

Article

Not peer-reviewed version

Pathogenic Variants in USH1G/SANS Alter the Protein Interaction With the Pre-RNA Processing Factors PRPF6 and PRPF31 of the Spliceosome

[Jacques S. Fritze](#) , Felizitas F. Stiehler , [Uwe Wolfrum](#) *

Posted Date: 9 November 2023

doi: 10.20944/preprints202311.0609.v1

Keywords: splicing, U4/U6.U5 tri-snRNP, Usher syndrome, protein-protein interaction, FRET, AlphaFold2, in silico structure predictions,



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Pathogenic Variants in USH1G/SANS Alter the Protein Interaction with the Pre-RNA Processing Factors PRPF6 and PRPF31 of the Spliceosome

Jacques S. Fritze, Felizitas F. Stiehler and Uwe Wolfrum *

Institute of Molecular Physiology, Johannes Gutenberg University Mainz, Mainz, Germany

* Correspondence: wolfrum@uni-mainz.de

Abstract: Pre-mRNA splicing is an essential process orchestrated by the spliceosome, a dynamic complex assembled stepwise on pre-mRNA. We have previously identified that USH1G protein SANS regulates pre-mRNA splicing by mediating the intra-nuclear transfer of the spliceosomal U4/U6.U5 tri-snRNP complex. During this process, SANS interacts with the U4/U6 and U5 snRNP-specific proteins PRPF31 and PRPF6 and regulates splicing, which is disturbed by variants of USH1G/SANS causative for human Usher syndrome (USH) the most common form of hereditary deaf-blindness. Here, we aimed to gain further insight into the molecular interaction of the splicing molecules PRPF31 and PRPF6 to the CENTn domain of SANS using fluorescence resonance energy transfer assays in cells and in silico deep learning-based protein structure predictions. This demonstrates that SANS directly binds via two distinct conserved regions of its CENTn to the two PRPFs. In addition, we provide evidence that these interactions occur sequentially and a conformational change of an intrinsically disordered region to a short α -helix of SANS CENTn2 is triggered by binding of PRPF6. Furthermore, we found that pathogenic variants of USH1G/SANS perturb the binding of SANS to both PRPFs implying a significance for the USH1G pathophysiology.

Keywords: splicing; U4/U6.U5 tri-snRNP; usher syndrome; protein-protein interaction; FRET; AlphaFold2; in silico structure predictions,

1. Introduction

Pre-mRNA splicing is a fundamental process in eukaryotic cells that almost exclusively occurs in the nucleus. It allows the inclusion/exclusion of exons/introns in the mRNA that can give rise to the synthesis of multiple alternative protein isoforms from a single genomic locus [1]. This process is catalyzed by the spliceosome, a highly dynamic macromolecular complex composed of five small nuclear ribonucleoproteins (U1, U2, U5, U4/U6 snRNPs) and numerous other polypeptides [2]. We have recently shown that the Usher syndrome type 1G (USH1G) protein SANS (scaffold protein containing ankyrin repeats and SAM domain) regulates pre-mRNA splicing [3]. Here, SANS mediates the intra-nuclear transfer of the spliceosomal U4/U6.U5 tri-snRNP complexes from the Cajal body and nuclear speckles, which is essential for the spliceosome activation [3,4]. We found that during these processes SANS interacts, among others, with the spliceosomal U4/U6 and U5 snRNP-specific proteins PRPF31 and PRPF6, respectively. *USH1G/SANS* deficiency altered the kinetics of spliceosome activation leading to perturbations in constitutive and alternative splicing of target genes especially genes related to the human Usher syndrome.

Human Usher syndrome (USH) is a complex autosomal-recessive rare genetic disorder (prevalence 1:5,000 -1:10,000) but represents the most common form of deaf-blindness [5]. To date, at least 11 genes have been classified into four clinical subtypes (USH1-4), among them USH1 is the most severe subtype with profound hearing loss, vestibular dysfunction, and pre-pubertal onset of retinal dysfunctions in the form of retinitis pigmentosa [5–7].

The human *USH1G* gene encodes SANS, a scaffold protein of ~ 52 kDa molecular weight consisting of 461 amino acids. SANS is composed of three N-terminal ankyrin repeats (ANK1-3), two central domains (CENTn and CENTc), a sterile alpha motif (SAM), and PDZ binding motif at the

very C-terminal end (Figure 1) [8,9]. In the auditory hair cells, SANS is essential for the correct arrangement of the stereocilia in the hair bundles during development and for the composition and function of the mechanosensitive tip-link complex at the tip of the stereocilia [10–12]. Less is known about the function of SANS in the retina. Previous work indicated that SANS associates with cellular modules such as intracellular transport, endocytosis, and primary ciliogenesis [9,13–17]. Additionally, the role of SANS and other USH proteins in stabilizing the outer segment of photoreceptors has been hypothesized due to their presence in the calyceal processes, which are microvilli-like processes of the inner segment of photoreceptor cells [18,19]. However, *USH1G/SANS* is not primarily expressed in photoreceptor cells but in Müller glial cells of the retina as recently highlighted by single-cell RNAseq data [20].

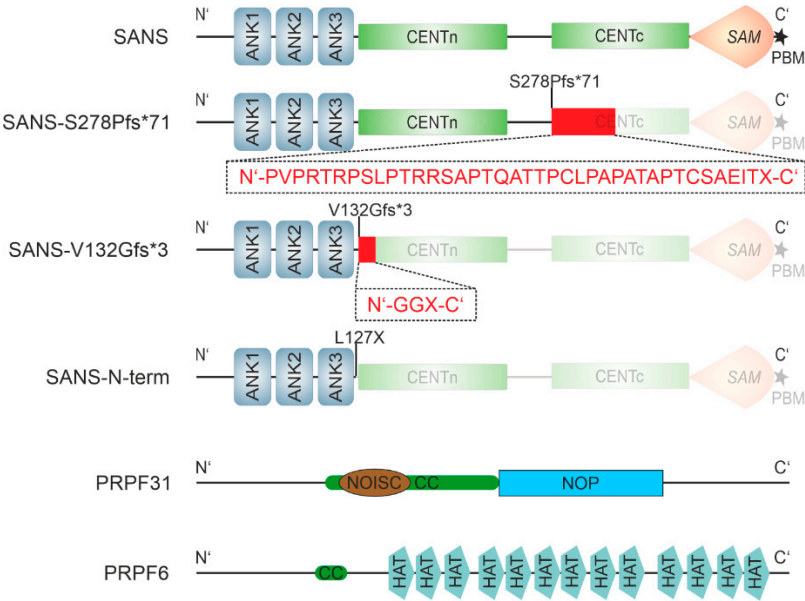


Figure 1. Domain structures of SANS variants and the pre-mRNA processing factors (PRPF) PRPF31, and PRPF6. Domain structure of SANS, USH-causing pathogenic variants SANS^{S278Pfs*71}, SANS^{V132Gfs*3}, SANS-Nterm, PRPF31, and PRPF6. SANS consists of three ankyrin repeats (ANK1-3), a central domain (CENT), divided into the N-terminal CENTn and the C-terminal CENTc, a sterile alpha motif (SAM), and PDZ binding motif (PBM, asterisk). Both USH-causing mutants S278Pfs*71 and V132Gfs*3 are frameshift mutations, which lead to “missense extension” (red boxes and amino acid sequences below) and a premature stop in CENTc (S278Pfs*71) or CENTn (V132Gfs*), respectively. SANS-Nterm is an artificial construct missing CENTn and onward. PRPF31 consists of coiled-coil (CC, green) regions, a NOISC domain (brown), named after the central domain in Nop56/SIK1-like proteins and a NOP domain (blue), a ribonucleoprotein (RNP) binding module, both found in various pre-RNA processing ribonucleoproteins (InterPro: <http://www.ebi.ac.uk/interpro/entry/InterPro/IPR012976> and IPR036070). PRPF6 exhibits thirteen half a TPR (HAT) domains as well as one CC domain.

In any case, the integrative scaffold functions of SANS are essential for the proper operation of diverse cellular modules, relying on context-dependent interactions with multiple protein partners. Others and we have identified numerous proteins that bind to all domains and motifs of the SANS molecule [9,13,15,17,21–23]. These studies also showed the CENT domain is the preferential binding site in SANS, with at least 50 different proteins binding to it [3,9,14,17,21]. Nevertheless, the mechanisms by which the CENT domain of SANS facilitates its interaction with multiple binding partners remain unexplained.

Here, we aimed to gain further insight into the molecular interaction of binding partners with the CENTn domain and to understand the effects of SANS-mediated splicing regulation. We investigated in detail the binary interaction of SANS with both PRPF31 and PRPF6 in cells using FRET assays and *in silico* deep learning-based protein structure predictions such as AlphaFold2-multimer and Metapredict [24,25]. Our data suggest that SANS CENTn domain binds directly via

two distinct conserved regions (CENTn1, CENTn2) to PRPF31 and PRPF6. We also provide evidence that these interactions occur sequentially and that PRPF6 binding introduces a conformational change of an intrinsically disordered region to an α -helix in SANS CENTn2. Finally, we found that perturbations in the binary interactions between the two PRPFs and pathogenic variants of USH1G/SANS were disturbed, implying a significance in the pathogenesis of USH1G.

2. Results

2.1. Binding of SANS to PRPF31 and PRPF6 in the nucleus revealed by FRET

We have previously demonstrated the interaction of the USH1G protein SANS with the splicing molecules PRPF31 and PRPF6 of the tri-snRNP complex by *in vitro* pull-down assays [3]. Here, we aimed to explore their binary binding in more detail by fluorescence resonance energy transfer (FRET) acceptor photobleaching assays in the cell [32]. For this, we co-expressed eYFP-SANS together with eCFP-tagged eCFP-PRPF31 or PRPF6-eCFP, in HEK293T cells. Confocal fluorescence microscopy showed that eYFP-SANS co-localized with both PRPF6-eCFP and eCFP-PRPF31, respectively, in the nucleus, as indicated by fluorescence intensity plots and confirmed by positive Pearson coefficient values (Figure 2A, B). In addition, eCFP-tagged harmonin, a USH-related scaffold protein known to co-localize *in situ* and to interact with SANS [8,21,33], co-localized with co-expressed eYFP-SANS as expected (Figure 2C). Intriguingly, the co-localization of eYFP-SANS with harmonin-eCFP was predominantly in the cytoplasm. The triple transfection of PRPF6-eCFP, eYFP-SANS, and mRFP-PRPF31 showed co-localization of all three proteins inside the nucleus (Figure 2D), consistent with the localization of endogenous proteins as previously reported [3].

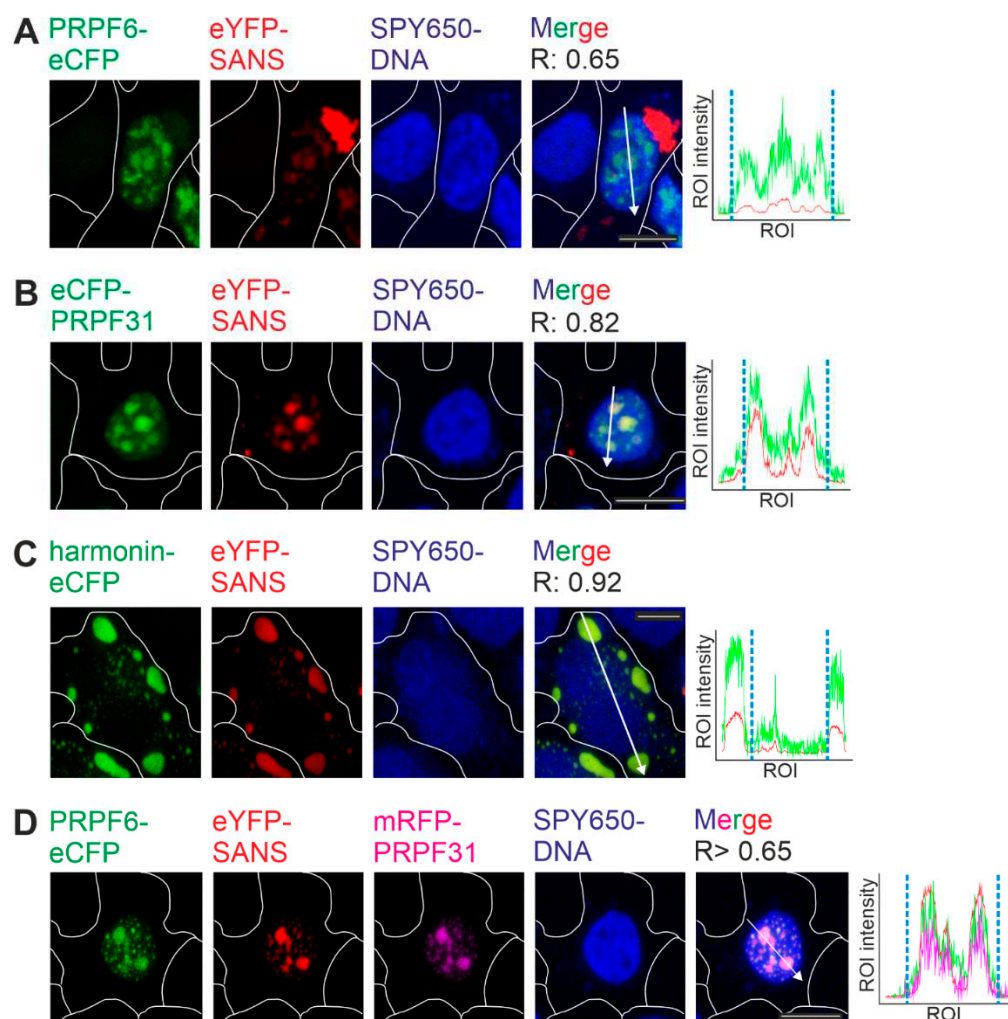


Figure 2. Co-localization of SANS and splicing proteins PRPF6 and PRPF31. **(A-D)** Confocal microscopy analyses of HEK293T cells co-transfected eYFP-SANS and eCFP/mRFP tagged proteins, counterstained with SPY650-DNA as a nuclear marker. Fluorescence intensity plots of eYFP SANS (red) and transfected eCFP- (green) and mRFP-tagged proteins for regions of interest (ROI) indicated by white arrows in merge images; blue dashed lines indicate nuclear extension. Positive Pearson coefficient R values indicate co-localization. **(A)** PRPF6-eCFP co-localized with eYFP-SANS in the nucleus. **(B)** eCFP-PRPF31 co-localized with eYFP-SANS in the nucleus. **(C)** USH1C protein harmonin-eCFP co-localized with eYFP-SANS predominantly in the cytoplasm outside of the nucleus. **(D)** eYFP-SANS co-localized with PRPF6-eCFP and mRFP-PRPF31 in the nucleus of HEK293T cell in a triple transfection. White lines indicate cell borders. Example images of n=3 experiments; scale bar = 10 μ m.

To study the binary interaction of SANS with both PRPFs in the cell, we performed FRET acceptor photobleaching assays in the nucleus (Figure 3A). FRET analyses revealed the interaction of the eCFP-eYFP FRET pairs SANS-PRPF31 and SANS-PRPF6 in nuclei (Figure 3B). The obtained FRET signals were at the same height as for the interaction between eYFP-SANS and the known SANS binding partner protein harmonin-eCFP in the cytoplasm (Figure 3B). In addition, we analyzed PRPF3, another PRPF of the spliceosomal tri-snRNP complex, previously shown not to interact with SANS as a negative control [3]. We did not observe notable FRET signals in assays with the SANS-PRPF3 eCFP-eYFP FRET pair (Figure 3C) and in all additional negative controls applied (Figure S1A-C). Taken together the FRET-based interacting assays revealed specific binding of SANS to PRPF31 and PRPF6, respectively in the nucleus.

During sequential spliceosome activation, the interaction of PRPF31 and PRPF6 as bridging factors between the small nuclear ribonucleoproteins snRNPs, U4/U6, and U5 of the tri-snRNP complex is essential [34]. This prompted us to test whether SANS interferes with the interaction between PRPF31 and PRPF6 within the tri-snRNP complex. As expected, we found significant FRET signals in HEK293T cells co-transfected with eCFP-PRPF31 and PRPF6-eYFP (Figure 3D). Additionally expressed mRFP-SANS or mRFP alone did not alter the FRET signals for eCFP-PRPF31 and PRPF6-eYFP. Next, we conducted FRET analysis for eCFP-eYFP FRET pairs of SANS-PRPF31 in the presence of PRPF6-mCherry (Figure 3E). The FRET signals between eYFP-SANS and eCFP-PRPF31 were unaffected by PRPF6-mCherry (Figure 3E). However, co-expression of mRFP-PRPF31 significantly reduced FRET signals between SANS-PRPF6 FRET pairs, which was not observed in the presence of mRFP (Figure 3F). However, FRET efficiency decreased only by approximately 16%, which still indicates the interaction of SANS and PRPF6 but with lower affinity than in the absence of PRPF31. This decrease may indicate the competition of binding partners and/or a change in conformation of the tertiary complex of PRPF31-PRPF6 and SANS.

Taken together, these data suggest that the binary binding of SANS to PRPF31 is independent of PRPF6, but that PRPF31 interferes with the interaction of SANS with PRPF6 (Figure 3G).

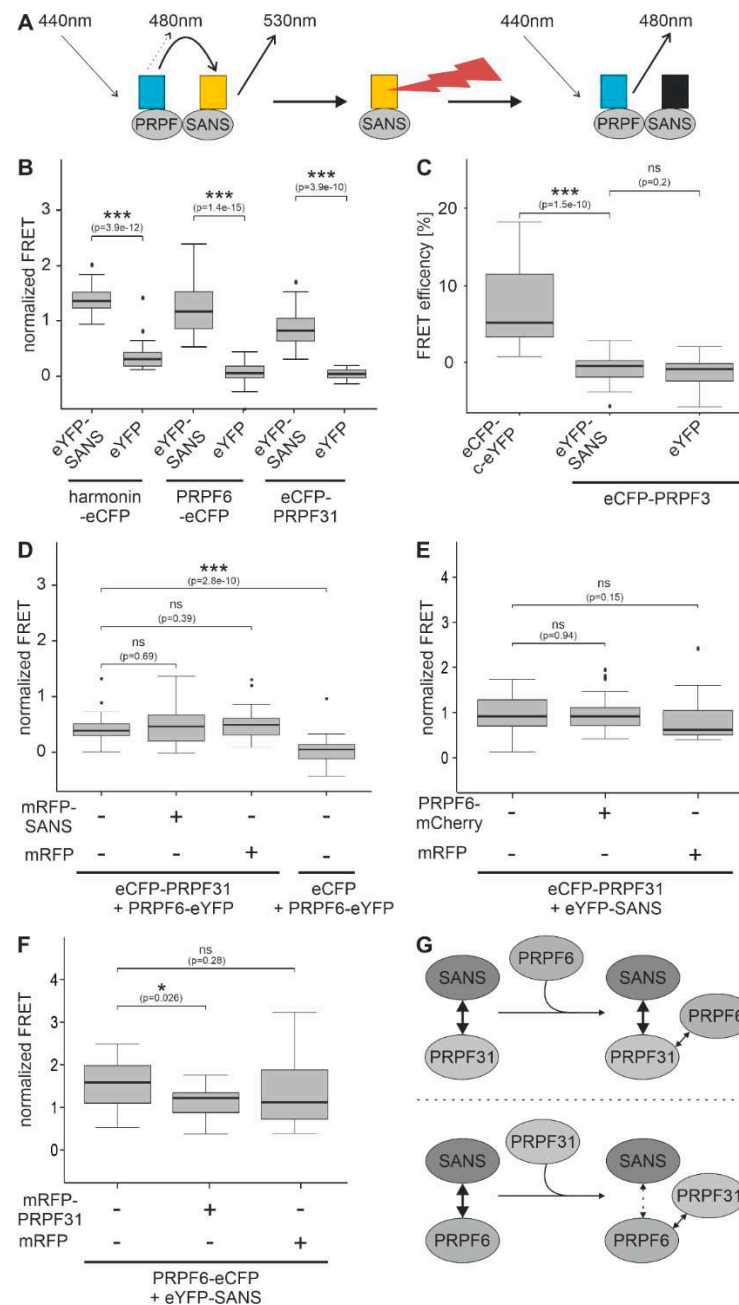


Figure 3. Interaction of SANS and binding partners analyzed by FRET acceptor bleach in HEK293T cells. **(A)** Illustration of the principle of the FRET acceptor bleach assay. Interaction of two proteins tagged with either eCFP (blue) or eYFP (yellow) leads to FRET (left). The acceptor (eYFP) is bleached (flash symbol) (middle), which leads to increased emission of the donor (eCFP) (right). **(B-F)** FRET assay in co-transfected HEK293T cells. In B, D-F, FRET efficiencies were normalized to the fused eCFP-c-eYFP FRET pair (positive control) as shown in C and Figure. S1. **(B)** FRET pairs eYFP-SANS-harmonin-eCFP, -PRPF6-eCFP, and -eCFP-PRPF31, respectively, show a significant increase in the normalized FRET efficiencies when compared to eYFP FRET pair negative controls. **(C)** FRET efficiency of eYFP-SANS-eCFP-PRPF3 shows a FRET efficiency similar to the eYFP negative control, significantly different from the positive control eCFP-c-eYFP. **(D)** FRET efficiencies of the eCFP-PRPF31-PRPF6-eYFP pair in the absence and presence of mRFP-SANS do not significantly differ. **(E)** Normalized FRET efficiencies of the eCFP-PRPF31-eYFP-SANS in the absence and presence of PRPF6-mCherry do not significantly differ. **(F)** Normalized FRET efficiency of the PRPF6-eCFP-eYFP-SANS is significantly decreased in the presence of mRFP-PRPF31. In E and F, mRFP was used as a control for the third interaction partner. Outliers are shown as dots above/below the boxplots. Wilcoxon

signed-rank test was performed for 3 independent experiments. **(G)** Schematic interaction of SANS, PRPF31 and PRPF6 from D-F. SANS interaction with PRPF6 is decreased in the presence of PRPF31.

2.2. PRPF31 and PRPF6 interact with different regions in SANS CENTn domain

Our FRET experiments indicated that SANS interacts differently with PRPF31 and PRPF6. To explore the structural basis for this interaction we used AlphaFold2-multimer as an *in silico* approach for protein-protein complex predictions [24]. The AlphaFold2 structural predictions aligned to the structures of PRPF31 and PRPF6 previously obtained by cryoelectron microscopy ([35], PDB: 6QW6) (Figure S2A,B). For SANS, AlphaFold2 predicted a central unstructured intrinsically disordered region (IDR) flanked by an α -helix representing the N-terminus of CENTn and the SAM domain in the C-terminus (Figure S2C).

To analyze the protein complexes, we used AlphaFold2-multimer [24], which predicts an alignment error (PAE) representing the accuracy of the predicted models, summarized in the PAE_{sum} score for each model (Figure S3A). We compared our prediction workflow with the harmonin-SANS complex previously determined by X-ray crystallography as a benchmark [33]. AlphaFold2-multimer predicted a complex similar to the experimentally resolved structure with a $PAE_{sum} \geq 12.24$ Å in the interaction region, namely SANS-SAM/PBM and the N-terminal harmonin homology domain of harmonin (Figure S3B) [36]. It should be noted that during development, AlphaFold2 was trained on accessible PDB data (Jumper et al. 2021, Evans et al. 2021), and the data for SANS/harmonin structures may have been among them.

AlphaFold2-multimer predicted low confidence ($PAE_{sum} \geq 25.66$ Å) for a complex of full-length PRPF31 and SANS, except for two small regions containing the N-terminal domains of both proteins (Figure S3C), namely, the N-terminal part of the CENTn domain (aa 128-173) of SANS, and the NOP domain (aa 215-333) of PRPF31 (Figure 1). Subsequent AlphaFold2-multimer predictions of these areas predicted a complex of SANS CENTn and the PRPF31 NOP domain with high confidence of a $PAE_{sum} \geq 11.83$ Å (Figure 4A). No other regions of SANS such as the ankyrin repeats, CENTc domain, and SAM/PBM domain were predicted in a complex with the PRPF31 NOP domain with such high confidence (Figure S3D-F). The amino acid sequence analyses of these domains predicted that the SANS CENTn-PRPF31 NOP complex is stabilized by multiple hydrophobic interactions (aa 140-165, Figure 4A, Table S2). Changes to less hydrophobic amino acids in SANS CENTn sequence ($\Delta Hydro$) resulted in an AlphaFold2-multimer prediction of low confidence for the PRPF31-SANS complex ($PAE_{sum} \geq 19.62$ Å) (Figure S3G, Table S2). In summary, these data suggest a protein complex of SANS N-terminal CENTn domain and PRPF31 NOP domain, stabilized by hydrophobic interactions.

AlphaFold2-multimer predictions for a complex of full-length PRPF6 and SANS were also of low confidence in the model ($PAE_{sum} \geq 31.02$ Å) (Figure S4A). However, AlphaFold2-multimer predicted a complex of the C-terminal part of PRPF6, containing HAT domains, and a short region of SANS CENTn (aa 166-194) domain with high confidence of a $PAE_{sum} \geq 10.21$ Å (Figure 4B). Interestingly, AlphaFold2-multimer predicted that a short α -Helix (aa 183-198) was formed in this complex by a part of the unstructured region of SANS CENTn domain. No other region of SANS could be predicted with PRPF6 in such high confidence (Figure S4B-E). Taken together, our *in silico* experiments suggest two different regions of the CENTn domain, hereafter referred to as CENTn1 (aa 128-173) and CENTn2 (aa 174-243). These data suggested direct binding of PRPF31 and PRPF6 to SANS. PRPF31 interacted exclusively with CENTn1 while PRPF6 interacted predominantly with CENTn2 and 7 residues of CENTn1 indicating that the binding site for both PRPFs do not overlap (Figure 4C).

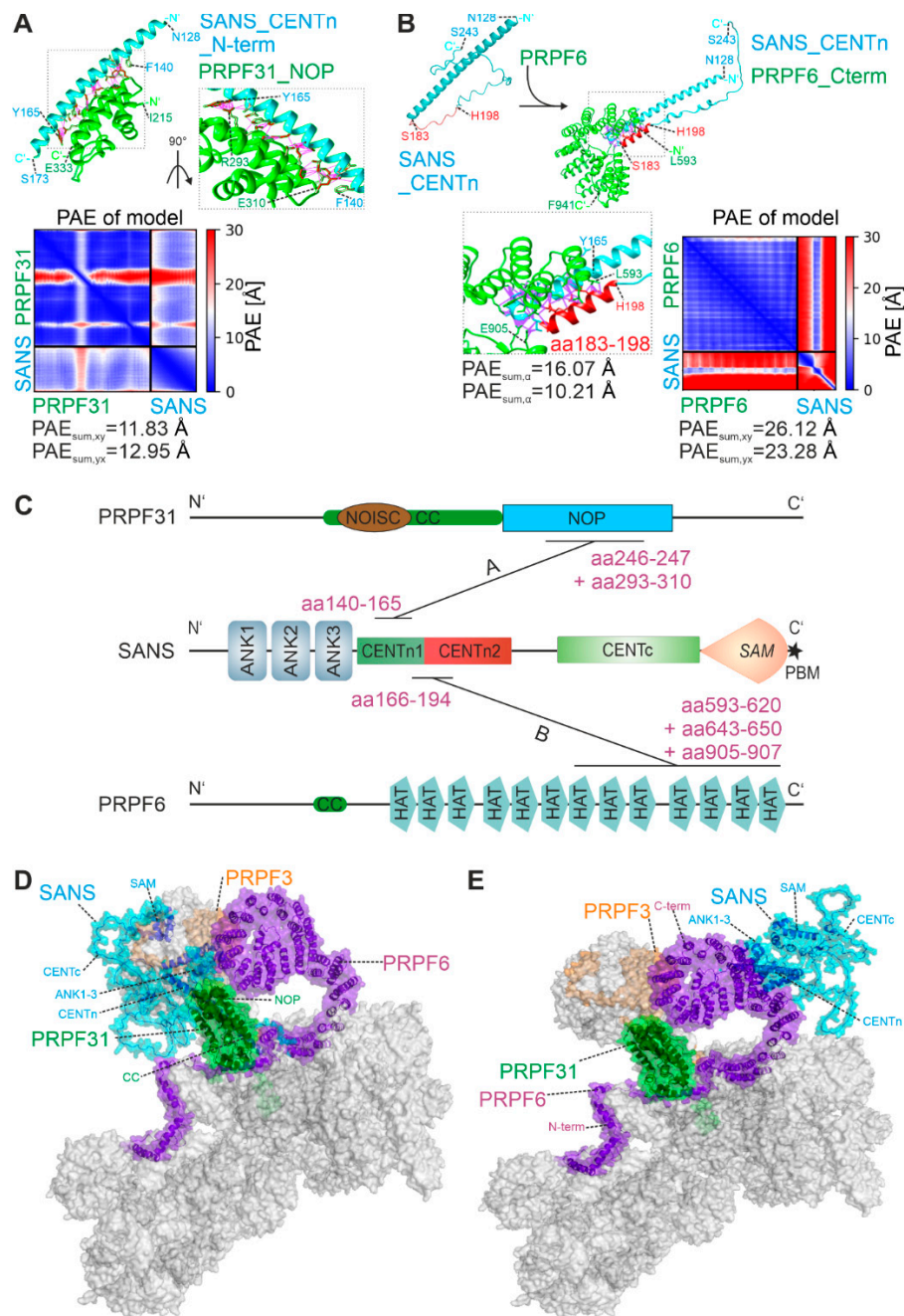


Figure 4. Molecular structure predictions of the interaction of SANS and PRPF31 and PRPF6. **(A-B)** *In silico* predictions of SANS and PRPFs complexes by AlphaFold2. Purple lines indicate hydrophobic interactions (Table S2). **(A)** PRPF31 NOP domain (green, amino acids (aa) 215-333) binds to the N-terminal part of SANS CENTn (blue, aa 128-173). Close-up is shown on the right (zoomed box). **(B)** PRPF6 binds via aa 593-941 of the C-terminus (green) to the SANS CENTn domain (blue, aa 128-243); PAE is predicted with high confidence for a small region of the two proteins. In SANS CENTn a small α -Helix (red, aa 183-198) is formed from an unstructured region upon interaction with PRPF6 (zoomed box). **(C)** Illustration of the binding regions on the domain structure of SANS, PRPF31 and PRPF6 predicted **A** and **B**, aa of binding sites are indicated in purple. Note: SANS CENTn is divided into CENTn1 (aa 128-173) and CENTn2 (aa 174-243) based on structural differences (see **A**, **B** and **Figure 5**). **(D, E)** Illustration of the SANS binding to the U5.U4/U6-tri-snRNP complex (cryoEM structure, Charenton et al., 2019, PDB: 6QW6) as predicted in **A** and **B**. SANS (blue) fits into the tri-snRNP complex in a small pocket between PRPF31 (green) and PRPF6 (purple) when interacting with PRPF31 **(D)**. SANS position in the tri-snRNP complex changes when interacting with PRPF6 **(E)**.

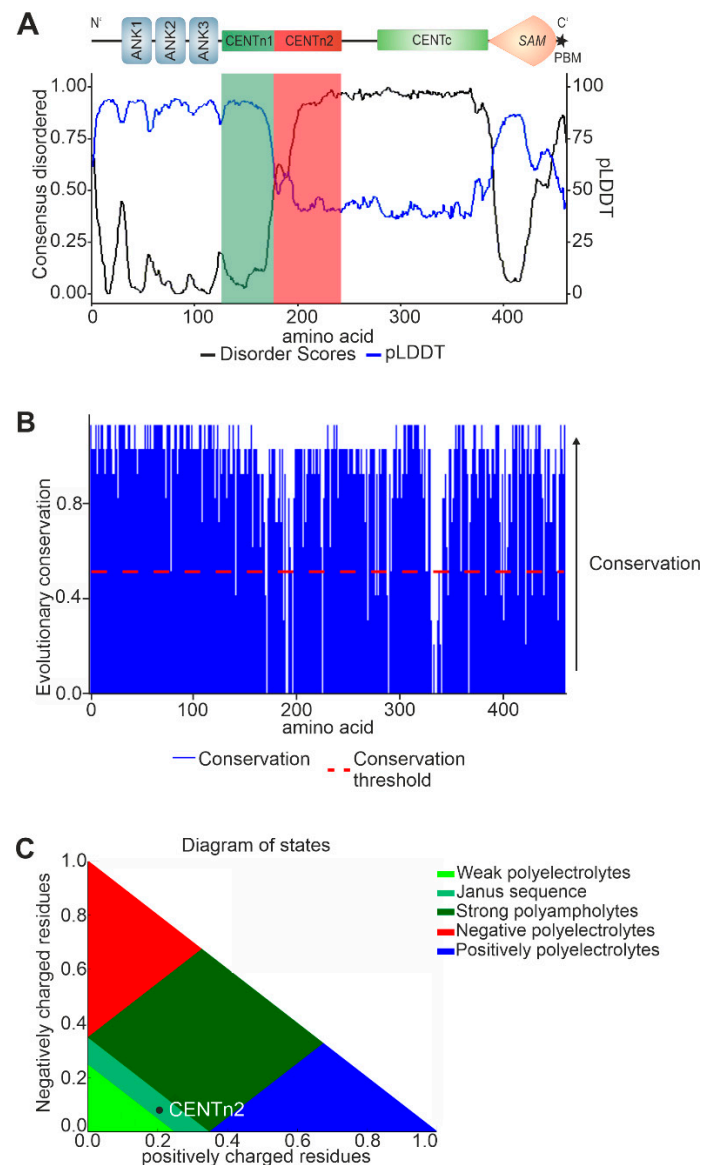


Figure 5. Predictions of SANS domain structure and their evolutionary conservation. **(A)** SANS predicted disordered state by Metapredict (black line) and AlphaFold2 (blue line). An increase of Metapredict prediction (disordered score) above 0.5 or a decrease in the AlphaFold2 prediction (shown by pLDDT) below 50 indicates an intrinsically disordered region (IDR). Metapredict and AlphaFold2 predicted SANS CENTn1 as structured (green box) and SANS CENTn2 (red box) as IDR. **(B)** Evolutionary conservation (blue) of SANS proteins for 266 vertebrates. The conservation threshold (red dotted line) indicates high conservation of SANS in vertebrates except for small regions of CENTn and CENTc domains. **(C)** SANS CENTn2 (black dot) in a Das-Pappu plot predicted by CIDER is identified as a flexible Janus sequence.

To predict the integration of SANS with the tri-snRNP complex, we made use of the AlphaFold2-multimer predictions for SANS structural complexes with PRPF31 and PRPF6 described above (Figure 4A, B) and fitted these in the experimental confirmed tri-snRNP complex ([35], PDB: 6QW6) (Figure 4D, E). However, SANS did not fit into the structure of tri-snRNP complex in a single conformation based on the predicted interactions with both PRPFs. Also, it should be noted that PRPF31 interaction with SANS (Figure 4D) may clash with SNU13 and the U4-snRNA in the U4.U6 di-snRNP. Nevertheless, in both predicted conformations SANS was structurally separated from PRPF3. Taken together, these data suggest a spatial separation of SANS binding platform in the tri-snRNP complex.

2.3. *In silico* analysis predicts evolutionary conserved multi-conformational intrinsically disordered regions (IDRs) for SANS

To decipher the structural characteristics of SANS domains we used the *in silico* tool Metapredict [25] in combination with pLDDT scores of AlphaFold2 (Figure 5A). Metapredict is a bidirectional recurrent neural network that gives a value of predicted disorder for each residue. While high values in Metapredict indicate intrinsically disordered regions (IDRs) the pLDDT scores of AlphaFold2 decrease in IDRs [37].

Metapredict and AlphaFold2 predicted the ankyrin repeats (aa 1-126), the CENTn1 (aa 128-177), and the SAM domain (aa 386-447) as folded and three unfolded IDRs in SANS, namely the CENTn2 domain (aa 178-243), the CENTc domain (aa 278-385), and the PBM (aa 451-461) (Figure 1, 5A). These predictions also strengthen the observation that SANS CENTn domain consists of two different regions, a structured region (CENTn1, aa 128-177) and an IDR (CENTn2, aa 178-243).

To analyze the evolutionary conservation of SANS domains, we performed multiple sequence alignments. The alignment of SANS amino acid sequences of over 98 mammals and 266 vertebrate species demonstrated that SANS overall is highly conserved among species (Figure 5B; Figure S5A, Table S3). Even IDRs, which are commonly known for evolutionary low conservation [38], namely the CENTn2, CENTc and the C-terminal PBM occur to be highly conserved overall vertebrate and mammal species (Figure 5B, Figure S5A).

Next, we analyzed the physicochemical properties of SANS CENTn2 domain by applying the CIDER analyzer [39]. The CENTn2 of SANS was predicted by CIDER to be a Janus sequence (Figure 5C). Janus sequences can change their structure depending on the binding partner [40]. Furthermore, CIDER predicted that SANS CENTn2 is positively charged overall (Figure S5B, C), a feature that is known to foster protein binding [41].

In summary, our results indicate three distinct conserved IDRs in CENTn2, CENTc, and PBM of SANS. Moreover, the physicochemical properties predicted for SANS CENTn2 indicate a high potential for protein-protein interactions, which is confirmed by the binding of PRPF6 to this domain described above.

2.4. *USH*-causing pathogenic variants of SANS show altered interaction with PRPFs

We have previously shown that pathogenic variants in the *USH1G* gene interrupt the interactions in the ternary *USH* protein complex of SANS, *USH2A*, and whirlin [9]. Here, we tested whether *USH1G* pathogenic frameshift mutations *USH1G*/SANS^{V132Gfs*3} and *USH1G*/SANS^{S278Pfs*71} (www.LOVD.nl/USH1G), which lead to a premature stop in the CENTn1 or CENTc domain, respectively (Figure 1), also alter the binary interaction between SANS and PRPFs.

We co-expressed eYFP-tagged versions of both *USH1G*/SANS mutations, with either eCFP-tagged PRPF31, PRPF6, or harmonin as control, in HEK293T cells (Figure S6). Confocal microscopy and fluorescent intensity plots demonstrated that both *USH1G*/SANS variants still co-localized with both PRPFs in the nucleus, confirmed by the positive Pearson coefficients (R) (Figure S6A, B). However, while in co-expressions with the *USH1G*/SANS variants, many PRPF6 droplets were still distributed all over the nucleus as in the *USH1G*/SANS wildtype expressing cells, fluorescent PRPF31 appeared in a few larger droplets in the nucleus. The expression of *USH1G*/SANS variants destroyed the specific localization of harmonin in the cytoplasm and led to an even distribution of the proteins throughout the cell (Figure S6C).

Subsequent FRET assays with eYFP-SANS^{S278Pfs*71} and both eCFP-tagged PRPFs showed significant reductions in the FRET efficiencies when compared to eYFP-SANS (Figure 6A, B). This indicated that the S278Pfs*71 mutation disrupts the binding of SANS to PRPF31 and PRPF6. However, we predicted no change in the IDR for SANS^{S278Pfs*71} compared to SANS (Figure S7A).

In FRET assays with eYFP-SANS^{V132Gfs*3} and PRPF6-eCFP the FRET efficiencies were also significantly reduced compared to the efficiency of eYFP-tagged full-length SANS and PRPF6-eCFP (Figure 6C).

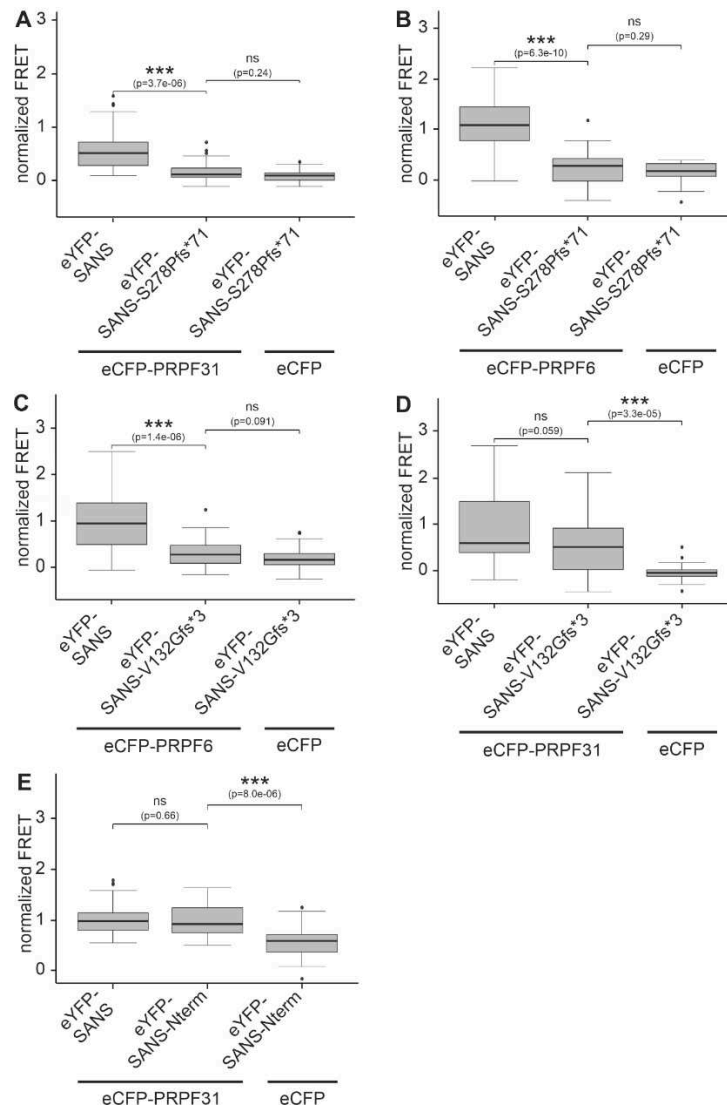


Figure 6. Interaction of pathogenic variants of *USH1G*/SANS to PRPF31 and PRPF6 analyzed by FRET in HEK293T cells. **(A-B)** FRET assays with eYFP-SANS, eYFP-SANS^{S278Pfs*71} paired with eCFP-PRPF31 **(A)** and PRPF6-eCFP **(B)** normalized to FRET of eCFP-c-eYFP fused tandem pair. Normalized FRET efficiencies are significantly decreased in both SANS^{S278Pfs*71}-PRPF pairs to levels of the negative control eYFP-SANS^{S278Pfs*71}-eCFP compared to eYFP-SANS, indicating no binding of eYFP-SANS^{S278Pfs*71} to both PRPFs. **(C)** Normalized FRET efficiency is also decreased in the eYFP-SANS^{V132Gfs*3}-PRPF6-eCFP pair indicating no binding of SANS^{V132Gfs*3} to PRPF6. **(D, E)** Normalized FRET efficiencies of the eYFP-SANS^{V132Gfs*3} **(D)** and eYFP-SANS-Nterm **(E)** paired with eCFP-PRPF31 are similar to eYFP-SANS-eCFP-PRPF31 pair values, but significantly different from eCFP control pairs indicating no binding of both truncated variants to PRPF31. Outliers are shown as dots above/below the boxplots. Wilcoxon signed-rank test was performed for 3 independent experiments.

Interestingly, in FRET assays with eYFP-SANS^{V132Gfs*3} and eCFP-PRPF31 the FRET efficiency was not reduced, indicating a binding of PRPF31 also to the N-terminal part of SANS (Figure 6D). Since this finding was contradictory to our previous results from *in vitro* pull-down assays [3] we validated this observation by additional FRET assays with the truncated eYFP-SANS-Nterm construct (Figure 1) and eCFP-PRPF31 (Figure 6E). eCFP-eYFP FRET pairs of SANS-Nterm-PRPF31 showed no reduction of the FRET efficiency compared to eYFP-tagged full-length SANS supporting binding of PRPF31 to the N-terminal part of SANS, which mainly consists of three ankyrin repeats (Figure 1).

Both *USH1G*/SANS variants (V132Gfs*3 and S278Pfs*71) did not show a significant FRET signal in all FRET controls performed (Figure S7B, C). In addition, FRET assays with harmonin, known to

bind to the SAM/PBM region of the C-terminus of SANS, which is absent in both *USH1G*/SANS variants, showed significantly reduced efficiencies when compared to full-length eYFP-SANS (Figure S7D). Thereby, the interaction with both fluorescent-tag proteins, eYFP and eCFP could be excluded. Moreover, we confirmed the correct expression of all eYFP-SANS constructs via Western blot (Figure S7E). Further, AlphaFold2-multimer did not predict a model for PRPF31 full-length or coiled-coil domain and SANS-Nterm with high confidence (Figure S8A, B).

Taken together, our data showed that both pathogenic variants *USH1G*/SANS^{V132Gfs*3} and *USH1G*/SANS^{S278Pfs*71} disrupted the interaction with PRPF6. In contrast, the interaction of PRPF31 was disrupted for *USH1G*/SANS^{S278Pfs*71} but not with the shorter eYFP-SANS^{V132Gfs*3} revealing a binding site for PRPF31 in the N-terminal part of SANS alternative to the binding to SANS CENTn1.

3. Discussion

Pre-mRNA splicing is a fundamental process in eukaryotic cells that almost exclusively occurs in the nucleus, catalyzed by the spliceosome, a highly dynamic macromolecular complex composed of small nuclear ribonucleoproteins (snRNPs) and numerous other polypeptides [1]. We previously identified the *USH1G* protein SANS as one of these proteins regulating pre-mRNA splicing by mediating the intra-nuclear transfer of the spliceosomal U4/U6.U5 tri-snRNP complexes from the Cajal body and nuclear speckles essential for the spliceosome activation [3]. During these processes, SANS interacts through the CENTn domain with the spliceosomal U4/U6 and U5 snRNP-specific proteins PRPF31 and PRPF6, respectively [3]. However, questions of the mode of binding of SANS to PRPFs, by direct or mediated by other spliceosome proteins or even by RNA, remained unanswered. In addition, the exact binding site in the CENTn domain of the SANS protein, which comprises more than one-fourth of the entire protein, has not yet been determined.

In our study, we found FRET efficiencies for both eCFP-eYFP FRET pairs PRPF31-SANS and PRPF6-SANS in the range of the tandem eCFP-c-eYFP construct (positive control), significantly higher than the negative controls. Since positive FRET signals are only achieved at a distance of 8-10 nm or less between the two complex partners [32,42], these findings did not only confirm the interactions of PRPF31 and PRPF6 to SANS but also suggest that both PRPFs bind directly to SANS without any mediator involved as previously considered [43]. Microscopic analyses, including fluorescence intensity plots and consistently positive Pearson coefficient R values, demonstrated the co-localization of PRPFs with SANS in the cell nucleus, further supporting a direct binary interaction.

Using AlphaFold2-multimer, we pinpointed the binding of PRPF31 and PRPF6 to different separate regions of the CENTn domain of SANS, namely to aa 140-165 of the long α -helix of CENTn1 and to aa 166-194, respectively. The predicted binding of PRPF31 via the NOP domain to the helical structures of SANS is also found for the interaction of PRPF31 with the helical structures in the small nuclear ribonucleoprotein 13 (SNU13) (also known as 15.5k protein or NHP2-like protein 1) of the U4/U6.U5 tri-snRNP complex [34]. Both observations are in line with the common interaction modes of NOP domains present in other ribonucleoproteins [44].

For the interaction of PRPF6 with SANS we confirmed C-terminal HAT domains as the interacting region in PRPF6 previously reported (Yildirim et al. 2021). These domains bind to the 7 amino acids of the α -helix of SANS CENTn1 but mainly to a region of SANS CENTn2, which was consistently predicted as an evolutionary conserved IDR by Metapredict, and AlphaFold2. Here, we have to note that the model of the PRPF6-SANS complex had low overall confidence by AlphaFold2-multimer probably since AlphaFold2-multimer is trained on structured protein data [24]. Nevertheless, the model is supported by the ability of PRPF6 to bind IDRs in other proteins such as the translation initiation factor eIF4E enrolled in splicing [45].

Our *in silico* predictions by AlphaFold2-multimer also suggested that this IDR of the CENTn2 forms a newly derived small α -helix upon binding to PRPF6. This structural change is supported by the observation that SANS CENTn2 represents a highly conserved Janus sequence, a protein region with the characteristic to specifically change the structure by binding different proteins [40]. The properties of a Janus sequence in the SANS CENTn domain are probably also the basis for allowing

the interaction of the numerous additional binding partners that have been previously identified [3,9,14,21].

Although our data consistently indicate that the binding of SANS to PRPF31 and PRPF6 is mediated by distinct subdomains of the SANS molecule we observed a decrease in FRET efficiency for the PRPF6-SANS pair upon co-transfection with PRPF31 indicating a competition between the two PRPFs. This finding is in line with the fact that the complexes predicted for SANS and the two PRPF, did not simultaneously fit into the structure of the U4/U6.U5 tri-snRNP complex [35]. These results suggest that SANS molecules bind to both PRPFs separately, either sequentially in the tri-snRNP complex or independently to the snRNP subunits. Sequential binding of proteins to the tri-snRNP complex was previously shown for U6-LSm ring proteins, which also change their spatial arrangement in the tri-snRNP complex, binding first to PRPF6 and then to PRPF3 during the formation of spliceosomal complex B [46]. Nevertheless, separate binding of SANS to the two PRPFs could also occur during the formation of the tri-snRNP complex in the Cajal bodies when the U4/U6 snRNP containing PRPF31 and U5 snRNP harboring PRPF6 mature separately from each other [47]. Alternatively, in the tri-snRNP complex, SANS may preferentially bind to PRPF31 as indicated by our FRET data and interacts with PRPF6 during the recycling process of the tri-snRNP complex components after activation of the spliceosome when the sufficient free U5 snRNP/PRPF6 is present [2,48]. Although the latter alternative is supported by previous work showing that SANS deficiency does not affect the maturation of the tri-snRNP complex but the recycling of U5-snRNP after activation of the spliceosome (Yildirim et al., 2021), in the tri-snRNP complex the SANS binding site of PRPF31 might already be occupied by SNU13 and U4-snRNA which could interfere with the binding of SANS [34].

More than 70 pathogenic variants in *USH1G*/SANS have been identified thus far in *USH1* patients (Fokkema et al., 2011; www.LOVD.nl/USH1G). However, the molecular mechanisms underlying the pathophysiology caused by these defects remain poorly understood. Our previous study provided evidence that SANS is part of the pre-mRNA splicing machinery and demonstrated that pathogenic variants in *USH1G*/SANS including both frameshift mutants, *USH1G*/SANS^{V132Gfs*3} and *USH1G*/SANS^{S278Pfs*71} investigated here, alter the pre-mRNA splicing of target genes [3]. These genes include *USH1C*, which is frequently alternatively spliced in the human retina [3,50]. Present FRET assays showed that for both pathogenic variants, the binding of PRPF6 was disrupted. *USH1G*/SANS^{S278Pfs*71} did not interact with either PRPF6 or PRPF31 in the cell, although the binding sites for both PRPFs in the CENTn domain are still present in the truncated SANS variant. Neither Metapredict nor AlphaFold2 predicted structural changes in *USH1G*/SANS^{S278Pfs*71} that could interfere with the binding of the two PRPFs. However, the content of proline residues in the frameshift sequence changes from 2 to 9 in *USH1G*/SANS^{S278Pfs*71} which might lead to steric changes in the CENT domain and the binding sites for the PRPFs interfering with the binding of PRPF6 and PRPF31.

The short, truncated SANS variants, SANS^{V132Gfs*3} and the SANS-Nterm construct lack the entire CENTn1 domain, which we identified above as the binding site of PRPF31, and all other downstream domains. However, both truncated SANS molecules consistently interact with PRPF31 in present FRET assays in cells, which was previously not found in *in vitro* pull-down assays [3]. The interaction of PRPF31 and SANS^{V132Gfs*3} is in line with the co-localization of both polypeptides in the nucleus of the cell observed by our microscopy analyses. Consequently, this novel finding suggests a binding site for PRPF31 in the three ankyrin repeats of the N-terminal portion of SANS alternative to SANS CENTn1. Ankyrin repeats are common protein-protein interaction platforms [51] and we have recently demonstrated that a set of intraflagellar transport (IFT) molecules bind to the ankyrin repeats of SANS [17]. Unfortunately, we failed to predict a site for PRPF31 that can bind to the region of ankyrin repeats in the N-terminus of SANS using the prediction tools employed.

In recent years, pathogenic variants leading to retinitis pigmentosa (RP) have been associated with many genes encoding proteins of the U4/U6.U5 tri-snRNP complex [4,52,53]. This list includes besides *PRPF31* (RP11) and *PRPF6* (RP60) several other components such as *PRPF8* (RP13), *PRPF3* (RP18), *SNRNP200* (RP33), and *PRPF4* (RP70). Here, we show that mutations in *USH1G*/SANS lead to disruption of binary binding properties with the key molecules of the U4/U6.U5 tri-snRNP

complex PRPF31 and PRPF6. *Vice versa*, several pathogenic RP mutations in the genes of the two PRPFs are also located in binding sites to SANS (Stenson et al., 2003; <https://www.hgmd.cf.ac.uk/ac/index.php>) and most likely also result in alterations in their interaction with SANS. Perturbations in the molecular interactions between the components of the U4/U6.U5 tri-snRNP complex are consistent with the fact that mutations in *PRPF6*, *PRPF31*, and *USH1G*/SANS have almost identical effects on splicing and that similar pathomechanisms leading to RP also caused similar retinal phenotypes.

Alternative splicing certainly increases the diversity of the transcriptome and proteome. Compared to other tissues, the retina exhibits the highest rates of alternative splicing [55,56]. Specific splicing programs are essential for the function and maintenance of the retina, uniquely regulated in a highly complex manner [57,58]. Thus, alterations in the splicing machinery may affect the retina more prominently than other tissues, which may explain why pathogenic variants in *USH1G* and in the genes of other splicing factors predominantly lead to retinal defects.

4. Materials and Methods

4.1. DNA constructs and primers

All nucleic acids used are listed in Table S1.

4.2. Cloning

cDNAs for expression of proteins were obtained by RT-PCR from human HEK293T cells and cloned into pENTR™ as described by the manufacturer (ThermoFisher, Karlsruhe, Germany, #K240020). Gateway™ LR reaction into the appropriate destination vector was performed according to the manufacturers protocol (ThermoFisher, #11791020). USH-causing mutants and SANS-Nterm, PRPF31 and PRPF6 were cloned into pENTR™ with the template described previously [3]. eCFP-c-eYFP was cloned from eYFP vector into pDest vector with eCFP.

4.3. Cell culture and cell lines

Dulbecco's modified Eagle's medium (DMEM #31966021) containing 10% heat-inactivated fetal calf serum (FCS) (Thermo Fisher #A4766801 or Cytiva, Freiburg, Germany, #SV30160.03) was used to culture HEK293T cells (ATCC: CRL-3216). Cells were fixed with 4% PFA in PBS for 10 min at room temperature. Nuclear DNA was stained by SPY650-DNA (tebubio, Offenbach, Germany, #SC501).

4.4. Fluorescence co-localization observation

HEK293T cells were seeded in 24 well plates with a density of 40,000 cells/well. Cells were transfected with Lipofectamine™ LTX as described by the manufacturer (ThermoFisher #15338100). Fixed HEK293T cells were observed with a Leica TCS SP5. Images were analyzed with Fiji (Schindelin et al., 2015, <https://imagej.net/software/fiji/downloads>). Pearson coefficient was performed with the Fiji Plugin *Coloc 2* for the whole nucleus. Intensity plot was performed over the indicated region of interest with the Fiji Plugin *Plot Profile*.

4.5. Fluorescence resonance energy transfer (FRET) acceptor photobleaching assay

HEK293T cells were seeded in 6 well plates with a density of 500.000 cells/well. Cells were transfected with Lipofectamine™ LTX as described by the manufacturer (ThermoFisher #15338100). Fixed HEK293T cells were analyzed with a Leica TCS SP8 and FRET acceptor photobleaching was performed following the Leica protocol (https://downloads.leica-microsystems.com/TCS%20SP8/Application%20Note/FRET_AB_with%20SP8-AppLetter_EN.pdf). eCFP was used as donor (D) and eYFP as acceptor (A). The acceptor was bleached at 100% laser intensity for 10 repeats. FRET efficiency was calculated via: $FRET_{eff} = \frac{(D_{post} - D_{pre})}{D_{post}}$. To exclude cellular structure the mean of six regions of interest (ROI) in the bleached area was calculated and FRET

efficiency of an unbleached region ($\geq 3 \mu\text{m}$ away from bleached ROI) was subtracted from the mean. Bleach efficiency was calculated via the formula $Bleach_{eff} = \left(1 - \frac{A_{post}}{A_{pre}}\right) * 100$. Only ROIs with $> 60\%$ bleach efficiency were used for analysis. Data was normalized by the positive control eCFP-c-eYFP if indicated. Example images of FRET acceptor bleaching are shown in Figure S9.

4.6. AlphaFold2-multimer

AlphaFold2-multimer prediction was done using the hetero-oligomer options of ColabFold [27]. Fasta sequence of the respective protein was used from UniProt (UniProt 2020) and a colon between the respective sequences simulated complexes. The following sequences were used: SANS: Q495M9; PRPF31: Q8WWY3; PRPF6: O94906; harmonin: Q9Y6N9. AlphaFold2 notebook was used in the ColabFold version (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=G4yBrceuFbf3) accessed last on 12.07.2023) and the standard options were extended by the template mode pdb70/pdb100, model type AlphaFold2-multimer_v2 recycling 24 times. Predicted structures are documented in GitHub (https://github.com/LabWolfrum/Fritze_et_al.2023.git). Structures were visualized, inspected, and superimposed using PyMOL (Schrödinger, LLC) or UCSF Chimera [28], which was also used to generate all structure images. Contacts for some protein complexes (Table S1) were predicted by the UCSF Chimera tool “Find Clashes/Contacts” (https://www.cgl.ucsf.edu/chimera/docs/ContributedSoftware/findclash/findclash.html). Following options were chosen: themselves, contact (default criteria -0.4). As output an overlap and a distance value can be observed.

The predicted protein complexes were fitted to a previously published PDB structure of the human tri-snRNP complex (Charenton et al., 2019, PDB: 6QW6)

4.7. Evolutionary Conservation

A previously described protocol was used as a foundation [29]. The complete code can be found on GitHub (https://github.com/LabWolfrum/Fritze_et_al.2023.git). The evolutionary Conservation was performed with the help of Jupyter Notebook and Jalview [30]. Python code was designed with the help of ChatGPT-3.5 (OpenAI). For graphs, areas not present in human SANS were excluded.

Western Blot

HEK293T cells were seeded in 6 well dishes with a confluency of 500,000 cells per/well. Cells were transfected with Lipofectamine™ LTX as described by the manufacturer (ThermoFisher #15338100). Cells were lysed 24 h after transfection with TritonX-100, protein amount was measured with a BCA-assay and proteins were eluted with 5x Laemmli buffer, separated by SDS-PAGE followed by Western blotting. Monoclonal mouse-anti-GFP (cross-reaction to eYFP, Proteintech # 66002-1-Ig) was used.

4.8. Additional bioinformatic analyses

CIDER was accessed via the online platform (http://pappulab.wustl.edu/CIDER/), and the FASTA sequence was analyzed with default options. SANS USH-causing mutants were chosen through the online Usher syndrome Database (www.LOVD.nl/USH1G). Metapredict was used by the Python library option [25]. PDB structures were obtained from PDBe webserver (https://www.ebi.ac.uk/pdbe/).

4.9. Statistics

Statistical analysis was performed with R-Studio [31]. The statistical methods and the significance criteria are listed in corresponding individual legends. Results are shown as a boxplot of data from at least 3 separate experiments. Significance was determined as: * $p \leq 0.05$, ** $p \leq 0.009$, *** $p \leq 0.0009$.

5. Conclusions

Our data strengthen that SANS participates in alternative splicing by the interaction with components of the U4/U6.U5 tri-snRNP complex. For this SANS directly binds with different binding sites of its CENTn to the splicing factors PRPF31 and PRPF6. Our data also suggest that SANS molecules bind sequentially to the two PRPFs in the tri-snRNP complex (Figure 7). Finally, we provide evidence that perturbations in the molecular interaction of SANS with both PRPF molecules underlie the pathophysiology of pathogenic variants in *USH1G/SANS* leading to the ocular phenotype in USH1.

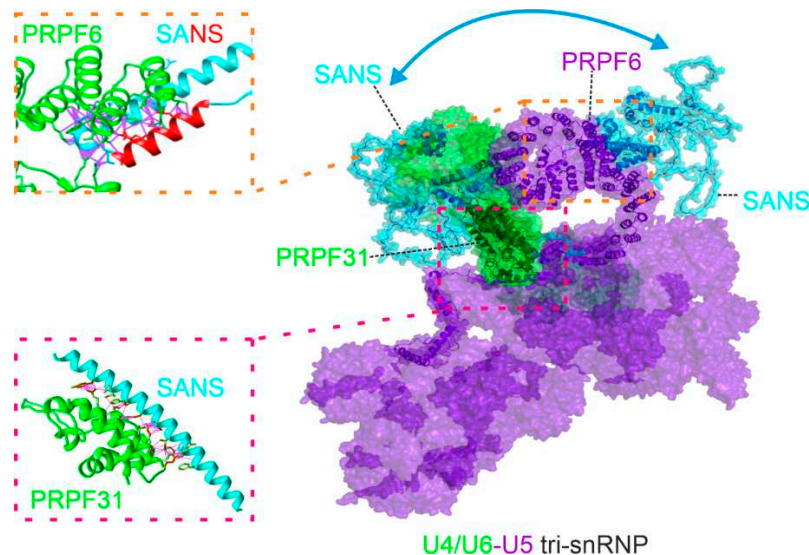


Figure 7. Graphical representation of proposed conformation of SANS in the tri-snRNP complex. SANS interacts with PRPF31 and PRPF6 in the tri-snRNP. Two conformations of SANS are used to bind to either PRPF6 or PRPF31, respectively.

Supplementary Materials: Supplementary_Figure_File.docx: all Figures S1–S8 referred to in this work; Table S1.xlsx: nucleic acids used in this study; Table S2.xlsx: contacts of PRPF31-NOP_SANS-CENTn1, contacts of PRPF6-Cterm_SANS-CENTn, contacts of PRPF31-NOP_SANS-ΔHydro. Table S3.xlsx: lists of vertebrates and mammals used for analyses of evolutionary conservation

Author Contributions: J.S.F. carried out most of the experiments. F.F.S. carried out cloning and FRET assays. J.S.F. and U.W. designed the studies and wrote the manuscript All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the FAUN-Stiftung, Nuremberg, The Foundation Fighting Blindness (FFB) [PPA-0717-0719-RAD], and the DFG (Project 521147274) to U.W.

Data Availability Statement: All raw data is available through contacting the corresponding author. Python code and AlphaFold2-multimer predictions used in this study are available via GitHub (https://github.com/LabWolfrum/Fritze_et_al.2023.git).

Acknowledgments: We thank Drs. Ronald Roepman (Human Genetics, Nijmegen), Robin Shaw (Cardiovascular Res. Inst., California) and Nasrin Soroush (iDN, JGU) for the contribution of several plasmids and constructs used in this study. We thank Christof Rickert (LMCF, JGU) for his technical support. Support by the IMB Microscopy Core Facility is gratefully acknowledged. We thank Drs. Lukas Stelzl (CMT, JGU) and Marcel Conrady (IMP, JGU) for the valuable discussion of the manuscript. J.S.F. was enrolled in the International PhD Programme (IPP) Mainz.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Papasaïkas, P.; Tejedor, J.R.; Vigevani, L.; Valcárcel, J. Functional Splicing Network Reveals Extensive Regulatory Potential of the Core Spliceosomal Machinery. *Mol Cell* 2015, 57, 7–22, doi:10.1016/j.molcel.2014.10.030.
- Will, C.L.; Lührmann, R. Spliceosome Structure and Function. *Cold Spring Harb Perspect Biol* 2011, 3, 1–2, doi:10.1101/cshperspect.a003707.
- Yildirim, A.; Mozaffari-Jovin, S.; Wallisch, A.K.; Schäfer, J.; Ludwig, S.E.J.; Urlaub, H.; Lührmann, R.; Wolfrum, U. SANS (USH1G) Regulates Pre-mRNA Splicing by Mediating the Intra-Nuclear Transfer of Tri-SnRNP Complexes. *Nucleic Acids Res* 2021, 49, 5845–5866, doi:10.1093/nar/gkab386.
- Matera, A.G.; Wang, Z. A Day in the Life of the Spliceosome. *Nat Rev Mol Cell Biol* 2014, 15, 108–121.
- Fuster-Garcia, C.; Garcia-Bohorquez, B.; Rodriguez-Munoz, A.; Aller, E.; Jaijo, T.; Millan, J.M.; Garcia-Garcia, G. Usher Syndrome: Genetics of a Human Ciliopathy. *Int J Mol Sci* 2021, 22, doi:10.3390/ijms22136723.
- Velde, H.M.; Reurink, J.; Held, S.; Li, C.H.Z.; Yzer, S.; Oostrik, J.; Weeda, J.; Haer-Wigman, L.; Yntema, H.G.; Roosing, S.; et al. Usher Syndrome Type IV: Clinically and Molecularly Confirmed by Novel ARSG Variants. *Hum Genet* 2022, 141, 1723–1738, doi:10.1007/s00439-022-02441-0.
- Peter, V.G.; Quinodoz, M.; Sadio, S.; Held, S.; Rodrigues, M.; Soares, M.; Sousa, A.B.; Coutinho Santos, L.; Damme, M.; Rivolta, C. New Clinical and Molecular Evidence Linking Mutations in ARSG to Usher Syndrome Type IV. *Hum Mutat* 2021, 42, 261–271, doi:10.1002/humu.24150.
- Weil, D.; El-Amraoui, A.; Masmoudi, S.; Mustapha, M.; Kikkawa, Y.; Laine, S.; Delmaghani, S.; Adato, A.; Nadifi, S.; Zina, Z. Ben; et al. Usher Syndrome Type I G (USH1G) Is Caused by Mutations in the Gene Encoding SANS, a Protein That Associates with the USH1C Protein, Harmonin. *Hum Mol Genet* 2003, 12, 463–471, doi:10.1093/hmg/ddg051.
- Sorusch, N.; Baub, K.; Plutniok, J.; Samanta, A.; Knapp, B.; Nagel-Wolfrum, K.; Wolfrum, U. Characterization of the Ternary Usher Syndrome SANS/Ush2a/Whirlin Protein Complex. *Hum Mol Genet* 2017, 26, 1157–1172, doi:10.1093/hmg/ddx027.
- Caberlotto, E.; Michel, V.; Foucher, I.; Bahloul, A.; Goodyear, R.J.; Pepermans, E.; Michalski, N.; Perfettini, I.; Alegria-Prévo, O.; Chardenoux, S.; et al. Usher Type 1G Protein sans Is a Critical Component of the Tip-Link Complex, a Structure Controlling Actin Polymerization in Stereocilia. *Proc Natl Acad Sci U S A* 2011, 108, 5825–5830, doi:10.1073/pnas.1017114108.
- He, Y.; Li, J.; Zhang, M. Myosin VII, USH1C, and ANKS4B or USH1G Together Form Condensed Molecular Assembly via Liquid-Liquid Phase Separation. *Cell Rep* 2019, 29, 974–986.e4, doi:10.1016/j.celrep.2019.09.027.
- Géléoc, G.G.S.; El-Amraoui, A. Disease Mechanisms and Gene Therapy for Usher Syndrome. *Hear Res* 2020, 394.
- Maerker, T.; van Wijk, E.; Overlack, N.; Kersten, F.F.J.; McGee, J.; Goldmann, T.; Sehn, E.; Roepman, R.; Walsh, E.J.; Kremer, H.; et al. A Novel Usher Protein Network at the Periciliary Reloading Point between Molecular Transport Machineries in Vertebrate Photoreceptor Cells. *Hum Mol Genet* 2008, 17, 71–86, doi:10.1093/hmg/ddm285.
- Overlack, N.; Kilic, D.; Bauss, K.; Marker, T.; Kremer, H.; van Wijk, E.; Wolfrum, U.; Bauß, K.; Märker, T.; Kremer, H.; et al. Direct Interaction of the Usher Syndrome 1G Protein SANS and Myomegalin in the Retina. *Biochim Biophys Acta* 2011, 1813, 1883–1892, doi:10.1016/j.bbamcr.2011.05.015.
- Bauss, K.; Knapp, B.; Jores, P.; Roepman, R.; Kremer, H.; Wijk, E. V.; Marker, T.; Wolfrum, U. Phosphorylation of the Usher Syndrome 1G Protein SANS Controls Magi2-Mediated Endocytosis. *Hum Mol Genet* 2014, 23, 3923–3942, doi:10.1093/hmg/ddu104.
- Papal, S.; Cortese, M.; Legendre, K.; Sorusch, N.; Dragavon, J.; Sahly, I.; Shorte, S.; Wolfrum, U.; Petit, C.; El-Amraoui, A. The Giant Spectrin BetaV Couples the Molecular Motors to Phototransduction and Usher Syndrome Type I Proteins along Their Trafficking Route. *Hum Mol Genet* 2013, 22, 3773–3788, doi:10.1093/hmg/ddt228.
- Sorusch, N.; Yildirim, A.; Knapp, B.; Janson, J.; Fleck, W.; Scharf, C.; Wolfrum, U. SANS (USH1G) Molecularly Links the Human Usher Syndrome Protein Network to the Intraflagellar Transport Module by Direct Binding to IFT-B Proteins. *Front Cell Dev Biol* 2019, 7, 216, doi:10.3389/fcell.2019.00216.
- Sahly, I.; Dufour, E.; Schietroma, C.; Michel, V.; Bahloul, A.; Perfettini, I.; Pepermans, E.; Estivalet, A.; Carette, D.; Aghaie, A.; et al. Localization of Usher 1 Proteins to the Photoreceptor Calyceal Processes, Which Are Absent from Mice. *J Cell Biol* 2012, 199, 381–399, doi:10.1083/jcb.201202012.
- May-Simera, H.; Nagel-Wolfrum, K.; Wolfrum, U. Cilia - The Sensory Antennae in the Eye. *Prog Retin Eye Res* 2017, 60, 144–180.
- Cowan, C.S.; Renner, M.; De Gennaro, M.; Gross-Scherf, B.; Goldblum, D.; Hou, Y.; Munz, M.; Rodrigues, T.M.; Krol, J.; Szikra, T.; et al. Cell Types of the Human Retina and Its Organoids at Single-Cell Resolution. *Cell* 2020, 182, 1623–1640.e34, doi:10.1016/j.cell.2020.08.013.

21. Adato, A.; Michel, V.; Kikkawa, Y.; Reiners, J.; Alagramam, K.N.; Weil, D.; Yonekawa, H.; Wolfrum, U.; El-Amraoui, A.; Petit, C. Interactions in the Network of Usher Syndrome Type 1 Proteins. *Hum Mol Genet* 2005, 14, 347–356, doi:10.1093/hmg/ddi031.
22. Sorusch, N.; Wunderlich, K.; Bauá, K.; Nagel-Wolfrum, K.; Wolfrum, U. Usher Syndrome Protein Network Functions in the Retina and Their Relation to Other Retinal Ciliopathies. *Adv Exp Med Biol* 2014, doi:10.1007/978-1-4614-3209-8_67.
23. Overlack, N.; Maerker, T.; Latz, M.; Nagel-Wolfrum, K.; Wolfrum, U. SANS (USH1G) Expression in Developing and Mature Mammalian Retina. *Vision Res* 2008, 48, 400–412, doi:10.1016/j.visres.2007.08.021.
24. Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Židek, A.; Bates, R.; Blackwell, S.; Yim, J.; et al. Protein Complex Prediction with AlphaFold-Multimer. *bioRxiv* 2021, 2021.10.04.463034, doi:10.1101/2021.10.04.463034.
25. Emenecker, R.J.; Griffith, D.; Holehouse, A.S. Metapredict: A Fast, Accurate, and Easy-to-Use Predictor of Consensus Disorder and Structure. *Biophys. J.* 2021, 120, 4312–4319, doi:10.1016/j.bpj.2021.08.039.
26. Schindelin, J.; Rueden, C.T.; Hiner, M.C.; Eliceiri, K.W. The ImageJ Ecosystem: An Open Platform for Biomedical Image Analysis. *Mol Reprod Dev* 2015, 82, 518–529, doi:10.1002/mrd.22489.
27. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat Methods* 2022, 19, 679–682, doi:10.1038/s41592-022-01488-1.
28. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J Comput Chem* 2004, 25, 1605–1612, doi:10.1002/jcc.20084.
29. Nunez-Castilla, J.; Siltberg-Liberles, J. An Easy Protocol for Evolutionary Analysis of Intrinsically Disordered Proteins. In *Methods mol biol*; Humana Press Inc., 2020; Vol. 2141, pp. 147–177.
30. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2-A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* 2009, 25, 1189–1191, doi:10.1093/bioinformatics/btp033.
31. Positteam Download RStudio - Posit Available online: <https://posit.co/downloads/> (accessed on 24 July 2023).
32. Karpova, T.S.; Baumann, C.T.; He, L.; Wu, X.; Grammer, A.; Lipsky, P.; Hager, G.L.; McNally, J.G. Fluorescence Resonance Energy Transfer from Cyan to Yellow Fluorescent Protein Detected by Acceptor Photobleaching Using Confocal Microscopy and a Single Laser. *J Microsc* 2003, 209, 56–70, doi:10.1046/j.1365-2818.2003.01100.X.
33. Yan, J.; Pan, L.; Chen, X.; Wu, L.; Zhang, M. The Structure of the Harmonin/sans Complex Reveals an Unexpected Interaction Mode of the Two Usher Syndrome Proteins. *Proc Natl Acad Sci U S A* 2010, 107, 4040–4045, doi:10.1073/pnas.0911385107.
34. Liu, S.; Li, P.; Dybkov, O.; Nottrott, S.; Hartmuth, K.; Lührmann, R.; Carlomagno, T.; Wahl, M.C. Binding of the Human Prp31 Nop Domain to a Composite RNA-Protein Platform in U4 SnRNP. *Science* 2007, 316, 115–120, doi:10.1126/science.1137924.
35. Charenton, C.; Wilkinson, M.E.; Nagai, K. Mechanism of 5' Splice Site Transfer for Human Spliceosome Activation. *Science* 2019, 364, 362–367, doi:10.1126/science.aax3289.
36. Colcombet-Cazenave, B.; Druart, K.; Bonnet, C.; Petit, C.; Sperandio, O.; Guglielmini, J.; Wolff, N. Phylogenetic Analysis of Harmonin Homology Domains. *BMC Bioinformatics* 2021, 22, 190, doi:10.1186/s12859-021-04116-5.
37. Alderson, T.R.; Pritišanac, I.; Kolarić, E.; Moses, A.M.; Forman-Kay, J.D. Systematic Identification of Conditionally Folded Intrinsically Disordered Regions by AlphaFold2. *bioRxiv* 2023, 2022.02.18.481080, doi:10.1101/2022.02.18.481080.
38. Nguyen Ba, A.N.; Yeh, B.J.; Van Dyk, D.; Davidson, A.R.; Andrews, B.J.; Weiss, E.L.; Moses, A.M. Proteome-Wide Discovery of Evolutionary Conserved Sequences in Disordered Regions. *Sci Signal* 2012, 5, doi:10.1126/scisignal.2002515.
39. Holehouse, A.S.; Das, R.K.; Ahad, J.N.; Richardson, M.O.G.; Pappu, R. V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* 2017, 112, 16–21, doi:10.1016/j.bpj.2016.11.3200.
40. Das, R.K.; Ruff, K.M.; Pappu, R. V. Relating Sequence Encoded Information to Form and Function of Intrinsically Disordered Proteins. *Curr Opin Struct Biol* 2015, 32, 102–112.
41. Sheinerman, F.B.; Norel, R.; Honig, B. Electrostatic Aspects of Protein-Protein Interactions. *Curr Opin Struct Biol* 2000, 10, 153–159.
42. Hevekerl, H.; Spielmann, T.; Chmyrov, A.; Widengren, J. Förster Resonance Energy Transfer beyond 10 Nm: Exploiting the Triplet State Kinetics of Organic Fluorophores. *J Phys Chem* 2011, 115, 13360–13370, doi:10.1021/jp206770s.
43. Roszik, J.; Tóth, G.; Szöllosi, J.; Vereb, G. Validating Pharmacological Disruption of Protein-Protein Interactions by Acceptor Photobleaching FRET Imaging. *Methods mol biol* 2013, 986, 165–178, doi:10.1007/978-1-62703-311-4_11.

44. Tancredi, T.; Carrà, G.; Guerrini, R.; Arduin, M.; Calò, G.; Regoli, D.; Salvadori, S.; Temussi, P.A. The Interaction of Highly Helical Structural Mutants with the NOP Receptor Discloses the Role of the Address Domain of Nociceptin/Orphanin FQ. *Chem Eur J* 2005, 11, 2061–2070, doi:10.1002/chem.200401095.
45. Ghram, M.; Morris, G.; Culjkovic-Kraljacic, B.; Mars, J.; Gendron, P.; Skrabanek, L.; Revuelta, M.V.; Cerchietti, L.; Guzman, M.L.; Borden, K.L.B. The Eukaryotic Translation Initiation Factor EIF4E Reprograms Alternative Splicing. *EMBO J* 2023, 42, doi:10.15252/embj.2021110496.
46. Bertram, K.; Agafonov, D.E.; Dybkov, O.; Haselbach, D.; Leelaram, M.N.; Will, C.L.; Urlaub, H.; Kastner, B.; Lührmann, R.; Stark, H. Cryo-EM Structure of a Pre-Catalytic Human Spliceosome Primed for Activation. *Cell* 2017, 170, 701–713.e11, doi:10.1016/j.cell.2017.07.011.
47. Novotny, I.; Blazikova, M.; Stanek, D.; Herman, P.; Malinsky, J. In Vivo Kinetics of U4/U6.U5 Tri-SnRNP Formation in Cajal Bodies. *Mol Biol Cell* 2011, 22, 513–523, doi:10.1091/mbc.E10-07-0560.
48. Stanek, D.; Pridalova-Hnilicova, J.; Novotny, I.; Huranova, M.; Blazikova, M.; Wen, X.; Sapra, A.K.; Neugebauer, K.M. Spliceosomal Small Nuclear Ribonucleoprotein Particles Repeatedly Cycle through Cajal Bodies. *Mol Biol Cell* 2008, 19, 2534–2543, doi:10.1091/mbc.E07-12-1259.
49. Fokkema, I.F.A.C.; Taschner, P.E.M.; Schaafsma, G.C.P.; Celli, J.; Laros, J.F.J.; den Dunnen, J.T. LOVD v.2.0: The next Generation in Gene Variant Databases. *Hum Mutat* 2011, 32, 557–563, doi:10.1002/humu.21438.
50. Nagel-Wolfrum, K.; Fadl, B.R.; Becker, M.M.; Wunderlich, K.A.; Schäfer, J.; Sturm, D.; Fritze, J.; Gür, B.; Kaplan, L.; Andreani, T.; et al. Expression and Subcellular Localization of USH1C/Harmonin in Human Retina Provides Insights into Pathomechanisms and Therapy. *Hum Mol Genet* 2023, 32, 431–449, doi:10.1093/hmg/ddac211.
51. Islam, Z.; Nagampalli, R.S.K.; Fatima, M.T.; Ashraf, G.M. New Paradigm in Ankyrin Repeats: Beyond Protein-Protein Interaction Module. *Int J Biol Macromol* 2018, 109, 1164–1173, doi:10.1016/j.ijbiomac.2017.11.101.
52. Tanackovic, G.; Ransijn, A.; Ayuso, C.; Harper, S.; Berson, E.L.; Rivolta, C. A Missense Mutation in PRPF6 Causes Impairment of Pre-mRNA Splicing and Autosomal-Dominant Retinitis Pigmentosa. *Am. J. Hum. Genet.* 2011, 88, 643–649, doi:10.1016/j.ajhg.2011.04.008.
53. Buskin, A.; Zhu, L.; Chichagova, V.; Basu, B.; Mozaffari-Jovin, S.; Dolan, D.; Droop, A.; Collin, J.; Bronstein, R.; Mehrotra, S.; et al. Disrupted Alternative Splicing for Genes Implicated in Splicing and Ciliogenesis Causes PRPF31 Retinitis Pigmentosa. *Nat Commun* 2018, 9, doi:10.1038/s41467-018-06448-y.
54. Stenson, P.D.; Ball, E. V.; Mort, M.; Phillips, A.D.; Shiel, J.A.; Thomas, N.S.T.; Abeyasinghe, S.; Krawczak, M.; Cooper, D.N. Human Gene Mutation Database (HGMD®): 2003 Update. *Hum Mutat* 2003, 21, 577–581.
55. Zelinger, L.; Swaroop, A. RNA Biology in Retinal Development and Disease. *Trends Genet* 2018, 34, 341–351, doi:10.1016/j.tig.2018.01.002.
56. Aisa-Marin, I.; Garcia-Arroyo, R.; Mirra, S.; Marfany, G. The Alter Retina: Alternative Splicing of Retinal Genes in Health and Disease. *Int J Mol Sci* 2021, 22, doi:10.3390/ijms22041855.
57. Murphy, D.; Cieply, B.; Carstens, R.; Ramamurthy, V.; Stoilov, P. The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina. *PLoS Genet* 2016, 12, e1006256, doi:10.1371/journal.pgen.1006256.
58. Sundar, J.; Matalkah, F.; Jeong, B.; Stoilov, P.; Ramamurthy, V. The Musashi Proteins MSI1 and MSI2 Are Required for Photoreceptor Morphogenesis and Vision in Mice. *J Biol Chem* 2020, 296, doi:10.1074/jbc.ra120.015714.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.