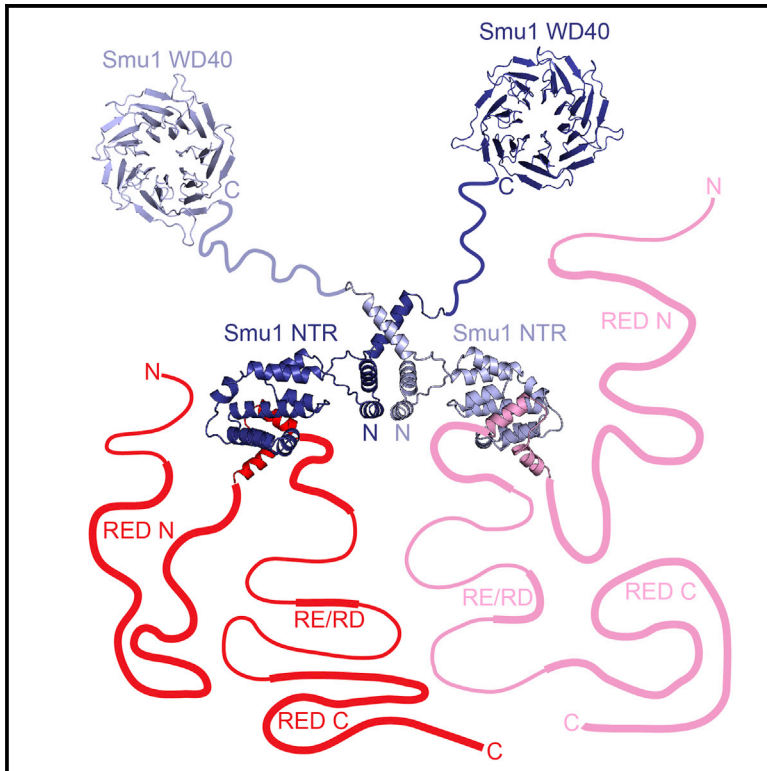# Structural Basis for the Functional Coupling of the Alternative Splicing Factors Smu1 and RED

Graphical Abstract

Authors

Alexander K.C. Ulrich, Jana F. Schulz, Antje Kamprad, Tonio Schütze, Markus C. Wahl

Correspondence

alexander.ulrich@fu-berlin.de (A.K.C.U.), mwahl@zedat.fu-berlin.de (M.C.W.)

In Brief

Smu1 and RED regulate splicing and mitosis. Systematic truncation of RED allowed Ulrich et al. to obtain the structure of the hetero-tetrameric Smu1-RED complex. Smu1 and RED form a functional module that displays multiple protein-protein interaction sites and exhibits intriguing similarities to transcriptional co-repressor complexes.

Highlights

- Crystal structures of dimeric Smu1 and a hetero-tetrameric Smu1-RED complex

- Smu1 dimerizes via a hydrophobic interface involving its LisH motif

- An N-terminal region of Smu1 interacts with short, central helices of RED

- Smu1-RED complex is structurally similar to transcriptional co-repressor complexes

Accession Numbers

5EN6
5EN7
5EN8

CellPress

Structural Basis for the Functional Coupling of the Alternative Splicing Factors Smu1 and RED

Alexander K.C. Ulrich,[1,*] Jana F. Schulz,[1,2] Antje Kamprad,[1,3] Tonio Schütze,[1,4] and Markus C. Wahl[1,*]

[1]Laboratory of Structural Biochemistry, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany
[2]Present address: Department of Pharmacology and Therapeutics, McGill University, 3655 Promenade Sir-William-Osler, Montréal, QC H3G 1Y6, Canada
[3]Present address: Department of Structural Systems Biology, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany
[4]Present address: Sinsheimer Laboratories, Department of Molecular, Cellular and Developmental Biology, Center for Molecular Biology of RNA, University of California Santa Cruz, Santa Cruz, CA 95064, USA
*Correspondence: alexander.ulrich@fu-berlin.de (A.K.C.U.), mwahl@zedat.fu-berlin.de (M.C.W.)
http://dx.doi.org/10.1016/j.str.2016.03.016

SUMMARY

The proteins Smu1 and RED have been jointly implicated in the regulation of alternative splicing, mitosis, and influenza virus infection, but how they interact and whether their diverse cellular functions are coupled is unknown. We identified an N-terminal region of Smu1 and a central region of RED that stably interact. Structural analyses revealed that the RED-binding region of Smu1 contains an N-terminal LisH motif linked to a core domain and a C-terminal α helix that folds back onto the LisH motif. Smu1 dimerizes via its LisH motif and C-terminal α helix and undergoes global conformational changes upon RED binding. In the ensuing hetero-tetrameric Smu1-RED complex, two molecules of RED use short α helices to bind hydrophobic grooves of two Smu1 core domains. Our results show how Smu1 and RED form a functional module that exhibits intriguing similarities to transcriptional co-repressor complexes, arranging multiple additional protein-protein interaction sites for contacting splicing and/or chromatin factors.

INTRODUCTION

Precursor mRNA (pre-mRNA) splicing is a mandatory step in the expression of most eukaryotic primary protein-coding genes, during which non-coding intervening sequences (introns) are removed and neighboring coding regions (exons) are concatenated. Most primary transcripts in higher eukaryotes contain more than one intron and can undergo alternative splicing (Wang et al., 2008), which gives rise to multiple mature mRNAs originating from the same gene. Splicing is catalyzed by a highly dynamic RNA-protein (RNP) molecular machinery, the spliceosome, which is assembled anew for each round of splicing from five small nuclear (sn) RNPs and numerous non-snRNP proteins (Brow, 2002; Wahl et al., 2009). During the initial stages of spliceosome assembly, the U1 and U2 snRNPs recognize the 5′ splice site and branchpoint region of an intron, respectively,

forming the A complex. Subsequently, the U4, U5, and U6 snRNPs join as a pre-formed tri-snRNP, giving rise to the pre-catalytic B complex. The B complex is then catalytically activated in multiple steps, yielding first the B[act] and then the B* complex. The latter can carry out the first of two transesterification reactions of a splicing event. The ensuing C complex catalyzes the second transesterification reaction, after which the spliceosome is disassembled and its subunits are recycled.

Transitions between functional stages in a splicing cycle are characterized by the ordered recruitment and release of specific snRNPs and many non-snRNP proteins (Brow, 2002; Wahl et al., 2009). For example, in higher eukaryotes, a group of nine non-snRNP proteins (Prp38, Snu23, MFAP1, Smu1, RED, NPW38, NPW38BP, UBL5, and FBP21) joins the spliceosome at the stage of B complex formation and is released again during the subsequent catalytic activation phase (Agafonov et al., 2011). All of these B-specific proteins, except Prp38, Snu23, and UBL5, are restricted to higher eukaryotes, in which alternative splicing is pervasive, suggesting that they may be important splicing regulators. Indeed, several of these proteins have been directly implicated in the regulation of splice site choice and alternative splicing (Chung et al., 2009; Ma et al., 2012; Mishra et al., 2011; Spartz et al., 2004; Spike et al., 2001; Sugaya et al., 2006, 2011).

Among the B-specific proteins, Smu1 (suppressor of *mec-8* and *unc-52* 1) and RED (named after a region rich in arginine (R)/glutamic acid (E) or arginine/aspartic acid (D) repeats (Assier et al., 1999) and also called Smu2 or cytokine IK), have also been jointly implicated in other cellular processes. Smu1 supports mitotic spindle integrity (Neumann et al., 2010; Rines et al., 2008), associates with chromatin (Ren et al., 2013), and has been attributed a role in genome stability (Paulsen et al., 2009; Ren et al., 2013; Sugaya et al., 2005, 2006). RED regulates mitotic kinases and phosphatases (Lee et al., 2014) and interacts with the spindle assembly checkpoint protein MAD1, supporting kinetochore localization of MAD1 (Yeh et al., 2012). Moreover, during influenza virus infection, the viral RNA polymerase recruits Smu1 and RED by direct binding of RED (Fournier et al., 2014).

The above observations suggest that Smu1 and RED form a functional module that acts in several cellular contexts. Consistently, the two proteins have been found to interact in pull-down (Chung et al., 2009; Spartz et al., 2004), yeast 2-hybrid (Y2H)
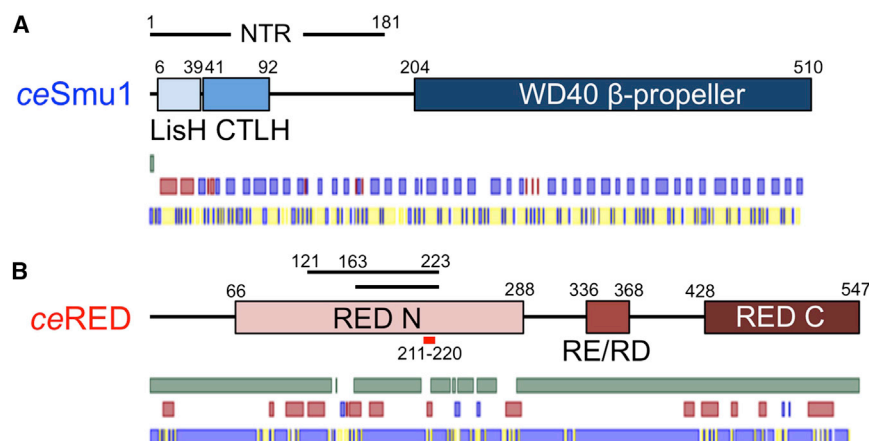
A

B

Figure 1. Domain Organization of *ce*Smu1 and *ce*RED

(A) *ce*Smu1.

(B) *ce*RED. Domains/regions are indicated by rectangles with numbers indicating borders. LisH, lissencephaly type-1-like homology motif; CTLH, C-terminal to LisH motif; RED N, RED-like protein N-terminal region; RED C, RED-like protein C-terminal region. Lines above the schemes indicate fragments used in this study. A short red line below RED N indicates the highly conserved Smu1-binding region of RED. Green bars, regions of predicted structural disorder; blue and red bars, predicted β strands and α helices, respectively; blue/yellow bar, predicted solvent exposed/buried regions.

([Hegele et al., 2012](#)), and luciferase complementation assays ([Fournier et al., 2014](#)), and their interaction stabilizes both proteins ([Fournier et al., 2014; Spartz et al., 2004](#)). Presently, further analysis of the coordinated functions of Smu1 and RED is hampered by lack of structural information on either of the proteins alone or in complex with each other. Here, we have experimentally delineated regions of Smu1 and RED that mediate their stable interaction and determined crystal structures of a truncated Smu1-RED complex and of the RED-binding N-terminal region (NTR) of Smu1 alone. The Smu1 NTR dimerizes and each protomer presents a binding site for RED. We verified the relevance of all interfaces by structure-guided, site-directed mutagenesis in combination with interaction assays.

RESULTS

Experimental Definition of Interacting Regions of Smu1 and RED

For biochemical and structural analyses, we resorted to the Smu1 and RED orthologs of *Chaenorhabditis elegans* (*ce*) ([Figure 1](#)), as we could produce the full-length proteins from this organism in soluble form in *Escherichia coli*. *Ce*Smu1 and *ce*RED are highly homologous to their human counterparts (62% and 38% sequence identity, respectively), except for some divergence in the equivalent of the RE/RD repeat region of *ce*RED, which lacks glutamic acid residues and contains RD and RS repeats instead.

After mixing and incubation, recombinantly produced full-length *ce*Smu1 and *ce*RED co-migrated on an analytical size-exclusion column and eluted earlier than either of the proteins alone ([Figure 2](#)A), indicating that they form a stable complex in vitro. However, the full-length *ce*Smu1-*ce*RED complex failed to crystallize, presumably due to high conformational flexibility. Sequence analyses suggested that *ce*Smu1 contains N-terminal lissencephaly type 1-like homology (LisH) and C-terminal-to-LisH (CTLH) motifs connected via a linker to a C-terminal WD40 domain ([Figure 1](#)A). The vast majority of residues in *ce*RED are predicted to reside in regions that lack a stable 3D fold (93.2%) and to be highly solvent exposed (81.2%), suggesting that RED belongs to the group of intrinsically disordered proteins ([Figure 1](#)B).

Treatment of *ce*Smu1 with chymotrypsin gave rise to a protease-resistant fragment that contained residues 186–508 based on mass spectrometric fingerprinting. Using this information and secondary structure predictions, we dissected the gene encoding *ce*Smu1 into an NTR (residues 1–181), which covered the LisH-CTLH motifs, and a C-terminal region (residues 181–510), which contained the WD40 domain, and produced the corresponding fragments in *E. coli*. Only *ce*Smu1^{NTR} stably bound *ce*RED during size-exclusion chromatography (SEC; [Figure 2](#)B). Thrombin treatment of an assembled *ce*Smu1^{FL}-*ce*RED^{FL} complex gave rise to a series of *ce*RED fragments, one of which contained residues 121–333 as characterized by mass spectrometry and still co-migrated with *ce*Smu1^{FL} and *ce*Smu1^{NTR} in SEC. We generated a series of further RED truncations based on secondary structure predictions and found that residues 163–223 of *ce*RED retained stable binding to *ce*Smu1^{NTR} ([Figure 2](#)C). *Ce*RED$^{163-223}$ contains a well-conserved region encompassing residues 210–221, with 50% sequence identity to human RED. Surface plasmon resonance (SPR) analysis revealed a dissociation constant (K_d) of 5.6 ± 1.5 μM for the *ce*Smu1^{NTR}-*ce*RED$^{163-223}$ interaction ([Figure 2](#)D).

Structural Basis of the Smu1-RED Interaction

To reveal the molecular basis of the *ce*Smu1-*ce*RED interaction, we produced *ce*Smu1^{NTR} and *ce*RED$^{163-223}$ separately in *E. coli* and assembled a complex from the purified components, which gave rise to crystals that diffracted to 2.94 Å resolution and belonged to space group P2$_1$ ([Table 1](#)). Selenomethionine (SeMet)-derivatized complex crystallized in a different space group, C222$_1$, and crystals diffracted to 3.10 Å resolution ([Table 1](#)). Thus, we first solved the structure of the SeMet-derivatized complex using a multiple anomalous diffraction strategy. After model building and refinement, the native structure was solved by molecular replacement.

Crystals of the SeMet-derivatized complex contained two *ce*Smu1^{NTR} and two *ce*RED$^{163-223}$ chains in an asymmetric unit, while an asymmetric unit of the native complex crystals contained four copies of each protein. All copies of *ce*Smu1^{NTR} adopt a very similar fold composed of 11 α helices (pairwise root-mean-square deviation [rmsd] of 0.22–0.78 Å for 126–161 equivalent Cα positions; [Figure 3](#)A). The N-terminal two α helices
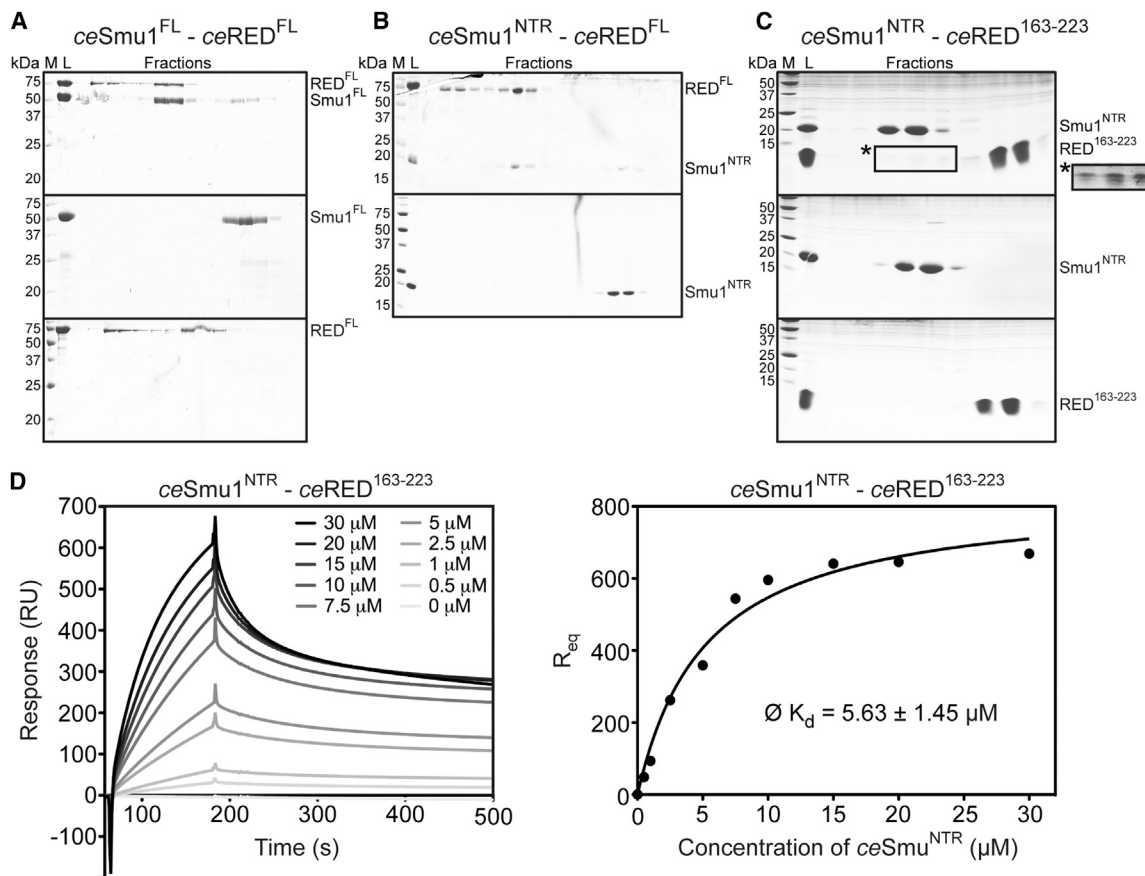
Figure 2. Experimental Definition of a Minimal *ce*Smu1-*ce*RED Complex

(A) SDS-PAGE gels of SEC experiments of *ce*Smu1^{FL} with *ce*RED^{FL} (top), isolated *ce*Smu1^{FL} (middle), and isolated *ce*RED^{FL} (bottom). A fraction of isolated *ce*RED^{FL} runs in the exclusion volume, likely due to formation of soluble aggregates.

(B) SDS-PAGE gels of SEC experiments of *ce*Smu1^{NTR} with *ce*RED^{FL} (top) and isolated *ce*Smu1^{NTR} (bottom).

(C) SDS-PAGE gels of SEC experiments of *ce*Smu1^{NTR} with *ce*RED$^{163-223}$ (top), isolated *ce*Smu1^{NTR} (middle), and isolated *ce*RED$^{163-223}$ (bottom). As a large excess of *ce*RED$^{163-223}$ was used in the analysis of the complex, only a small fraction of total *ce*RED$^{163-223}$ co-eluted with *ce*Smu1^{NTR} and stained weakly. The region of the gel containing the *ce*RED$^{163-223}$ fraction co-eluting with *ce*Smu1^{NTR} (black rectangle and asterisk) is again shown on the right with higher contrast.

(D) SPR sensorgrams (left) and K$_d$ determination (right) of the interaction between His$_6$-tagged *ce*RED$^{163-223}$ immobilized on an Ni^{2+}-loaded sensor chip NTA (GE Healthcare) and *ce*Smu1^{NTR} (analyte). Graphs shown are from one representative set of experiments. Ø K$_d$, mean ± SD of the K$_d$ values of four independent experiments.

α1 and α2 (residues 8–35) run antiparallel to each other and are part of the LisH motif. They are connected through a nine-residue linker to a globular α-helical core (GAC) domain (residues 45–152) comprising helices α3-α10. Helix α5 forms a central rod within the GAC domain, around which the other seven helices are arranged. Following the GAC domain, a seven-residue linker leads back to the LisH motif. The C terminus of *ce*Smu1^{NTR} is formed by helix α11 that towers almost perpendicularly on top of helix α1 of the LisH motif.

The two crystallographically independent *ce*Smu1^{NTR} molecules in the SeMet complex each form a dimer with an identical, crystallographically related copy, while the four *ce*Smu1^{NTR} molecules of the native structure form two fully crystallographically independent dimers. These four dimers exhibit very similar head-to-head arrangements (pairwise rmsd of 0.43–3.61 Å for 291–342 equivalent Cα positions; Figures 3B and 3C). The proteins dimerize via a mainly hydrophobic interface created by the LisH motif helices α1 and α2 and the C-terminal helix α11 (Figures

3B and 3C). The dimer interface is formed by three layers of pairs of equivalent α helices on the two monomers (α1-α1′, α2-α2′, and α11-α11′; primes indicate elements of the second monomer of a dimer; Figures 3C and 4A). The layers formed by helices α1-α1′ and α2-α2′ are stacked on top of each other forming an antiparallel four-helix bundle with a hydrophobic core and hydrophilic interactions at the periphery (Figure 4A). The positions of helices α11/α11′ on top of helices α1/α1′ are stabilized by hydrophobic interactions (Figure 4A). Helices α11 and α11′ cross each other at an angle of about 70° and engage in direct, reciprocal interactions along their entire lengths (Figure 4A).

All four *ce*Smu1^{NTR} dimers in the SeMet-derivatized and native complex crystals form hetero-tetramers, in which each *ce*Smu1^{NTR} chain is associated with one molecule of *ce*RED$^{163-223}$ (Figures 3B and 3C). In each crystallographically independent *ce*Smu1^{NTR}-*ce*RED$^{163-223}$ hetero-dimer, the highly conserved *ce*RED$^{163-223}$ residues 210–221 form a C-terminal α helix that lies along a groove formed by helices

Table 1. Crystallographic Data and Refinement Statistics

Dataset	ceSmu1^{NTR}-ceRED$^{163-223}$ (SeMet) Peak Data	Inflection Point Data	High Remote Data	ceSmu1^{NTR}-ceRED$^{163-223}$ (Native)	ceSmu1^{NTR}
PDB ID	5EN6			5EN7	5EN8
Data collection					
Wavelength (Å)	0.979531	0.979771	0.977427	0.918	0.976
Space group	C222$_1$	C222$_1$	C222$_1$	P2$_1$	C2
Unit cell parameters					
a, b, c (Å)	166.5, 183.1, 41.0	166.7, 183.5, 41.1	166.9, 184.0, 41.1	122.2, 41.0, 126.3	160.4, 35.2, 79.3
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 95.9, 90.0	90.0, 93.3, 90.0
Resolution (Å)[a]	45.77–3.16 (3.35–3.16)	45.87–3.10 (3.28–3.10)	45.99–3.18 (3.38–3.18)	46.16–2.94 (3.11–2.94)	40.04–2.22 (2.36–2.22)
Unique reflections[a]	20,766 (3,265)[b]	22,137 (3,493)[b]	20,426 (3,227)[b]	26,994 (3,995)	21,482 (3,198)
Completeness (%)[a]	99.6 (98.0)	99.7 (98.5)	99.7 (98.1)	98.3 (92.6)	95.9 (89.1)
Redundancy[a]	6.0 (5.7)	6.0 (5.6)	6.2 (6.0)	3.3 (3.5)	3.2 (3.2)
R_{meas} (%)[a,c]	9.2 (87.0)	10.5 (130.9)	9.5 (115.7)	14.8 (101.8)	5.0 (90.9)
$<I/\sigma(I)>$[a]	16.1 (2.2)	14.6 (1.4)	15.4 (1.7)	10.6 (1.4)	11.1 (1.1)
CC(1/2)[a]	99.9 (85.1)	99.9 (74.9)	99.9 (80.2)	99.5 (65.8)	99.8 (85.2)
Wilson B factor (Å^2)	102.6	102.6	102.6	62.7	55.3
Phasing					
No. of sites	4	4	4		
P (acentric)[d]					
Isomorphous	0.000	0.402	0.366		
Anomalous	1.015	0.456	0.428		
R_{Cullis} (acentric)[e]					
Isomorphous	0.000	0.760	0.735		
Anomalous	0.822	0.953	0.964		
FOM (acentric)[f]		0.281			
Refinement					
Unique reflections		11,881		26,984	21,192
Resolution (Å)[a]		45.86–3.10 (3.42–3.10)		46.16–2.94 (3.04–2.94)	40.04–2.23 (2.33–2.23)
R_{work} (%)[a,g]		28.3 (36.6)		22.2 (34.7)	22.9 (47.7)
R_{free} (%)[a,h]		32.1 (39.0)		26.9 (39.0)	26.0 (53.6)
No. atoms protein		3,130		6,320	2,757
No. atoms water		56		141	92
B factor (Å^2) protein		111.6		68.0	77.8
B factor (Å^2) water		91.9		53.7	76.2
Rmsd[i]					
Bond lengths (Å)		0.003		0.005	0.003
Bond angles (°)		0.722		0.898	0.656
Ramachandran plot					
Favored (%)		99.2		98.6	99.1
Allowed (%)		0.8		1.3	0.9
Outliers (%)		0.0		0.1	0.0
Rotamer outliers (%)		2.0		2.3	1.3

[a]Data for highest resolution shell in parentheses.

[b]Friedel pairs were treated as separate reflections.

[c]$R_{meas} = \Sigma_h[n/(n-1)]^{1/2}\Sigma_i|I_h - I_{h,i}|/\Sigma_h\Sigma_i I_{h,i}$, in which I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.

[d]P = phasing power = $\Sigma_n|F_{H,calc}|/\Sigma_n|E| = |F_{PH,obs}| - |F_{PH,calc}|$ = mean lack of closure error; n, number of observed scattering amplitudes for the derivate; $F_{PH,obs}$, $F_{PH,calc}$, observed and calculated structure factor amplitudes of the derivative; $F_{H,calc}$, calculated structure factor amplitude of the heavy atom substructure.

[e]$R_{Cullis} = \Sigma_{hkl}||F_{PH} \pm F_P| - F_{H,calc}|/\Sigma_{hkl}|F_{PH} \pm F_P|$; F_{PH}, F_P, observed structure factor amplitudes of the derivative, native; $F_{H,calc}$, calculated structure factor amplitudes of the heavy atom substructure; + if signs of F_{PH} and F_P are equal, − if opposite.
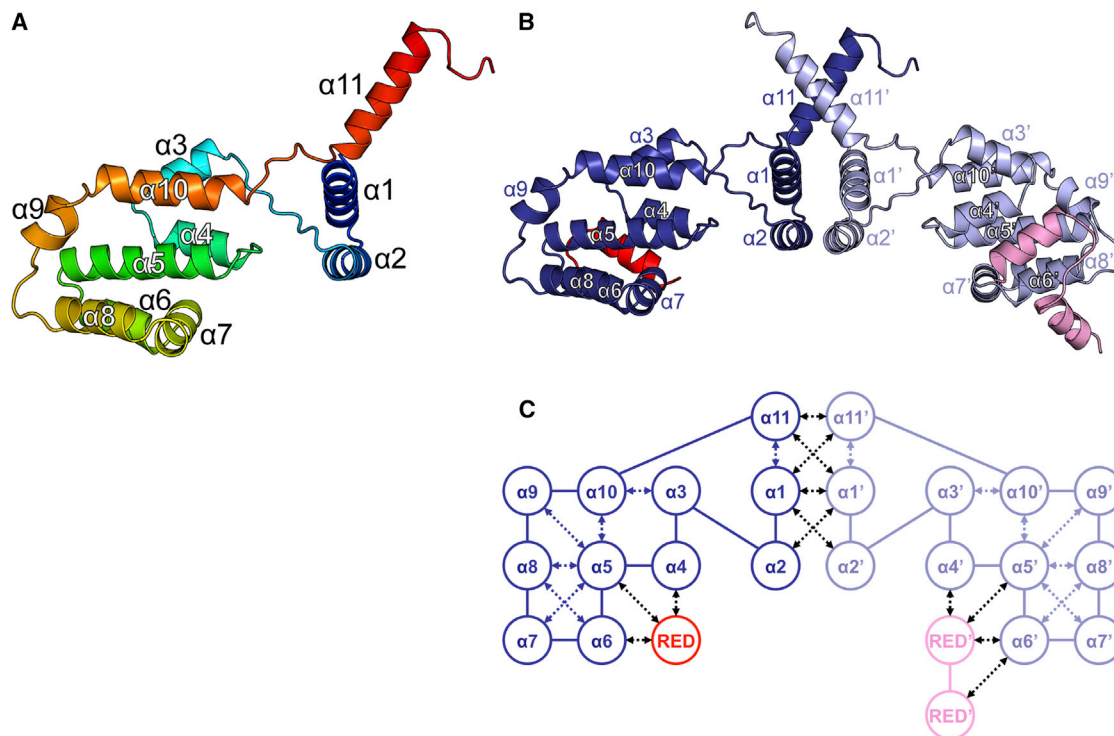
Figure 3. Crystal Structure of the Hetero-Tetrameric *ce*Smu1-*ce*RED Complex

(A) Cartoon plot of a *ce*Smu1[NTR] monomer colored blue to red from the N to the C terminus with α helices labeled.

(B) Cartoon plot of a *ce*Smu1[NTR]-*ce*RED[163−223] hetero-tetramer. *Ce*Smu1[NTR] is depicted in dark and light blue, *ce*RED[163−223] is depicted in red and pink. The left *ce*Smu1[NTR] is shown in the same view as in (A).

(C) Structural arrangement of α helices (circles) in the *ce*Smu1[NTR]-*ce*RED[163−223] hetero-tetramer. Solid lines represent covalent inter-helix connections and dashed arrows represent non-covalent inter-helix connections. Coloring as in (B).

α4, α5, and α6 of the *ce*Smu1[NTR] GAC domain (Figures 3B, 3C, and 5A). A preceding loop of *ce*RED[163−223] (residues 205–209) lies across the *ce*Smu1 α5-α6 loop and the N-terminal part of helix α6 (Figure 5A). In one case, residues 194–204 of *ce*RED[163−223] form an additional N-terminal α helix that further encircles *ce*Smu1[NTR] helix α6 (Figure 5A).

Smu1 Forms LisH Motif-Based Dimers in Solution

The head-to-head mode of *ce*Smu1[NTR] dimerization seen in *ce*RED[163−223]-complexed structures is the only self-interaction between *ce*Smu1[NTR] molecules that is observed for all crystallographically independent *ce*Smu1[NTR] copies in our crystals. An analysis of all *ce*Smu1[NTR]-*ce*Smu1[NTR] crystal contacts via the PISA server (Krissinel and Henrick, 2007) suggested that only this dimerization mode involving the LisH motifs is physiologically relevant (probability of getting a lower than obtained solvation free energy gain purely based on hydrophobicity, $P(\Delta G_i)$, 0.048; maximal fraction of the total free energy of binding that belongs to the interface in stable assemblies, CSS, 1.0; free energy of assembly dissociation, ΔG^{diss}, 12.7 kcal/mol). To experimentally assess the oligomerization state of *ce*Smu1[NTR] and of

the full-length protein in solution, we carried out analytical SEC coupled to multi-angle light scattering (MALS). For full-length *ce*Smu1 (calculated monomeric molecular mass 57.44 kDa) we determined a molecular mass of 113.8 ± 1.4 kDa, very close to the calculated molecular mass of a dimer (114.88 kDa; Figure 4B). Likewise, for His6-tagged *ce*Smu1[NTR] (calculated monomeric molecular mass 23.68 kDa), we obtained a molecular mass of 46.48 ± 0.73 kDa, which closely matched the calculated molecular mass of a dimer (47.37 kDa; Figure 4C).

To further test if dimerization in solution also occurs via the LisH motif, we recombinantly produced *ce*Smu1[NTR] variants bearing residue exchanges within the LisH motif, which based on our structures should interfere with stable dimerization. Wild-type (wt) *ce*Smu1[NTR] and variants were produced, purified without cleavage of the His6 tags and subjected to analytical SEC. Compared with the wt protein (elution peak at 1.15 ml), variants with disrupted dimerization potential should elute at larger volumes (Figure 4D). I15 and F18 are positioned in the center of the LisH-LisH′ interface (Figure 4A). A I15A exchange led to only partial disruption or weakening of the dimerization (main elution peak at 1.21 ml, secondary peak at 1.15 ml). Residue exchanges
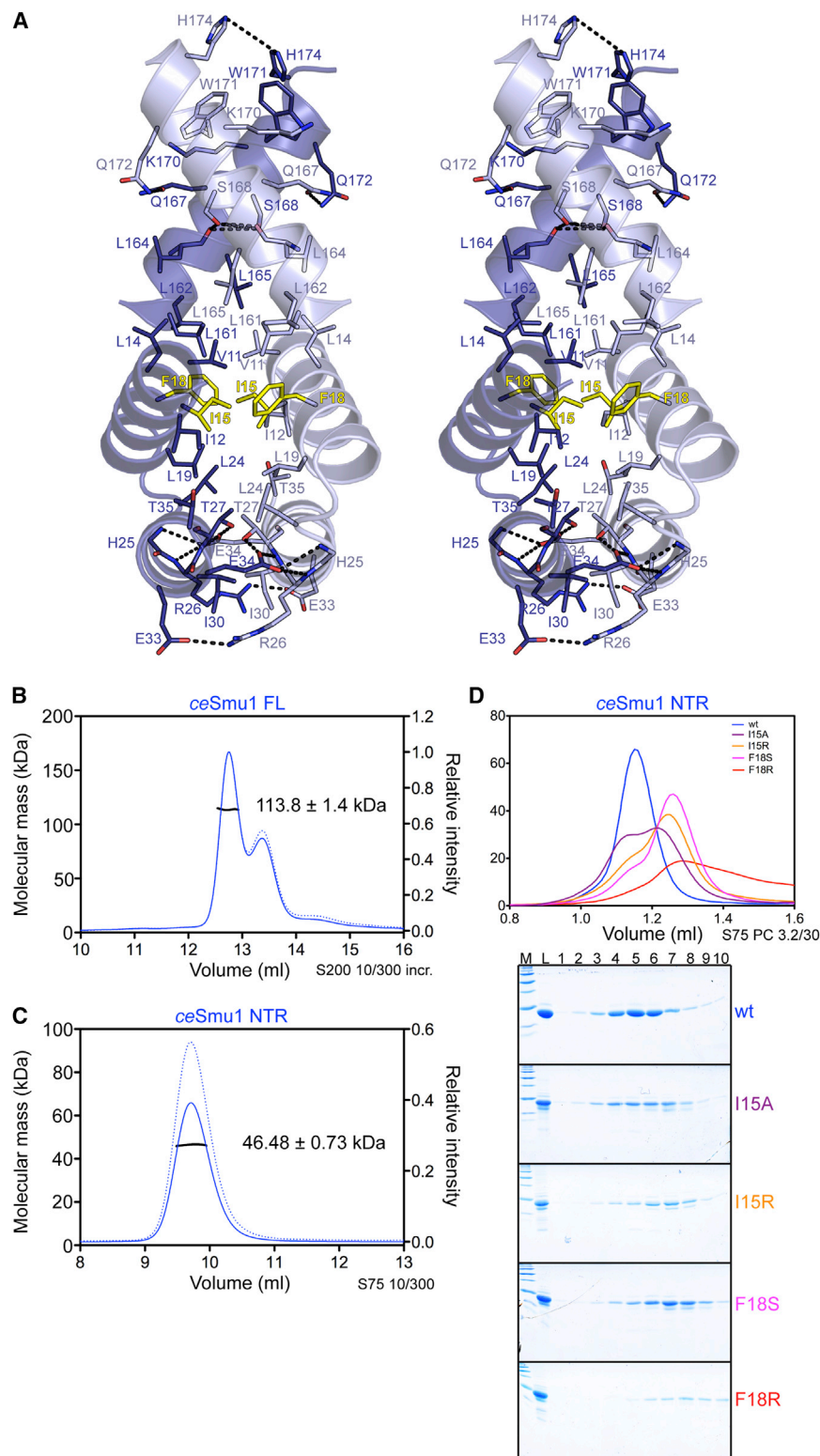
Figure 4. ceSmu1^{NTR} Forms a Dimer in Solution Mediated by Its LisH Motif

(A) Stereo of α helices forming the ceSmu1^{NTR}-dimerization interface. Residues involved in direct contacts between the monomers are shown as sticks, residues exchanged to probe dimerization are highlighted in yellow.

(B and C) SEC/MALS analysis of ceSmu1^{FL} (B) and His$_6$-tagged ceSmu1^{NTR} (C). Columns used are indicated on the lower right of the chromatograms. Solid, blue curve, UV$_{280nm}$ trace; dashed, blue curve, refractive index trace; black curve, molecular mass at the corresponding elution volumes. The averaged molecular masses across the main peaks are indicated, the errors represent three SDs of the individual molecular mass measurements.

(D) Elution profiles (UV$_{280nm}$; top) and SDS-PAGE analyses of SEC runs (bottom) with the indicated ceSmu1^{NTR} variants.

the F18R variant compared with the other two variants may be due to a weak interaction with the size-exclusion resin and/or due to misfolding of the LisH motif region. Together, these results show that ceSmu1 forms dimers in solution and that dimerization is mediated by head-to-head interaction of its LisH motif and additional contacts provided by helix α11 as seen in our crystal structures.

We also analyzed the oligomeric state of ceRED fragments by SEC-MALS. The ceRED$^{163−223}$ construct employed in crystallization was monomeric in solution (9.91 ± 0.59 kDa; calculated monomeric molecular mass of 9.95 kDa; Figure S1A). SEC-MALS analysis had to be carried out at 18°C, whereby larger fragments, such as ceRED$^{121−223}$, eluted in two peaks. The first peak at the void volume (presumably containing unspecific aggregates) trailed into the second peak at larger volumes, preventing accurate molar mass determination of the sample in the latter peak. Thus, we cannot rule out that larger ceRED fragments oligomerize.

Mutational Analysis of the ceSmu1-ceRED Interaction

To investigate whether the ceSmu1^{NTR}-ceRED$^{163−223}$ interface seen in our crystal structures also mediates complex formation in solution, we performed analytical SEC experiments with wt ceSmu1^{NTR} and ceRED$^{121−223}$ as well as variants, in which we exchanged contact-mediating amino acid residues. For these experiments, we employed a larger ceRED construct than used for co-crystallization, which was easier to monitor by SDS-PAGE. Although we could not determine the

I15R (elution peak at 1.25 ml), F18S (1.26 ml), and F18R (1.28 ml), by which a polar residue was introduced into the hydrophobic center of the LisH-LisH′ dimerization interface, disrupted the dimers completely (Figure 4D). The slightly larger elution volume of

exact oligomerization state of ceRED[121–223] (see above), the wt fragment and its variants did not show aggregation at 4°C and displayed specific, differential effects with ceSmu1[NTR] in the interaction tests. Wt ceSmu1[NTR]-ceRED[121–223] complex eluted at 1.05 ml from our analytical SEC column, ceSmu1[NTR] alone eluted at 1.15 ml and ceRED[121–223] eluted at 1.22 ml (Figure 5B). In our crystal structures, L96 of ceSmu1[NTR] is engaged in hydrophobic interactions with F220 of ceRED[163–223] (Figure 5A). All combinations involving protein variants, in which these residues were individually exchanged with arginines, showed broad elution peaks at larger volumes than the wt complex (ceSmu1[NTR,L96R] plus ceRED[121–223], 1.12 ml; ceSmu1[NTR] plus ceRED[121–223,F220R], 1.14 ml; ceSmu1[NTR,L96R] plus ceRED[121–223,F220R], 1.15 ml), suggesting disintegration of the complex (Figure 5C). SDS-PAGE analysis confirmed that the elution peaks were made up of separately migrating ceSmu1[NTR] and ceRED[121–223] variants, whose elution profiles overlapped (Figure 5C). The same result was obtained when we converted ceRED residue V212, which engages in hydrophobic interactions with ceSmu1 L85, E90, and T93 at the center of the ceSmu1[NTR]-ceRED[163–223] interface (Figure 5A), to an arginine (Figure 5C). All variants individually eluted like the respective wt protein (Figure 5B). Thus, these results indicate that the interface observed in the crystallographic ceSmu1[NTR]-ceRED[163–223] complexes is also used for complex formation in solution.

Conformational Changes upon Complex Formation

To investigate whether binding of RED is accompanied by any conformational changes in ceSmu1[NTR], we determined a crystal structure of ceSmu1[NTR] alone at 2.23 Å resolution (Table 1). An asymmetric unit of these crystals contains two molecules of ceSmu1[NTR]. Structure solution by molecular replacement required a search with the separated ceSmu1 LisH and GAC domains, already indicating that these elements are arranged differently in the structure of the isolated protein compared with the structures in complex with ceRED[163–223]. Individually, the LisH and GAC domains adopt very similar folds in isolated and ceRED[163–223]-bound ceSmu1[NTR] (pairwise rmsd of 0.17–0.23 Å for 24–27 equivalent Cα positions for the LisH domain and 0.33–0.66 Å for 92–98 equivalent Cα positions for the GAC domain). Thus, binding of ceRED[163–223] does not induce significant conformational changes in the ceSmu1 LisH or GAC domains.

The two ceSmu1[NTR] monomers in an asymmetric unit engage in very similar head-to-head contacts between the two LisH domains as seen in the structures in complex with ceRED[163–223] (pairwise rmsd of 0.23–0.38 Å for 72–80 equivalent Cα positions; Figure 6A). However compared with isolated ceSmu1[NTR], the GAC domains of ceRED[163–223]-bound ceSmu1[NTR] are repositioned relative to the LisH dimerization region such that ceRED[163–223]-binding surfaces on the GAC domains are rotated away from the LisH dimerization region (Figure 6A). A ceRED[163–223] peptide as seen in the ceSmu1[NTR]-ceRED[163–223] complex structures could be modeled onto the GAC domains of isolated ceSmu1[NTR] without steric conflict. However, the 31 residues contained in the peptide but not traceable in our electron densities obviously add steric bulk that might favor an alternative conformation in the ceRED[163–223]-bound complex, in which steric hindrance is minimized.

To assess conformational changes in ceRED[163–223] upon complex formation, we analyzed the fragment via circular dichro-

ism (CD) spectroscopy (Figure S1B). The CD spectrum of ceRED[163–223] at 20°C revealed a major negative peak below 200 nm, a signature of unstructured proteins, and a small negative shoulder around 220 nm, which could indicate a small percentage of α helix (ca. 16% calculated with the program CDNN; Bohm et al., 1992). However, no significant differences were seen in the CD spectrum at 90°C (ca. 19% calculated α helix content), where the protein is expected to be fully unfolded. Visible α-helical portions of ceRED[163–223] in the Smu1[NTR]-ceRED[163–223] complex crystal structures already amount to 26% of the residues present in ceRED[163–223]. These observations suggest that ceRED[163–223] is unstructured in isolation and that its ceSmu1[NTR]-interacting α helices are induced upon complex formation.

## Smu1 Is Structurally Related to the TOPLESS Family of Transcriptional Co-repressors

A structure similarity search with the Dali server (Holm and Rosenstrom, 2010) revealed that ceSmu1[NTR] is highly similar to the recently published structure of the NTR of plant TOPLESS-related protein 2 (TPR2; Z score 11.8; rmsd 4.7 Å for 151 aligned Cα-atoms of 206 total Cα atoms; 15% sequence identity), a member of the TOPLESS (TPL) and TOPLESS-related (TPR) families of transcriptional co-repressors (PDB: 4ZHE; Ke et al., 2015). TPR2[NTR] and ceSmu1[NTR] structures share the LisH motif, the GAC domain, and the terminal α helix. In both proteins, the NTRs are also predicted to be followed by one (ceSmu1) or two (TPR2) WD40 domains. TPR2 additionally harbors a C-terminal zinc finger that is lacking in Smu1 orthologs.

Like ceSmu1[NTR], TPR2[NTR] utilizes an LisH motif in concert with a terminal helix for homotypic dimerization (Figure 6B). The dimerization regions in ceSmu1[NTR] and TRP2[NTR] are highly similar (rmsd of 0.47–0.55 Å for 74–77 equivalent Cα positions). However, while in isolated and ceRED[163–223]-bound ceSmu1[NTR] the GAC domains extend laterally from the central dimerization region, they are positioned below the LisH region in TPR2[NTR] (Figures 6A and 6B), underscoring the structural flexibility of the linkers between the LisH motif/C-terminal helix and the GAC domains. The domain arrangement in TPR2[NTR] allows two LisH-based TPR2[NTR] dimers to additionally associate via the distal surfaces of their GAC domains (helices equivalent to α6 and α8 of ceSmu1[NTR]), forming a tetramer (Figure 6B). This second TPR2[NTR] oligomerization interface is different from any ceSmu1[NTR] contact seen in our crystal structures.

TPR2 binds transcriptional repressors via ethylene response factor-associated amphiphilic repression (EAR) motifs, and structures of TPR2 in complex with three small EAR motif-containing peptides have also been determined (Ke et al., 2015). Superimposition of the peptide-bound structures of TPR2[NTR] (PDB: 5C7F, 5C7E, and 5C6V) and ceSmu1[NTR] revealed different peptide-binding surfaces on the GAC domains. While ceRED[163–223] binds to helices α4, α5, and α6 of the ceSmu1 GAC domain, EAR motif-containing peptides bind to elements that correspond to ceSmu1[NTR] helices α5, α7, α8, α9, and α10 on the opposite side of the GAC domain of TPR2 (Figure 6C).

DISCUSSION

We have determined the crystal structures of interacting portions of the ceSmu1 and ceRED proteins as well as of the
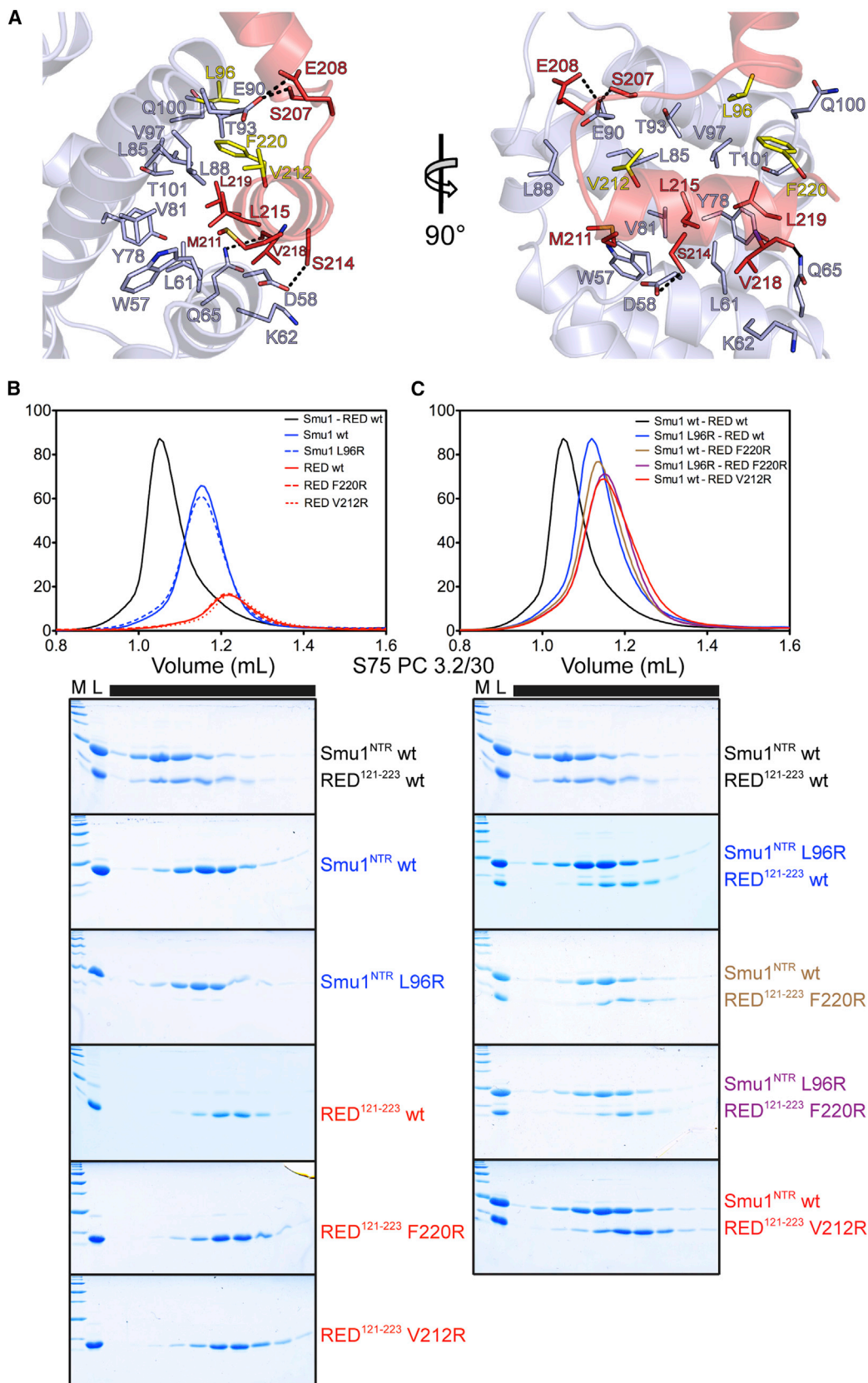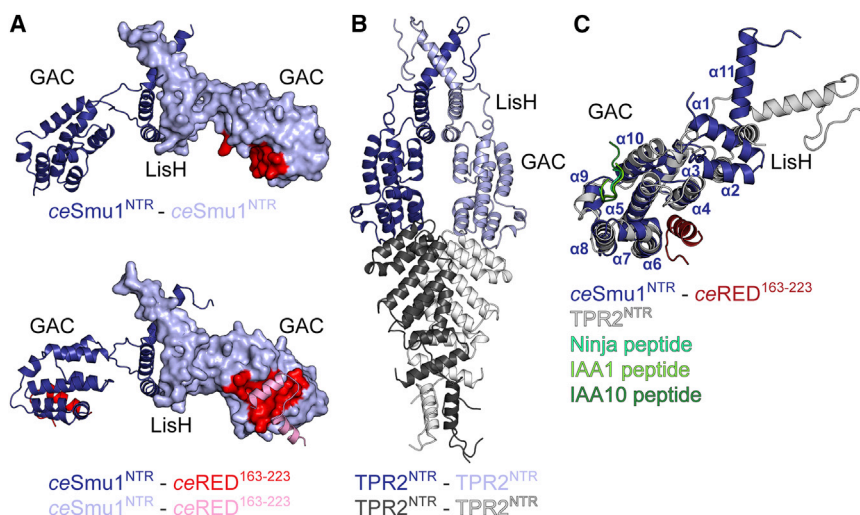
(legend on next page)

Figure 6. Structural Comparisons

(A) Combined cartoon and surface plot of isolated (top) and *ce*RED[163−223]-bound (bottom) *ce*Smu1[NTR] dimer in dark and light blue. The *ce*RED[163−223]-binding surface on *ce*Smu1[NTR] is highlighted in red, emphasizing the conformation change upon complex formation.

(B) Isolated TRP2[NTR] tetramer, with the two *ce*Smu1[NTR]-like dimers in dark/light blue and dark/light gray (PDB: 4ZHE) (Ke et al., 2015).

(C) Superposition of one monomer of *ce*Smu1[NTR] (blue) in *ce*RED[163−223]-bound state and of TPR2[NTR] (gray) in complex with EAR motif-containing peptides (PDB: 5C6V, 5C7F, 5C7E). The RED α helix (in red) binds to GAC domain α helices 4, 5, and 6, whereas the EAR motif-containing peptides Ninja, IAA1 and IAA10 (in different shades of green) bind to the opposite site of the GAC domain (α helices 5, 7, 8, 9, and 10). View rotated by −54° about the y axis and by −27° about the x axis relative to (A).

*ce*RED-binding portion of *ce*Smu1 alone. Apart from the atomic details underlying the *ce*Smu1-*ce*RED interaction, our results suggest that the *ce*Smu1 LisH motif-based dimerization region and the GAC domains are repositioned upon *ce*RED binding, and that the *ce*Smu1-binding fragment of *ce*RED undergoes a folding-upon-binding transition during complex formation. By structure-guided mutational analyses and interaction studies, we showed that the observed modes of *ce*Smu1 dimerization and the binding of the *ce*RED fragment via the GAC domain also occur in solution. Due to the high conservation of Smu1 and of the Smu1-binding region of RED, the observed structures and modes of interaction are expected to be conserved in other eukaryotic organisms that harbor these two proteins. Furthermore, comparison of our structures and those of other LisH motif-containing proteins shows a conserved mode of dimerization, suggesting that LisH motifs generally represent protein dimerization elements.

A Smu1-RED Complex as a Multi-Protein Interaction Module

Both Smu1 and RED harbor long regions in addition to the elements contained in our crystal structures. Full-length Smu1 encompasses a C-terminal WD40 domain connected by a 28-residue linker to its NTR. Human Smu1 showed Y2H interactions with several other spliceosomal proteins (hLIN1, LSm4, LSm8, MFAP1, RED, YB-1) (Hegele et al., 2012; Schütze et al., 2016), suggesting that regions beyond the NTR might provide additional protein-protein interaction (PPI) sites, consistent with the WD40 domain as a bona fide PPI element. RED harbors ca. 200 and 300 residues N- and C-terminal, respectively, of the Smu1-binding site, which are predicted to be largely intrinsically disordered. In line with this prediction, no electron density was observed in our *ce*Smu1[NTR]-*ce*RED[163−223] crystal structures

for *ce*RED residues that were not directly involved in Smu1 binding. Intrinsically disordered proteins often exhibit closely spaced peptide motifs that can serve as protein binding sites (Tompa et al., 2014) with spacer regions serving as dynamic linkers (Kalmar et al., 2012). A similar organization of RED is indicated by our finding of a short peptide motif in *ce*RED required for *ce*Smu1 binding and by previously reported Y2H interactions of human RED with numerous other spliceosomal proteins (CTNNBL1, FRA10AC1, hSYF1, MGC20398, MFAP1, Prp38, Prp6, RACK1, RBM10, RNF114A, Smu1, Snip1, Snu114, THOC1, TOE1, TTC14, ZCCHC10, ZNF830; Hegele et al., 2012; Schütze et al., 2016).

The above observations suggest that a hetero-tetrameric Smu1-RED complex, sustained by the Smu1-RED interaction delineated here, constitutes a multi-protein interaction module that could provide versatile scaffolding functions in the various cellular contexts in which the proteins are implicated. Consistent with the notion that Smu1 and RED function as a stably interacting module, the proteins have been found to regulate alternative splicing of a similar set of pre-mRNAs involved in development, cell death, and survival (Papasaikas et al., 2015). Moreover, the effects of Smu1 or RED knockdown on alternative splicing of 35 test genes had the highest pairwise correlation of all splicing factor pairs (Papasaikas et al., 2015). Our delineation of single point mutations that abrogate or disturb Smu1 dimerization or Smu1-RED complex formation provides useful tools for future tests of the importance of these interactions for Smu1 and RED functions.

Similar Organization of Smu1-RED and Transcriptional Co-repressor Complexes

We suggest that the functional mechanisms of the Smu1-RED complex resemble those of TPL and TPR transcriptional

Figure 5. Molecular Details of the *ce*Smu1[NTR]-*ce*RED[163−223] Interaction

(A) Orthogonal, close-up views of the *ce*Smu1[NTR]-*ce*RED[163−223] interaction region. Residues relevant for the interaction are shown as sticks, residues exchanged to interfere with complex formation are colored in yellow. Dashed, black lines, hydrogen bonds. The left panel is rotated by 138° about the z axis, by 88° about the y axis, and by 5° about the x axis compared with Figure 3A.

(B and C) Elution profiles (UV[280nm]; top) and SDS-PAGE analyses of SEC runs (bottom) with the indicated *ce*Smu1[NTR] and *ce*RED[121−223] variants.

co-repressor complexes, of which closely related families exist in plants, fungi, insects, and mammals. Co-repressor complexes consist of scaffolding components, such as TPL/TPR, chromatin-modifying enzymes, and transcriptional repressor proteins that anchor the complexes to specific gene regulatory elements (Buscarlet and Stifani, 2007; Schoch and Abel, 2014). The architecture of Smu1 is very similar to that of plant TPR2. Furthermore, both proteins bind short peptide motifs in certain interaction partners via their GAC domains. TPR2 uses its GAC domain to interact with EAR motifs of the transcriptional repressors NINJA and auxin-responsive proteins IAA1 and IAA10. Like RED, the EAR motif-containing proteins are predicted to be intrinsically disordered. EAR motif-containing adaptor proteins tether TPL/TPR co-repressors to chromatin-bound transcription factor complexes to repress gene expression (Pauwels et al., 2010; Peer, 2013). Likewise, RED might act as an adaptor that recruits splicing-regulatory Smu1-bound complexes to the spliceosome at the B complex stage (Agafonov et al., 2011). In a similar fashion, RED might recruit Smu1 and associated proteins to mitosis-regulating kinases and phosphatases (Lee et al., 2014), to chromatin-associated complexes (Ren et al., 2013), or to influenza virus RNA polymerase (Fournier et al., 2014).

As Smu1 has been found in nuclear speckles (Sugaya et al., 2011), which are enriched in pre-mRNA splicing factors but low in chromatin (Spector, 2001), as well as in chromatin fractions (Ren et al., 2013), Smu1-RED complexes may exist in two different nuclear pools associated with separate cellular processes. On the other hand, speckles often localize to sites of active transcription (Spector, 2001), and there is by now ample evidence that pre-mRNA splicing is functionally coupled to chromatin organization (Iannone and Valcarcel, 2013). Therefore, the involvement of Smu1 and RED in splicing and chromatin-related functions together with our structural analyses would be in line with the intriguing, yet speculative, possibility that splicing-associated and chromatin-associated roles of Smu1-RED complexes are functionally integrated. For example, the multiplication of protein binding sites by the observed Smu1 dimerization would in principle allow the hetero-tetrameric Smu1-RED complex to engage splicing factors via some of its PPI sites and chromatin-associated factors via others. Alternatively, Smu1-RED complexes could cycle between speckle- and chromatin-associated states in an ordered fashion or RED could hand over Smu1 and Smu1-interacting factors from one functional context to the other.

EXPERIMENTAL PROCEDURES

Protein Sequence Analyses
The RD/RE region of ceRED was assigned manually. The WD40 domain of ceSmu1 was predicted by the WD40 repeat protein Structure Predictor (Wang et al., 2013). All other regions of the proteins (Figure 1) were predicted by the NCBI conserved domain database (Marchler-Bauer et al., 2015). Regions of structural disorder were predicted with Meta-Disorder (Schlessinger and Rost, 2005). Programs of the PredictProtein package (Rost et al., 2004) were used for secondary structure analyses (REPROFSec) and solvent exposed or buried regions (PROFAcc).

Plasmids
Synthetic open reading frames encoding full-length ceSmu1 and ceRED in pETM11 plasmids were ordered from GeneArt (Thermo Fisher Scientific). The pETM11 vector guides production of N-terminally His$_6$-tagged, tobacco

etch virus (TEV)-cleavable fusion proteins. Mutations leading to truncations and residue exchanges were introduced by inverse PCR as described (Ulrich et al., 2012). All expression constructs were verified by sequencing.

Protein Production and Purification
Proteins were produced in *E. coli* Rosetta 2 (DE3) or BL21 (DE3) RIL cells by auto-induction in ZY medium or, for expression of SeMet-containing proteins, in auto-inducing medium for SeMet labeling (Studier, 2005) at 18°C for 24–40 hr. Cells were harvested by centrifugation. The following steps were performed at 4°C. Cells were resuspended in solubilization buffer (50 mM sodium phosphate [pH 8.0], 500 mM NaCl, 30 mM imidazole, 5 mM β-mercaptoethanol) and disrupted using a Sonopuls Ultrasonic Homogenizer HD 3100 (Bandelin). The soluble fractions were incubated with Ni^{2+}-NTA resin (GE Healthcare) for 1 hr, washed three times with ten column volumes of solubilization buffer, and bound protein was eluted with 250 mM imidazole (pH 8.0), 300 mM NaCl, 5 mM β-mercaptoethanol. For crystallographic studies, the His$_6$ tag was cleaved with TEV protease (mass ratio 1:30) during overnight dialysis against 10 mM sodium phosphate (pH 8.0), 300 mM NaCl, 30 mM imidazole, 5 mM β-mercaptoethanol, and cleaved samples were again passed over Ni^{2+}-NTA resin. The flow-through was collected, concentrated, and further purified by SEC in SEC buffer (10 mM Tris-HCl [pH 8.0], 300 mM NaCl, 0.1 mM EDTA, 1 mM DTT).

For crystallization of the ceSmu1^{NTR}-ceRED$^{163-223}$ complex, proteins were individually purified by Ni^{2+}-NTA affinity chromatography. After cleavage of the ceSmu1^{NTR} His$_6$ tag, ceSmu1^{NTR} and His$_6$-ceRED$^{163-223}$ were combined, passed over a 5 ml His-Trap FF crude column (GE Healthcare) and eluted with elution buffer. The eluate was collected, the His$_6$ tag on ceRED$^{163-223}$ was cleaved, and the sample was again passed over a 5 ml His-Trap FF crude column to remove uncleaved protein. The flow-through was collected, pooled, and further purified by SEC in SEC buffer. Peak fractions were pooled, concentrated and shock-frozen in liquid nitrogen.

Limited Proteolysis
Protein or protein complex preparations were diluted in SEC buffer to 300 μg/ml. 18 μl of protein/complex and 2 μl of different proteases at varying concentrations were incubated for 30 min at room temperature. Reactions were stopped by addition of 5 μl of 5× SDS loading buffer, boiled for 5 min, and analyzed by SDS-PAGE.

Analytical Size-Exclusion Chromatography
50 μM protein alone or with an equimolar amount of binding partner were incubated in SEC buffer for 30 min at 4°C. 50 μl of the samples were analyzed on a Superdex 75 PC 3.2/30 size-exclusion column (GE Healthcare) at 4°C using an ÄKTAmicro system (GE Healthcare). The peak fractions were inspected by SDS-PAGE.

Multi-Angle Light Scattering
MALS experiments were performed at 18°C. Samples were passed over a Superdex 200 increase 10/300 (ceSmu1^{FL}) or a Superdex 75 10/300 size-exclusion column (all other samples) coupled to a miniDAWN TREOS three-angle light scattering detector (Wyatt Technology) and a RefractoMax520 refractive index detector (ERC). Detectors were aligned, corrected for band broadening, and photodiodes were normalized with BSA as a reference. For calculation of the molecular mass, protein concentrations were determined from the differential refractive index with a specific refractive index increment (dn/dc) of 0.185 ml/g. Data were analyzed with ASTRA 6.1.4.25 (Wyatt Technology).

CD Spectroscopy
CeRED$^{163-223}$ was dialyzed against 10 mM sodium phosphate (pH 8.0), 200 mM sodium perchlorate at 4°C overnight, and diluted to a final concentration of 19.5 μM. Spectra between 190 and 260 nm were recorded at 20°C and 90°C with a Jasco J-810 spectropolarimeter using a quartz cuvette with 0.1 mm path length.

Surface Plasmon Resonance Analysis
SPR experiments were carried out at 25°C, using a Biacore 3000 instrument (GE Healthcare) with 50 mM HEPES (pH 8.0), 300 mM NaCl, 0.5 mg/ml BSA,

0.005% NP-40 as running buffer. Ni^{2+} was loaded at 190–210 response units (RUs) on two channels of a four-channel sensor chip NTA (GE Healthcare) according to the supplier's manual. His$_6$-tagged ceRED$^{163-223}$ was subsequently immobilized at 110–160 RUs. CeSmu1^{NTR} at concentrations between 0 and 80 μM was injected to all four channels for 2 min at a flow rate of 20 μl/min. Subsequently, the analyte was allowed to dissociate for 15 min at a flow rate of 20 μl/min. The signals of nonspecific binding to the reference channels were subtracted from the signals of the ceRED$^{163-223}$ channels. K$_d$ values were obtained by fitting a steady-state model to the observed equilibrium RU values (R$_{eq}$).

Crystallographic Procedures

Crystals of native and SeMet-containing ceSmu1^{NTR}-ceRED$^{163-223}$ were obtained by vapor diffusion at 18°C with drops containing 1 μl of complex at 28 mg/ml and 1 μl of reservoir solution (0.1 M HEPES [pH 7.5], 0.02 M MgCl$_2$, 22% [w/v] polyacrylic acid 5,100, or 0.1 M HEPES [pH 7.0], 0.02 M MgCl$_2$, 19% [w/v] polyacrylic acid 5,100, respectively). Crystals of ceSmu1^{NTR} were obtained by vapor diffusion at 4°C with drops containing 0.1 μl of protein at 23 mg/ml and 0.1 μl of reservoir solution (0.1 M MES [pH 6.5], 12% [w/v] PEG 20,000).

For diffraction data collection, crystals were transferred into reservoir solution supplemented with 20% (v/v) ethylene glycol (native ceSmu1^{NTR}-ceRED$^{163-223}$), 22% (v/v) glycerol (SeMet-containing ceSmu1^{NTR}-ceRED$^{163-223}$) or 25% (v/v) glycerol (ceSmu1^{NTR}) and flash-cooled in liquid nitrogen. Diffraction data of native and SeMet-containing ceSmu1^{NTR}-ceRED$^{163-223}$ crystals were collected at beamline 14.1 of the BESSY II storage ring (Berlin, Germany) (Mueller et al., 2012). Diffraction data of ceSmu1^{NTR} were collected at beamline P14 of the PETRA III storage ring (Hamburg, Germany). All data were processed by XDS (Kabsch, 2010).

Initial phases for ceSmu1^{NTR}-ceRED$^{163-223}$ were obtained by the multiple anomalous dispersion strategy using data collected at the selenium peak wavelength, the inflection point wavelength, and a high-energy remote wavelength (Table 1). Four selenium sites were located and used for phasing with autoSHARP (Vonrhein et al., 2007). The experimentally phased map was of high quality (Figure S2) and allowed manual building of the structure with COOT (Emsley and Cowtan, 2004). The structures of native ceSmu1^{NTR}-ceRED$^{163-223}$ and native ceSmu1^{NTR} were solved by molecular replacement using MOLREP (Vagin and Teplyakov, 2010) from the CCP4 package (Collaborative Computational Project, Number 4, 1994) with the structure coordinates of the refined SeMet-containing ceSmu1^{NTR}-ceRED$^{163-223}$ model. All structural models were completed through alternating rounds of automated refinement using phenix.refine (Afonine et al., 2012) and manual model building using COOT.

ACCESSION NUMBERS

Structure coordinates and diffraction data have been deposited with the PDB (http://www.pdb.org) under accession codes PDB: 5EN6 (SeMet ceSmu1^{NTR}-ceRED$^{163-223}$), 5EN7 (native ceSmu1^{NTR}-ceRED$^{163-223}$), and 5EN8 (ceSmu1^{NTR}).

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at http://dx.doi.org/10.1016/j.str.2016.03.016.

AUTHOR CONTRIBUTIONS

All authors designed experiments, analyzed the data, and interpreted the results. A.K.C.U., J.S., A.K., and T.S. performed experiments. A.K.C.U. and M.C.W. wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367.

Agafonov, D.E., Deckert, J., Wolf, E., Odenwalder, P., Bessonov, S., Will, C.L., Urlaub, H., and Luhrmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. Mol. Cell Biol. 31, 2667–2682.

Assier, E., Bouzinba-Segard, H., Stolzenberg, M.C., Stephens, R., Bardos, J., Freemont, P., Charron, D., Trowsdale, J., and Rich, T. (1999). Isolation, sequencing and expression of RED, a novel human gene encoding an acidic-basic dipeptide repeat. Gene 230, 145–154.

Bohm, G., Muhr, R., and Jaenicke, R. (1992). Quantitative analysis of protein far UV circular dichroism spectra by neural networks. Protein Eng. 5, 191–195.

Brow, D.A. (2002). Allosteric cascade of spliceosome activation. Annu. Rev. Genet. 36, 333–360.

Buscarlet, M., and Stifani, S. (2007). The 'Marx' of Groucho on development and disease. Trends Cell Biol. 17, 353–361.

Chung, T., Wang, D., Kim, C.S., Yadegari, R., and Larkins, B.A. (2009). Plant SMU-1 and SMU-2 homologues regulate pre-mRNA splicing and multiple aspects of development. Plant Physiol. 151, 1498–1512.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Fournier, G., Chiang, C., Munier, S., Tomoiu, A., Demeret, C., Vidalain, P.O., Jacob, Y., and Naffakh, N. (2014). Recruitment of RED-SMU1 complex by Influenza A Virus RNA polymerase to control viral mRNA splicing. PLoS Pathog. 10, e1004164.

Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Luhrmann, R., and Stelzl, U. (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell 45, 567–580.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549.

Iannone, C., and Valcarcel, J. (2013). Chromatin's thread to alternative splicing regulation. Chromosoma 122, 465–474.

Kabsch, W. (2010). Xds. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132.

Kalmar, L., Acs, V., Silhavy, D., and Tompa, P. (2012). Long-range interactions in nonsense-mediated mRNA decay are mediated by intrinsically disordered protein regions. J. Mol. Biol. 424, 125–131.

Ke, J., Ma, H., Gu, X., Thelen, A., Brunzelle, J.S., Li, J., Xu, H.E., and Melcher, K. (2015). Structural basis for recognition of diverse transcriptional repressors by the TOPLESS family of corepressors. Sci. Adv. 1, e1500107.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Lee, S., Han, S., Jeong, A.L., Park, J.S., and Yang, Y. (2014). Depletion of IK causes mitotic arrest through aberrant regulation of mitotic kinases and phosphatases. FEBS Lett. *588*, 2844–2850.

Ma, L., Gao, X., Luo, J., Huang, L., Teng, Y., and Horvitz, H.R. (2012). The *Caenorhabditis elegans* gene mfap-1 encodes a nuclear protein that affects alternative splicing. PLoS Genet. *8*, e1002827.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., et al. (2015). CDD: NCBI's conserved domain database. Nucleic Acids Res. *43*, D222–D226.

Mishra, S.K., Ammon, T., Popowicz, G.M., Krajewski, M., Nagel, R.J., Ares, M., Jr., Holak, T.A., and Jentsch, S. (2011). Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature *474*, 173–178.

Mueller, U., Darowski, N., Fuchs, M.R., Forster, R., Hellmig, M., Paithankar, K.S., Puhringer, S., Steffien, M., Zocher, G., and Weiss, M.S. (2012). Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. J. Synchrotron Radiat. *19*, 442–449.

Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., et al. (2010). Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature *464*, 721–727.

Papasaikas, P., Tejedor, J.R., Vigevani, L., and Valcarcel, J. (2015). Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. Mol. Cell *57*, 7–22.

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., et al. (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell *35*, 228–239.

Pauwels, L., Barbero, G.F., Geerinck, J., Tilleman, S., Grunewald, W., Perez, A.C., Chico, J.M., Bossche, R.V., Sewell, J., Gil, E., et al. (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature *464*, 788–791.

Peer, W.A. (2013). From perception to attenuation: auxin signalling and responses. Curr. Opin. Plant Biol. *16*, 561–568.

Ren, L., Liu, Y., Guo, L., Wang, H., Ma, L., Zeng, M., Shao, X., Yang, C., Tang, Y., Wang, L., et al. (2013). Loss of Smu1 function de-represses DNA replication and over-activates ATR-dependent replication checkpoint. Biochem. Biophys. Res. Commun. *436*, 192–198.

Rines, D.R., Gomez-Ferreria, M.A., Zhou, Y., DeJesus, P., Grob, S., Batalov, S., Labow, M., Huesken, D., Mickanin, C., Hall, J., et al. (2008). Whole genome functional analysis identifies novel components required for mitotic spindle integrity in human cells. Genome Biol. *9*, R44.

Rost, B., Yachdav, G., and Liu, J. (2004). The PredictProtein server. Nucleic Acids Res. *32*, W321–W326.

Schlessinger, A., and Rost, B. (2005). Protein flexibility and rigidity predicted from sequence. Proteins *61*, 115–126.

Schoch, H., and Abel, T. (2014). Transcriptional co-repressors and memory storage. Neuropharmacology *80*, 53–60.

Schütze, T., Ulrich, A.K., Apelt, L., Will, C.L., Bartlick, N., Seeger, M., Weber, G., Luhrmann, R., Stelzl, U., and Wahl, M.C. (2016). Multiple protein-protein interactions converging on the Prp38 protein during activation of the human spliceosome. RNA *22*, 265–277.

Spartz, A.K., Herman, R.K., and Shaw, J.E. (2004). SMU-2 and SMU-1, *Caenorhabditis elegans* homologs of mammalian spliceosome-associated proteins RED and fSAP57, work together to affect splice site choice. Mol. Cell Biol. *24*, 6811–6823.

Spector, D.L. (2001). Nuclear domains. J. Cell Sci. *114*, 2891–2893.

Spike, C.A., Shaw, J.E., and Herman, R.K. (2001). Analysis of smu-1, a gene that regulates the alternative splicing of unc-52 pre-mRNA in *Caenorhabditis elegans*. Mol. Cell Biol. *21*, 4985–4995.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. *41*, 207–234.

Sugaya, K., Hongo, E., and Tsuji, H. (2005). A temperature-sensitive mutation in the WD repeat-containing protein Smu1 is related to maintenance of chromosome integrity. Exp. Cell Res. *306*, 242–251.

Sugaya, K., Hongo, E., Ishihara, Y., and Tsuji, H. (2006). The conserved role of Smu1 in splicing is characterized in its mammalian temperature-sensitive mutant. J. Cell Sci. *119*, 4944–4951.

Sugaya, K., Ishihara, Y., and Sugaya, K. (2011). Enlargement of speckles of SF2/ASF due to loss of function of Smu1 is characterized in the mammalian temperature-sensitive mutant. RNA Biol. *8*, 488–495.

Tompa, P., Davey, N.E., Gibson, T.J., and Babu, M.M. (2014). A million peptide motifs for the molecular biologist. Mol. Cell *55*, 161–169.

Ulrich, A., Andersen, K.R., and Schwartz, T.U. (2012). Exponential megapriming PCR (EMP) cloning–seamless DNA insertion into any target plasmid without sequence constraints. PLoS One *7*, e53360.

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. *66*, 22–25.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. *364*, 215–230.

Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wang, Y., Jiang, F., Zhuo, Z., Wu, X.H., and Wu, Y.D. (2013). A method for WD40 repeat detection and secondary structure prediction. PLoS One *8*, e65705.

Yeh, P.C., Yeh, C.C., Chen, Y.C., and Juang, Y.L. (2012). RED, a spindle pole-associated protein, is required for kinetochore localization of MAD1, mitotic progression, and activation of the spindle assembly checkpoint. J. Biol. Chem. *287*, 11704–11716.