kanamycin and $^{15}NH_4Cl$ and ^{13}C-glucose as the sole nitrogen and carbon sources. The protein was purified according to the protocol described in Guerrero et al. (2015). The purified protein was dialyzed against 20.0 mM sodium phosphate buffer (pH 6.0) and concentrated to 250 μL. d_{10}-DTT (1,4-ditrithreitol) (Sigma-Aldrich, CAS No. 302912-05-6) was added to the buffer to a final concentration of 2.0 mM. The 0.2 mM protein sample of doubly [^{13}C, ^{15}N]-labeled U11/U12-65K C-RRM (381–516) containing 5% D_2O (v/v) was transferred into a 5.0 mm microcell NMR tube (Shigemi Inc.). The ^{15}N-labeled U11/U12-65K C-$RRM_{(381–516)}$ P474T sample was expressed and purified as described above and concentrated to 250 μL (a final concentration of 30 μM). The fractionally [^{15}N, ^{13}C]-labeled U11/U12-65K C-$RRM_{(381–501)}$ 502X mutant sample was expressed and purified as described above except for the M9 medium that contained 20% ^{13}C-D-glucose and 80% natural isotope abundance D-glucose (w/w), and 100% $^{15}NH_4Cl$ as sole nitrogen and carbon sources, and concentrated to 250 μL with the final concentration of 88 μM.

FIGURE 7. A model for molecular defects in isolated growth hormone deficiency with 65K mutations. IGHD patients are compound heterozygous for two mutations, P474T and R502X, which either destabilize the *65K* mRNA through nonsense-mediated decay (R502X) or lead to misfolding and presumably increased decay of the 65K protein (P474T). Due to structural instability and the resulting U12 snRNA binding defect, the P474T mutant forms U11/U12 di-snRNPs with reduced efficiency compared to the WT protein. However, the mutant may be stabilized upon incorporation into the U11/U12 di-snRNP complex, allowing formation of stable di-snRNPs. Reduced level of the U11/U12 di-snRNP causes defects in intron recognition and U12-type intron missplicing. Through unknown mechanisms, this leads to impaired development of the pituitary gland.

results from defects in intron recognition, MOPD1 mutations impair the later stages of the minor spliceosome assembly. Presently, it is not known if this molecular level difference leads to the dramatic difference in severity between the two diseases: Whereas IGHD is a mild disease, mainly affecting pituitary development and postnatal growth, MOPD1 is usually a very severe disorder causing a number of developmental abnormalities and death at an early age, although milder cases have also been reported (Abdel-Salam et al. 2016; Krøigård et al. 2016). Presumably, the splicing defects in IGHD are largely tolerated in other tissues but not in the

Electrophoretic mobility shift assay

RNA–protein complexes were formed by incubating ~1000 cpm (~5 nM) of [γ-^{32}P]-ATP-labeled RNA oligo with different concentrations of GST-tagged 65K C-RRM (WT, P474T, or R502X) in a total volume of 10 µL. RNA oligonucleotides used in the assay (U12, U6atac, control) were bought from Sigma-Aldrich and are listed in Supplemental Table 2. Gel analysis of either untreated or heat-denaturated and snap-cooled oligonucleotides was used to confirm that hairpin oligonucleotides do not form dimeric structures. The binding buffer contained 20 mM HEPES–KOH (pH 8.0), 100 mM KCl, 1.5 mM MgCl$_2$, 5% glycerol, 0.1 µg/µL BSA, 1 µg/µL yeast RNA (Roche) and 1 U/µL RiboLock RNase inhibitor (Thermo Scientific). After 1 h incubation on ice, 2.5 µL of 5× loading buffer (20 mM HEPES–KOH, 100 mM KCl, 1.5 mM MgCl$_2$, 50% [v/v] glycerol, 0.1% [w/v] bromophenol blue, 0.1% [w/v] xylene cyanol) was added, and 6 µL of each reaction mixture was loaded onto a 6% native polyacrylamide gel with 0.5× TBE and 5% glycerol. Gels were run at 120 V for ~2 h at 4°C and visualized and quantified using the FLA-5100 phosphorimager and AIDA image analysis software (Raytest, Germany). Binding curves were fit to data from three independent experiments and dissociation constants determined by nonlinear regression (one-site specific binding with Hill slope) using GraphPad Prism 6.

Ultraviolet crosslinking assay

RNA–protein complexes were formed as described above for electrophoretic mobility shift assay, except that 10^5 cpm of RNA was used for each sample and BSA and yeast RNA were omitted from the binding buffer. After incubation, samples were transferred into the wells of a 96-well plate and UV-irradiated for 10 min on ice using the UVP CL-1000 Ultraviolet Crosslinker. For each protein, a "no UV" control was also prepared. After UV or control treatment, samples were digested with 20 µg of RNase A for 15 min at 37°C. Samples were then run on a 12% SDS–PAGE gel and the gel was visualized using FLA-5100 (Fuji).

Differential scanning fluorimetry

Untagged 65K C-RRM (WT or P474T mutant) was incubated on ice at a concentration of 5 µM in buffer 20 mM HEPES–KOH (pH 8.0), 100 mM KCl, 1.5 mM MgCl$_2$ for 1 h without RNA or in the presence of a 10-fold molar excess of U12 RNA oligo. After incubation, SYPRO Orange (Sigma) was added to 25× final concentration. The samples were then heated from 20°C to 95°C at a rate of ~1°C/min while measuring fluorescence from the SYPRO Orange dye using the Roche LightCycler 480 real-time PCR machine (excitation = 483 nm, emission = 568 nm). Protein melting temperature (T_m) was estimated as the temperature corresponding to the maximum value of the first derivative of fluorescence with respect to temperature.

Cell culture and transfection

HeLa cells were cultured in DMEM supplemented with 10% FBS, 1% penicillin–streptomycin and 2 mM L-glutamine. Transfections were carried out using the Lipofectamine 2000 reagent (Thermo Fisher Scientific). Lymphoblastoid cell lines derived from two con-

trol individuals and two IGHD patients (Argente et al. 2014) were cultured in RPMI 1640 with 10% FBS, 1% penicillin–streptomycin and 2 mM L-glutamine. For the nonsense-mediated decay experiments (Fig. 3), cells were treated with 100 µg/mL cycloheximide or an equal volume of DMSO for 5 h.

RNA extraction and RT-PCR

Total RNA was extracted from cells using the TRIzol reagent and treated with RQ1 RNase-free DNase (Promega) to remove potential genomic DNA contamination. cDNA synthesis was carried out using the Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific) and random primers. For PCRs, either the Phire Hot Start II or Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) was used. Primers used in RT-PCR are listed in Supplemental Table 2.

Immunoprecipitation and northern blot

HeLa cells were transfected in six-well plate format with 4 µg of pCI-neo plasmids expressing V5-tagged full-length 65K (wild-type, P474T or R502X) or empty vector. Twenty-four hours after transfection, cells were washed with PBS, scraped into ice-cold NP-40 lysis buffer (20 mM HEPES, pH 8.0, 137 mM NaCl, 10% glycerol, 1% NP-40, 2 mM EDTA, 1× cOmplete protease inhibitor cocktail [Roche], 1 U/µL RiboLock) and sonicated 5 × 30 sec using a Bioruptor Twin sonicator (Diagenode). After centrifugation (16,000g, 15 min) to remove cell debris, 2 µg of anti-V5 tag antibody (Invitrogen, R96025) or control antibody (anti-FLAG, Sigma F3165) bound to Dynabeads Protein G (Invitrogen) was added and incubated with the lysate for 1 h at 4°C with end-to-end rotation. After IP, beads were washed 3× with lysis buffer, followed by proteinase K digestion, phenol:chloroform extraction and ethanol precipitation. For input samples, RNA was extracted from 1/25 volume of the pre-IP lysate.

For northern blot, samples were run on an 8% urea–polyacrylamide gel and transferred onto Hybond-XL nylon membrane (GE Healthcare) using an Owl semi-dry blotter. Prehybridization, hybridization, and washes were done as described by Tarn and Steitz (1996b), except that LNA oligonucleotides U11-6L and U12-9L (Supplemental Table 2) were used as probes and hybridization was carried out at 45°C overnight. Northern blots were visualized and quantified using the FLA-5100 phosphorimager and AIDA image analysis software (Raytest, Germany).

Western blotting

Total protein (10 µg) was run on 10% SDS–PAGE gel and transferred onto Hybond-P membrane (GE Healthcare). Membranes were incubated with anti-V5 tag (Invitrogen, R96025) or anti-GAPDH (Cell Signaling 14C10) antibody, followed by HRP-conjugated secondary antibodies (NA931 and NA934, GE Healthcare). Chemiluminescence was detected using the LAS-3000 imager (Fuji). Blots were quantified using the Aida software (Raytest, Germany) and V5-65K signal was normalized to GAPDH signal.

Surveyor assay

65K short-3′UTR and long 3′UTR isoforms were amplified using isoform-specific exon-spanning primers (65K-long-R or 65K-short-R) and a common forward primer (65K-ex11-F; Supplemental Table 2) using cDNA from cycloheximide or DMSO-treated lymphoblastoid cell lines as a template. PCR products were column-purified, re-annealed to themselves to form homo- and heteroduplexes and digested with Surveyor nuclease according to the instructions from the Surveyor Mutation Detection Kit (IDT). Digestion products were analyzed by agarose gel electrophoresis. Sanger sequencing of the PCR products (Fig. 3C) was carried out using primer 65K-ex11-F.

NMR spectroscopy

All NMR experiments were carried out on Bruker 600 MHz or 850 MHz Avance III HD NMR spectrometers equipped with a cryogenically cooled TCI probe head. NMR spectra were recorded at 298.3K or 303.3K. For the sequential backbone assignment, the following NMR experiments were recorded: [^{1}H, ^{15}N]-HSQC, HNCA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, CON, CACO (Sattler et al. 1999; Kostic et al. 2002; Bermel et al. 2005). The aliphatic side-chain assignment was carried out using [^{1}H, ^{13}C]-HSQC, (H)CCCONH, H(CCO)NH, HBHACONH, HCCH-COSY, HCCH-TOCSY and ^{15}N-edited [^{1}H, ^{1}H]-NOESY spectra. The assignments for the aromatic side-chains were based on the spectra of CBCGCDHD, CBCGCDCEHE and ^{13}C-edited [^{1}H, ^{1}H]-NOESY. NMR spectra were processed using Bruker Topspin 3.2. The chemical shift assignment was carried out using CcpNmr Analysis 2.4.1 software (Vranken et al. 2005).

^{15}N NMR relaxation measurements

For backbone dynamics analysis, the longitudinal (T$_1$) and transverse (T$_2$) relaxation rates and heteronuclear ^{15}N{^{1}H}-NOEs for backbone ^{15}N atoms were determined at 298 K using the well-established NMR pulse sequences described previously (Kay et al. 1989; Barbato et al. 1992). T$_1$(^{15}N) and T$_2$(^{15}N) relaxation times were determined using the following delay times: 10, 50, 100, 200, 300, 500, 800, 1000, 1200, and 2000 msec for T$_1$ and 16, 64, 96, 128, 156, 196, 224, and 256 msec for CMPG pulse train with 1.0 msec interval for T$_2$ relaxation rates, respectively. Relaxation rates were obtained by fitting a single exponential decay to peak intensity values: $I(t) = I_0 \times \exp(-t/T_1)$ or $I_0 \times \exp(-t/T_2)$, where $I(t)$ is the peak volume at a time t. Heteronuclear ^{15}N{^{1}H}-NOEs were obtained with a relaxation delay of 5 sec with or without saturation of protons. Heteronuclear ^{15}N{^{1}H}-NOEs (η) were derived from the volumes of the HSQC peaks using the equation of $\eta = I/I_0$. The relaxation data were processed and analyzed using Bruker Dynamic Center software (Version 2.1.8).

NMR structure determination

Upper distance restraints were obtained from NOE intensities from three-dimensional ^{15}N- or ^{13}C-edited [^{1}H, ^{1}H]-NOESY spectra with a 75-msec mixing time for NOE building up. The NMR solution conformers were generated using CYANA 3.97

software based on the automated NOESY cross-peak assignments (Güntert et al. 1997; Güntert 2004, 2009) and dihedral angle constraints generated using TALOS-N software (Shen and Bax 2013). The restraint energy minimization of the final 20 best conformers was performed using AMBER 14 (Cornell et al. 1995). The structure validation was performed using PSVS 1.5. (Bhattacharya et al. 2007). The structures were visualized with MOLMOL (Koradi et al. 1996), which was also used for calculation of the electrostatic surface potential. The final structural statistics are shown in Table 1.

TABLE 1. Experimental data for the NMR structure calculation and the structural statistics for the 20 energy-minimized conformers of wild-type 65K C-RRM

Completeness of resonance assignments (%)[a]	
Backbone	98.8
Side chain	94.0
Aromatic	97.0
Distance restraints	
Total	2501
Sequential ($|i-j| \leq 1$)	1283
Medium range ($1 < |i-j| < 5$)	476
Long range ($|i-j| \geq 5$)	742
Dihedral angle restraints	
No. of restraints per residue	18.4
No. of long-range restraints per residue	5.5
Residual restraint violations	
Average no. of distance violations per structure	
0.1–0.2 Å	5.75
0.2–0.3 Å	0
>0.3 Å	0
Average no. of dihedral angle violations per structure	
>1°	0
Model quality[b]	
Rmsd backbone atoms (Å)	0.5
Rmsd heavy atoms (Å)	1.0
Rmsd bond lengths (Å)	0.014
Rmsd bond angles (°)	2.1
Ramachandran statistics[b]	
Molprobity	
Most favored regions (%)	94.8
Allowed regions (%)	5.0
Disallowed regions (%)	0.3
Global quality scores[b] (raw/Z score)	
Verify3D	0.41/−0.80
Prosall	0.80/0.62
PROCHECK (ϕ-ψ)	−0.41/−1.30
PROCHECK (all)	−0.43/−2.54
MolProbity clash score	1.73/1.23
Model contents	
Ordered residue ranges	388–501
Total no. of residues	136
BMRB accession number	34155
PDB ID code	5OBN

[a]Calculated excluding N-terminal, Lys, amino and Arg guanido groups, hydroxyls of Ser, Thr and Tyr, carboxyls of Asp and Glu, and nonprotonated aromatic carbons. Backbone includes HN, Hα, N, C′, Cα, and Cβ atoms.
[b]Calculated by using PSVS version 1.5 (Bhattacharya et al. 2007).

Protein–RNA interaction by chemical shift perturbation

For the RNA interaction studies, an equimolar protein–RNA (U12) solution was prepared and transferred into a siliconized 5 mm symmetrical microcell NMR tube (Shigemi Inc.). Unlabeled RNA U12 (CCCGCCUACUUUGCGGG) oligonucleotide was purchased from Sigma-Aldrich. [^{1}H, ^{15}N]-HSQC spectrum of the protein–RNA solution was recorded and overlaid with the spectrum of the U11/U12-65K C-RRM protein for the comparison. The chemical shift perturbations (CSPs) were calculated using the equation CSP (p.p.m.) $= [(\Delta^1H)^2 + 0.157 \times (\Delta^{15}N)^2]^{1/2}$ between a peak in the HSQC spectrum of the free C-RRM protein and the nearest shifted peak in the HSQC spectrum of the protein–RNA complex. CSPs were not calculated for the peaks that are apparently disappeared in the protein–RNA spectrum but marked.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

REFERENCES

Abdel-Salam GM, Emam BA, Khalil YM, Abdel-Hamid MS. 2016. Long-term survival in microcephalic osteodysplastic primordial dwarfism type I: evaluation of an 18-year-old male with g.55G>A homozygous mutation in *RNU4ATAC*. *Am J Med Genet* **170:** 277–282.

Argente J, Flores R, Gutiérrez-Arumí A, Verma B, Martos-Moreno GÁ, Cuscó I, Oghabian A, Chowen JA, Frilander MJ, Pérez-Jurado LA. 2014. Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol Med* **6:** 299–306.

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A. 1992. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31:** 5269–5278.

Benecke H, Lührmann R, Will CL. 2005. The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. *EMBO J* **24:** 3057–3069.

Bermel W, Bertini I, Duma L, Felli IC, Emsley L, Pierattelli R, Vasos PR. 2005. Complete assignment of heteronuclear protein resonances by protonless NMR spectroscopy. *Ange Chem Int Ed Engl* **44:** 3089–3092.

Betts MJ, Russell RB. 2003. Amino acid properties and consequences of substitutions. In *Bioinformatics for geneticists* (ed. Barnes MR, Gray IC), pp. 289–316. John Wiley & Sons, Ltd., Chichester, UK.

Bhattacharya A, Tejero R, Montelione GT. 2007. Evaluating protein structures determined by structural genomics consortia. *Proteins* **66:** 778–795.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* **117:** 5179–5197.

Dietrich RC, Padgett RA, Shukla GC. 2009. The conserved 3′ end domain of U6atac snRNA can direct U6 snRNA to the minor spliceosome. *RNA* **15:** 1198–1207.

Doktor TK, Hua Y, Andersen HS, Brøner S, Liu YH, Wieckowska A, Dembic M, Bruun GH, Krainer AR, Andresen BS. 2016. RNA-sequencing of a mouse-model of spinal muscular atrophy reveals tissue-wide changes in splicing of U12-dependent introns. *Nucleic Acids Res* **45:** 395–416.

Edery P, Marcaillou C, Sahbatou M, Labalme A, Chastang J, Touraine R, Tubacher E, Senni F, Bober MB, Nampoothiri S, et al. 2011. Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* **332:** 240–243.

Elsaid MF, Chalhoub N, Ben-Omran T, Kumar P, Kamel H, Ibrahim K, Mohamoud Y, Al-Dous E, Al-Azwani I, Malek JA, et al. 2017. Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. *Ann Neurol* **81:** 68–78.

Frilander MJ, Steitz JA. 1999. Initial recognition of U12-dependent introns requires both U11/5′ splice-site and U12/branchpoint interactions. *Genes Dev* **13:** 851–863

Frilander MJ, Steitz JA. 2001. Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. *Mol Cell* **7:** 217–226.

Golas MM, Sander B, Will CL, Lührmann R, Stark H. 2005. Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. *Mol Cell* **17:** 869–883.

Guceva Z, Polenakovicb M, Tasica V, LeBoucc Y, Klammtd J, Pfaeffled J, Filipovskae A. 2015. Severe isolated growth hormone deficiency and myopathy in two brothers with RNPC3 mutation. *Horm Res Paediatr* **84**(Suppl 1)**:**447.

Guerrero F, Ciragan A, Iwaï H. 2015. Tandem SUMO fusion vectors for improving soluble protein expression and purification. *Protein Expr Purif* **116:** 42–49.

Güntert P, Mumenthaler C, Wüthrich K. 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **273:** 283–298.

Güntert P. 2004. Automated NMR structure calculation with CYANA. In *Protein NMR techniques* (ed. Downing AK), pp. 353–378. Humana Press, Totowa, NJ.

Güntert P. 2009. Automated NMR protein structure calculation. *Eur Biophys J* **38:** 129–143

Hall SL, Padgett RA. 1996. Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* **271:** 1716–1718.

Hastings ML, Krainer AR. 2001. Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA* **7:** 471–482.

He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, Li W, Sebastian N, Wen B, Xin B, et al. 2011. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* **332:** 238–240.

Jafarifar F, Dietrich RC, Hiznay JM, Padgett RA. 2014. Biochemical defects in minor spliceosome function in the developmental disorder MOPD I. *RNA* **20:** 1078–1089.

Kay LE, Torchia DA, Bax A. 1989. Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28:** 8972–8979.

Kolossova I, Padgett RA. 1997. U11 snRNA interacts in vivo with the 5′ splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA* **3:** 227–233.

Koradi R, Billeter M, Wüthrich K. 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Grap* **14:** 51–55.

Kostic M, Pochapsky SS, Pochapsky TC. 2002. Rapid recycle ^{13}C′,^{15}N and ^{13}C,^{13}C′ heteronuclear and homonuclear multiple quantum

coherence detection for resonance assignments in paramagnetic proteins: example of Ni^{2+}-containing acireductone dioxygenase. *J Am Chem Soc* **124:** 9054–9055.

Krøigård AB, Frost M, Larsen MJ, Ousager LB, Frederiksen AL. 2016. Bone structure in two adult subjects with impaired minor spliceosome function resulting from RNU4ATAC mutations causing microcephalic osteodysplastic primordial dwarfism type 1 (MOPD1). *Bone* **92:** 145–149.

Madan V, Kanojia D, Li J, Okamoto R, Sato-Otsubo A, Kohlmann A, Sanada M, Grossmann V, Sundaresan J, Shiraishi Y, et al. 2015. Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat Comm* **6:** 6042.

Markmiller S, Cloonan N, Lardelli RM, Doggett K, Keightley MC, Boglev Y, Trotter AJ, Ng AY, Wilkins SJ, Verkade H, et al. 2014. Minor class splicing shapes the zebrafish transcriptome during development. *Proc Natl Acad Sci* **111:** 3062–3067.

Matera AG, Wang Z. 2014. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* **15:** 108–121.

Merico D, Roifman M, Braunschweig U, Yuen RKC, Alexandrova R, Bates A, Reid B, Nalpathamkalam T, Wang Z, Thiruvahindrapuram B, et al. 2015. Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. *Nat Comm* **6:** 8718.

Netter C, Weber G, Benecke H, Wahl MC. 2009. Functional stabilization of an RNA recognition motif by a noncanonical N-terminal expansion. *RNA* **15:** 1305–1313.

Niemelä EH, Verbeeren J, Singha P, Nurmi V, Frilander MJ. 2015. Evolutionarily conserved exon definition interactions with U11 snRNP mediate alternative splicing regulation on U11–48K and U11/U12–65K genes. *RNA Biol* **12:** 1256–1264.

Park SJ, Jung HJ, Nguyen Dinh S, Kang H. 2016. Structural features important for the U12 snRNA binding and minor spliceosome assembly of Arabidopsis U11/U12-small nuclear ribonucleoproteins. *RNA Biol* **13:** 670–679.

Reber S, Stettler J, Filosa G, Colombo M, Jutzi D, Lenzken SC, Schweingruber C, Bruggmann R, Bachi A, Barabino SM, et al. 2016. Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. *EMBO J* **35:** 1504–1521.

Sattler M, Schleucher J, Griesinger C. 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* **34:** 93–158.

Schneider C, Will CL, Makarova OV, Makarov EM, Lührmann R. 2002. Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Molecular and cellular biology* **22** (10): 3219–3229

Shen Y, Bax A. 2013. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56:** 227–241.

Shen H, Zheng X, Luecke S, Green MR. 2010. The U2AF35-related protein Urp contacts the 3′ splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes Dev* **24:** 2389–2394.

Singh RK, Cooper TA. 2012. Pre-mRNA splicing in disease and therapeutics. *Trend Mol Med* **18:** 472–482.

Singh J, Sikand K, Conrad H, Will CL, Komar AA, Shukla GC. 2016. U6atac snRNA stem-loop interacts with U12 p65 RNA binding protein and is functionally interchangeable with the U12 apical stem-loop III. *Sci Rep* **6:** 31393.

Tarn WY, Steitz JA. 1996a. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* **273:** 1824–1832.

Tarn WY, Steitz JA. 1996b. A novel spliceosome containing U11, U12 and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* **84:** 801–811.

Tidow H, Andreeva A, Rutherford TJ, Fersht AR. 2009. Solution structure of the U11-48K CHHC zinc-finger domain that specifically binds the 5′ splice site of U12-type introns. *Structure* **17:** 294–302.

Turunen JJ, Niemelä EH, Verma B, Frilander MJ. 2013a. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* **4:** 61–76.

Turunen JJ, Verma B, Nyman TA, Frilander MJ. 2013b. HnRNPH1/H2, U1 snRNP and U11 snRNP co-operate to regulate the stability of the U11-48K pre-mRNA. *RNA* **19:** 380–389.

Turunen JJ, Will CL, Grote M, Lührmann R, Frilander MJ. 2008. The U11-48K protein contacts the 5′ splice site of U12-type introns and the U11-59K protein. *Mol Cell Biol* **28:** 3548–3560.

Verbeeren J, Niemelä EH, Turunen JJ, Will CL, Ravantti JJ, Lührmann R, Frilander MJ. 2010. An ancient mechanism for splicing control: U11 snRNP as an activator of alternative splicing. *Mol Cell* **37:** 821–833.

Verbeeren J, Verma B, Niemelä EH, Yap K, Makeyev EV, Frilander MJ. 2017. Alternative exon definition events control the choice between nuclear retention and cytoplasmic export of U11/U12-65K mRNA. *PLoS Genet* **13:** e1006824.

Verma B, Akinyi MV, Norppa AJ, Frilander MJ. 2018. Minor spliceosome in disease. *Sem Cell Dev Biol*. doi: 10.1016/j.semcdb. 2017.09.036.

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59:** 687–696.

Wahl MC, Will CL, Lührmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136:** 701–718.

Wassarman KM, Steitz JA. 1992. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol Cell Biol* **12:** 1276–1285.

Will CL, Lührmann R. 2011. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3:** a003707.

Will CL, Schneider C, Hossbach M, Urlaub H, Rauhut R, Elbashir S, Tuschl T, Lührmann R. 2004. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **10:** 929–941.

Will CL, Schneider C, Reed R, Lührmann R. 1999. Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science* **284:** 2003–2005.

Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, et al. 2011. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478:** 64–69.

Yu YT, Steitz JA. 1997. Site-specific crosslinking of mammalian U11 and U6atac to the 5′ splice site of an AT–AC intron. *Proc Natl Acad Sci* **94:** 6030–6035.