Molecular Cell, Volume *49*

Supplemental Information

Structural Basis of Assembly Chaperone-Mediated snRNP Formation

Clemens Grimm, Ashwin Chari, Jann-Patrick Pelz, Jochen Kuper, Caroline Kisker, Kay Diederichs, Holger Stark, Hermann Schindelin, and Utz Fischer

Supplemental Experimental Procedures

Expression and Purification of recombinant proteins

pICln variants: Expression clones encoding His-tagged *Drosophila melanogaster* pICln fragments were constructed as described in table S3. The respective vectors were transformed into *E. coli* Rosetta 2(DE3)pLysS cells (Novagen, Merck KGaA, Darmstadt, Germany). Transformed cells were grown at 37°C in Terrific Broth (TB) medium containing 25µg/ml Kanamycin and 30µg/ml Chloramphenicol to an OD_{600} of 0.8. Prior to induction with 1mM isopropyl ß-D-thiogalactopyranoside (IPTG), the culture was cooled down to 20°C and grown overnight at 15°C. Cells were resuspended in phosphate buffered saline (PBS) pH 7.5 supplemented with 2mM ß-mercaptoethanol (ß-ME) and lysed by sonication. A cleared lysate was prepared by ultracentrifugation for 45 minutes, 45000 rpm (rotor 45Ti, Beckman Coulter, Krefeld, Germany). The cleared lysate was afterwards incubated with Ni-NTA matrix (Qiagen, Hilden, Germany) for 30 minutes. The matrix was washed twice with 15 bead volumes of a buffer containing 0.3M NaCl, 50mM HEPES pH 7.5 and 2mM ß-ME followed by a wash with 15 bead volumes of a buffer containing 150mM NaCl, 50mM HEPES pH 7.5 and 2mM ß-ME. The proteins were eluted from the Ni-NTA matrix with a buffer containing 150mM NaCl, 50mM HEPES pH 7.5 and 2mM ß-ME and 200mM imidazole. Fractions containing the proteins were concentrated and subsequently further purified by size exclusion gel filtration on a Superdex75 column (GE Healthcare, Munich, Germany). Purified proteins were concentrated to 20mg/ml and stored at -80°C.

SMN/ Gemin2: For bacterial over-expression of *D. melanogaster* SMNΔC and Gemin2, *E. coli* Rosetta 2(DE3)pLysS (Novagen) cells were transformed separately with either plasmid dSMN(1-122):pETM30 or plasmid dGemin2:pETM30, encoding for an N-terminal His6-GST-tag. Cells were grown to an OD of 0.8 and induced overnight at 15°C in TB containing 25µg/ml Kanamycin and 30µg/ml Chloramphenicol. Both cultures were then combined and processed as described above for pICln. The cleared lysate was subjected to sequential affinity purifications on Ni-NTA and glutathione affinity chromatography, respectively. The affinity-tag was cleaved by addition of 1% (w/w) TEV protease. The protease and the cleaved affinity tag were depleted by a second incubation with Ni-NTA. Purification was completed by gel filtration on a Superdex200 (GE Healthcare) column. The protein complex was concentrated to 20mg/ml and stored at -80°C.

Sm protein expression and purification: The SmD1/D2 heterodimer and the SmE/F/G heterotrimer were expressed and purified as described previously (Chari et al., 2008; Kambach et al., 1999).

Reconstitution of the 6S and 8S complexes

Reconstitution of 6S and 8S complexes was performed by mixing the respective subunits in presence of 1 M NaCl and subsequent dialysis to 150 mM NaCl. All buffers contained 5 mM tris(2-carboxyethyl)phosphine (TCEP) and 20mM HEPES-NaOH at pH 7.5. The reconstituted complexes were separated by gel filtration chromatography on a Superdex200 column (GE Healthcare) and fractions were analyzed by SDS-PAGE. Only those fractions containing all subunits of the respective complexes in stoichiometric amounts were pooled and concentrated to 11 mg/ mL (6S complex) or 15 mg/ mL (8S complex).

Crystallogenesis, Data Collection and Structure Determination of 6S

Initially, we obtained two 6S crystal forms diffracting to less than 4 Å. To improve crystal quality, we mutated several amino acids on those parts of the pICln surface that were involved in contact formation within the 8S crystals as soon as this information was available. Several different mutations of surface residues combined with C-terminal truncations as well as different truncations of the loop region in the range of residues 90-125 were screened for crystallization and five more crystal forms were obtained. A 6S complex preparation with dpICln(Δ90-125, Δ160-215, H144A) yielded crystals after two weeks of growth at 20°C (reservoir composition: 15% Jeffamine ED-2003 titrated to pH 7, 10% ethanol; 1ul mixed with 1ul protein preparation within the hanging drop) that diffracted to 1.9 Å on beamline ID14-4 of the ESRF Grenoble. The obtained dataset was processed with XDS and solved by molecular replacement with PHASER / CPP4, using the SmD1/D2/E/F/G/pICln ring from the previously obtained 8S model as search model. The resulting solution, which contained two molecules in the asymmetric unit was subjected to refinement in PHENIX. After an initial manual inspection and correction of the model, three more cycles of automated refinement and manual model building were performed. This included TLS refinement, placement of 372 water molecules and modeling of 20 residues in alternative conformations.

Crystallization, Data Collection and Structure Determination of 8S

Diffraction quality 8S (human Sm proteins, Drosophila SMN(1-122), Drosophila pICln (Δ181-215) and Drosophila full-length Gemin2) crystals were obtained by hanging drop vapor diffusion after 3 weeks of growth at 16°C. Each sample was composed of 6µL protein solution mixed with 2µl of reservoir solution (24% PEG 4000, 10% Ethanol and 150mM NaCl). Crystals diffracted highly anisotropically, up to 3.8 Å resolution in the preferred direction on beamline ID14-4 of the European Synchrotron Radiation Facility, Grenoble, France (ESRF). After screening of approx. 500 crystals, a single crystal was discovered that diffracted up to 3.1 Å. The collected dataset was processed with XDS (Kabsch, 2010). The dataset was subjected to ellipsoidal truncation with resolution limits along a*, b* and c* of 3.1 Å, 3.8 Å and 4.0 Å, respectively, followed by anisotropic scaling (Strong et al., 2006). Useful signal (Karplus and Diederichs, 2012) was detected up to 3.1 Å resolution within the data. The structure was solved by molecular replacement with PHASER (McCoy et al., 2007) using the coordinates of PDB entry 3S6N (Zhang et al., 2011) as a search model. The initial model containing 20 molecules in the asymmetric unit was then refined with PHENIX (Felder et al., 2007) using strong NCS restraints and the resulting electron density was inspected. The missing pICln unit was inserted manually in all 20 molecules using the coordinates from PDB entry 1ZYI (Furst et al., 2005b). Within 8 cycles of manual rebuilding in coot (Huber et al., 1998) including the exchange of the Gemin2, SMN and pICln sequences to those from *D. melanogaster* and the placement of 10 sulphate ions as well as automated refinement in PHENIX the R/ Rfree factors converged. For the final rounds of refinement, the 1.9 Å 6S model was available and used as a reference model for refinement with BUSTER (Gruber and

Pleiss, 2011) However, the NCS and reference model restraints for divergent parts of the structure(s) were released by omitting the previously set "-autoncs_noprune" option. The final model has a R / Rfree-factor of 0.232 / 0.256 and contains residues 2-83 of SmD1, 7-117 of SmD2, 14-90 of SmE, 6-76 of SmF, 5-74 of SmG, 1-89 and 131-168 of pICln, 8-24 of SMN and 30-245 of Gemin2. Unless otherwise noted, the complex molecule comprising chains A, B, C, D, E, F, G and H was used for evaluation and figure generation, as it featured the clearest electron density.

Molecular Graphics and Electrostatic Calculations

All figures were produced using PYMOL (Schrodinger, LLC.). Poisson-Boltzmann calculations were performed with APBS (Baker et al., 2001) using the corresponding PYMOL interface and standard settings (150mM monovalent ions, protein dielectric 2.0, solvent dielectric 78.0, linearized Poisson-Boltzmann equation). The electrostatic potential was then mapped to a PYMOL-generated Connolly surface. The protein dipole moment was determined with the Protein Dipole Moments Server (Felder et al., 2007). Sequence Conservation scores were calculated with the program Consurf (Landau et al., 2005) and mapped on the molecular surface within Pymol.

Normal Mode Analysis

To create the 'open' conformation, all subunits of chain B from the 8S crystal structure were superposed separately on their respective counterparts from structure 3S6N. However, pICln/SmD1 was treated as a single unit. The such generated 'open' model was then submitted to the ElNemo server (Suhre and Sanejouand, 2004) and modes 7-11 with an amplitude of -250 to 250 were calculated.

Molecular Dynamics simulations

Simulations were performed within GROMACS (Van Der Spoel et al., 2005) using the Gromos 43a1 force field and the explicit SPC water model with dodecahedral boundary conditions. The 6S or 8S monomeric complex from the respective crystal structure was placed in a water box and the total charge was brought to zero by replacing an appropriate number of water molecules by sodium or chloride ions, respectively. For a third simulation run, the pICln cover loop was modeled according to the NMR structure by Furst et al. (Furst et al., 2005b) and the pICln C-terminus was manually modeled in an extended conformation until residue 175. Likewise, the SmD2 L2 loop was reconstructed manually in an open conformation. The respective systems were then submitted to an isobaric-isothermal simulation at 1bar and 310 K through a Berendsen thermostat-barostat. Under equilibrium conditions, Debye-Waller factors were calculated from the coordinate differences of 1500 states (corresponding to 3ns simulation time) at the end of the trajectory and mapped to the color of the models depicted in Figure 5C.

Native Gel Electrophoresis of RNA-Protein Complexes

Band shift assays were performed essentially as described (Chari et al., 2008). In brief, 3pmol proteins were added to 0.3pmol radio-labeled RNA in a 10µl reaction, containing 10% glycerol, 0.1U/µl RNAsin (Promega) and 0.1µg/µl tRNA. For the preincubation experiment with pICln, 3 pmol 7S complex were mixed with 3 pmol pICln in the presence of 10% glycerol, 0.1U/µl RNAsin (Promega), 0.1µg/µl tRNA and incubated at 4°C for 20 minutes prior to the addition of 0.3 pmol of radio-labeled RNA. The mixtures were incubated for 15min at 30°C and 45min at 37°C. After incubation the mixtures were briefly centrifuged, supplemented with heparin to 1µg/µl and separated on 6% native polyacrylamide gels

(acrylamide/bisacrylamide ratio 80:1) containing 4% glycerol and 0.5×TBE buffer. Gels were pre-run for 1h at 4°C at a constant current of 15 mA in 1×TBE, and run with samples for 2h at 4°C at a constant current of 35mA. Gels were exposed wet at -80°C.

Electron Microscopy and Single-Particle Image Processing

For EM sample preparation, the 8S complex (Chari et al., 2008) was subjected to the GraFix approach using a 5–20% sucrose gradient (Kastner et al., 2008) and centrifuged in a TH-660 rotor at 40000 rpm for 20 hours at 4°C. Specimens were negatively stained with uranyl formate as previously described (Chari et al., 2008). Images were taken on a Philips CM200 electron microscope (Philips/FEI, Eindhoven, The Netherlands) operated at 160kV using a twofold binned 4kx4k CCD camera (TemCam-F415, TVIPS, Gauting, Germany) at a magnification of 119000 fold (Sander et al., 2005) resulting in a pixel size of 2.5 Å. 10000 single-particle images were selected for the dataset using Boxer from the EMAN package (Ludtke et al., 1999). The particle images were further subjected to CTF correction (Sander et al., 2003a). Using iteratively refined class averages, CTF- corrected single-particle images were aligned via an exhaustive alignment (Sander et al., 2003b) and subsequent multivariate statistical analysis (MSA)-based hierarchical ascendant classification (HAC) in the context of IMAGIC-5 into classes of ≈20 class members in average (van Heel and Frank, 1981). We used a low-pass filtered X-ray structure of the 8S complex to assign initial angles to the class averages by angular reconstitution (Van Heel, 1987). The structure was iteratively refined by several rounds of projection matching using an increasingly finer angular sampling of reprojections used as references during the alignments. The resulting structure had a resolution of 20 Å as determined by the Fourier Shell Correlation 0.5 σ criterion (Harauz et al., 1988).

Supplemental Description to Figure 4A: Molecular Details of the pICln-SmD1 Interface

The canonical, Sm-Sm type interaction of the assembly chaperone with SmD1 is centered on the antiparallel β-strand pair between β4 of SmD1 and β5 of pICln. It involves the formation of five backbone-backbone H-bonds comprising residues SmD1 56-61 and pICln 61-65. In the direct neighborhood of the β-pair, a sixth H-bond is formed between the sidechain hydroxyl group of pICln-Ser66 and the backbone carbonyl of SmD1-Gln54 (Fig 4A). Further interactions involve a hydrophobic cluster located beneath the tapered ring face comprised of Leu62, Tyr74, Met76 and Val133 of pICln and Leu19, Met45, Leu55, Ile58, and Ile60 of SmD1. Located on the flat ring face, pICln's extended C-terminal α-helix is positioned in a related manner as the N-terminal α-helix present in the Sm proteins and therefore also contributes to the interactions of the chaperone with SmD1. However, it remains at a greater distance to the neighboring Sm protein as observed for a canonical Sm-Sm contact (compare Figs 1A and 2A). As a consequence, direct interactions between the C-terminal helix of pICln and SmD1 are relatively sparse and are restricted to H bonds between pICln Tyr150-OH and SmD1 Ser159-OH. In addition, two water-mediated hydrogen bonds are formed between D1-Thr30-OH and pICln Tyr150-OH as well as between D1-Asn37-Nδ2 and the backbone carbonyl of pICln-Ala158. Finally, H-bonds pICln Lys71 – SmD1 Glu56, pICln Val60 (carbonyl O) – SmD1 Arg61 and pICln Ser61/Lys58 (carbonyl O atoms) to SmD1 Asn64 and a salt bridge between pICln Glu135 and SmD1 Arg50 are noteworthy for this contact.

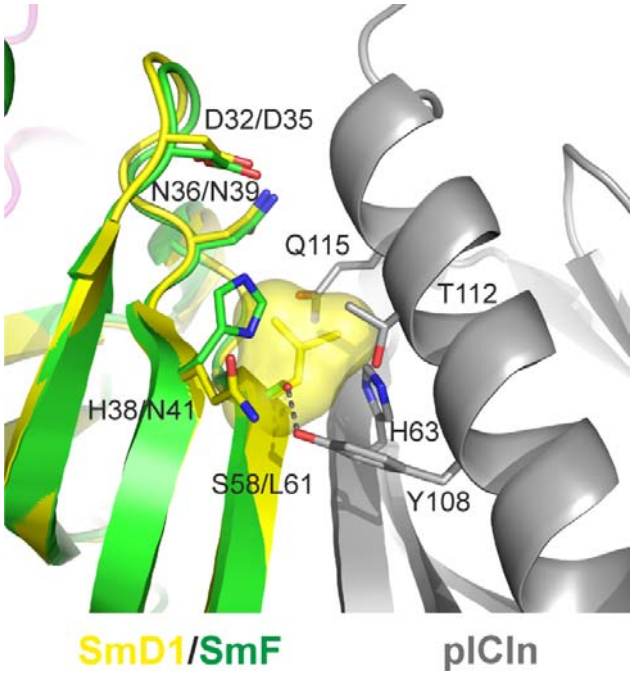

Figure S1. Superposition of SmF onto SmD1 within the SmD1-pICln interface, Related to Figure 1

Subunits are color-coded as in Figure 1. Labels indicate the SmD1 residue number before and the SmF residue number after the slash (/). Differences between SmD1 and SmF illustrate the molecular basis for pICln's function as a topological organizer of the Sm core. The bulky Leu61 of SmF in position of Ser58 of SmD1 would produce a clash with pICln and is shown in surface representation.

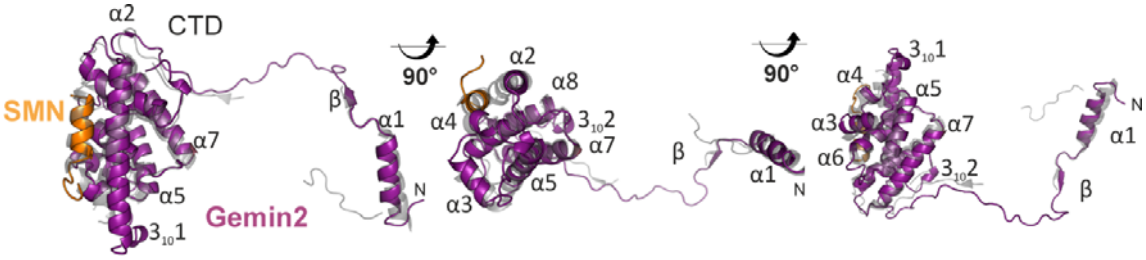

Figure S2. Superposition of the SMN/Gemin2 units from the 8S structure and from the structure of the late assembly intermediate (3S6N, (Zhang et al., 2011)), Related to Figure 2

Subunits are color-coded as in Figure 2, the late assembly intermediate is shown in grey, views similar to Figure 2B. Note the differences between both structures in the area of the two 3_{10}-helices, the N-terminus and the connecting region between β and α2 as well as the different orientation of α1 due to the more open conformation of the late assembly intermediate.

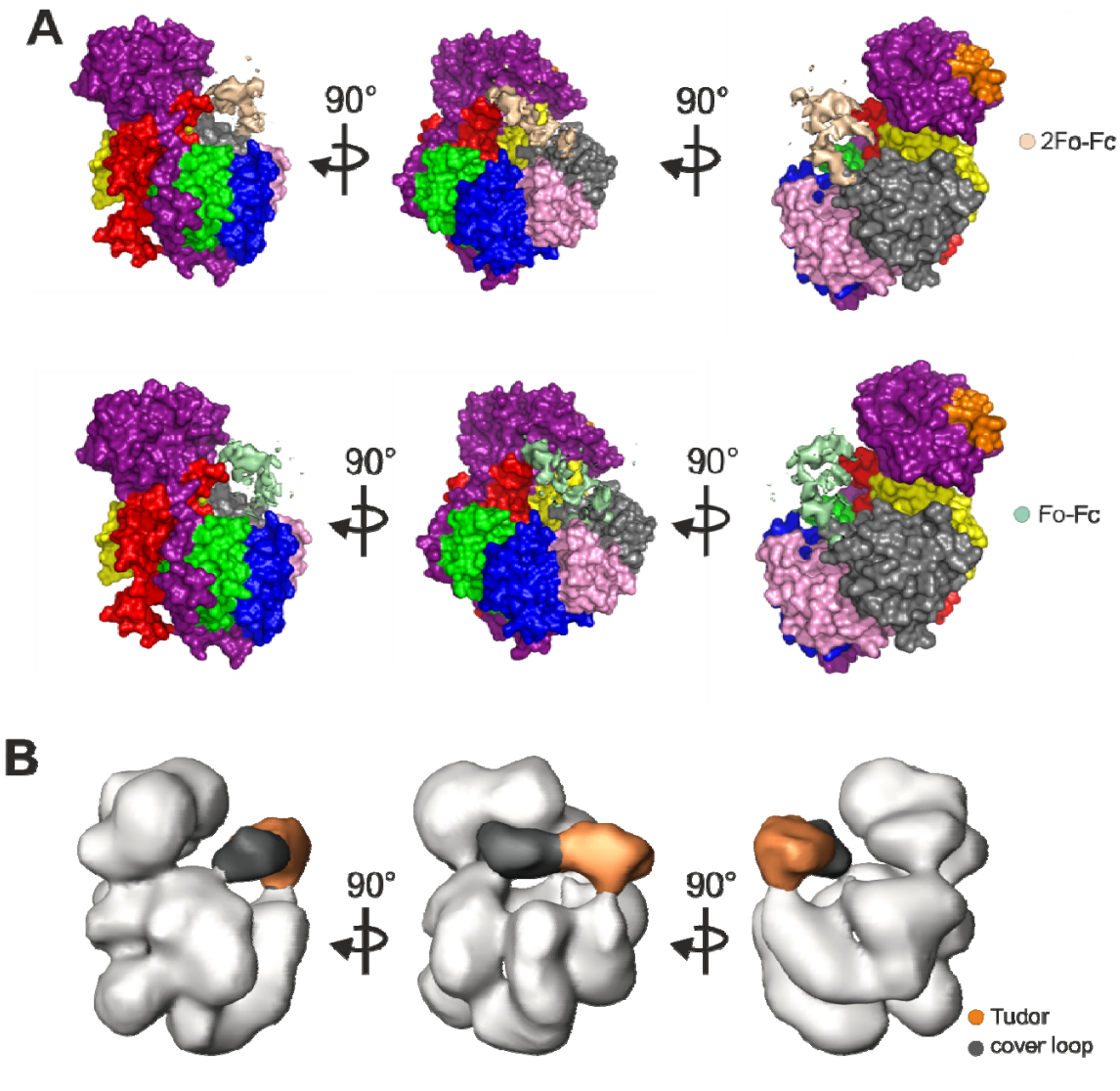
A
90°
90°
2Fo-Fc
90°
90°
Fo-Fc
B
90°
90°
Tudor
cover loop
C
V85
pICln
SmG
V85
N84
Y86
W83
E82
V81
K80
SmD1
Cover Loop
(disordered)
Eisenberg Hydrophobicity Scale
-2.5
+1.4

Figure S3. Additional density that can be attributed to the Tudor domain and the cover loop, Related to Figure 3

(A) Crystallographic electron density at the position of the EM extra density. The model is shown as a Conolly surface color coded as in Figure 2. The crystallographic map is B-factor smoothed (+50A^2) and contoured on the 0.8σ (2Fo-Fc, wheat) and 2.0σ (Fo-Fc, turquoise) level. (B) Three different views of the 8S EM density depicted in Figure 3 (opaque surface). The interpretation of the additional EM density is color coded as follows, Tudor domain in orange and the pICln cover loop in dark grey. **(C)** Surface view of pICln colored according to the Eisenberg hydrophobicity scale. Residues flanking the ends of the cover loop that were defined in the electron density were excluded from the surface representation and are depicted in stick model style.

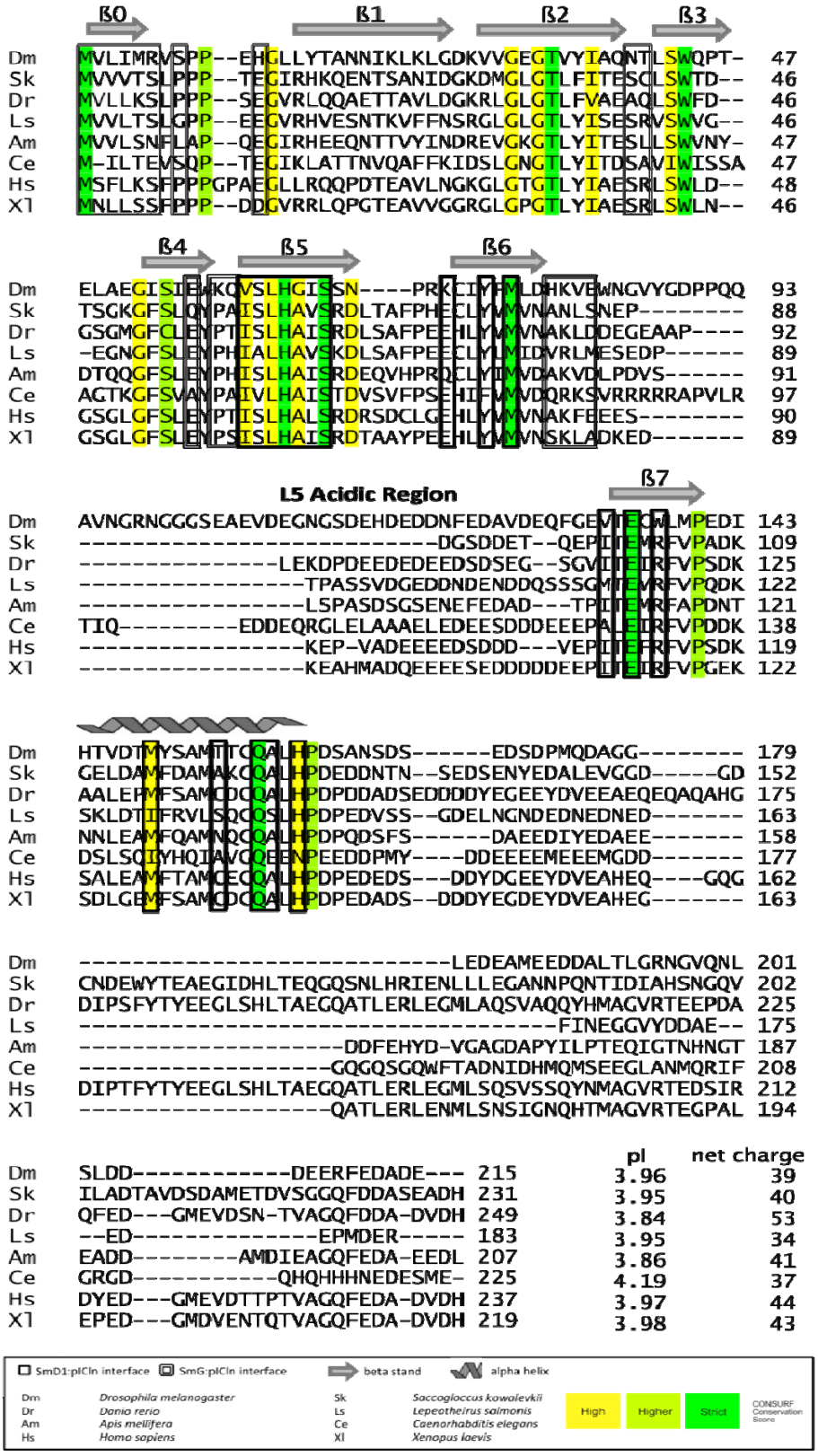
ß0
ß1
ß2
ß3
Dm MVLIMRVSPP--EHGLLYTANNIKLKLGDKVVGEGTVYIAQNTLSWQPT- 47
Sk MVVVTSLPPP--TEGIRHKQENTSANIDGKDMGLGTLFITESCLSWTD-- 46
Dr MVLLKSLPPP--SEGVRLQQAETTAVLDGKRLGLGTLFVAEAQLSWFD-- 46
Ls MVVLTSLGPP--EEGVRHVESNTKVFFNSRGLGLGTLYISESRVSWVG-- 46
Am MVVLSNFLAP--QEGIRHEEQNTTVYINDREVGKGTLYITESLLSWVNY- 47
Ce M-ILTEVSQP--TEGIKLATTNVQAFFKIDSLGNGTLYITDSAVIWISSA 47
Hs MSFLKSFPPPGPAEGLLRQQPDTEAVLNGKGLGTGTLYIAESRLSWLD-- 48
Xl MNLLSSFPPP--DDGVRRLQPGTEAVVGGRGLGPGTLYIAESRLSWLN-- 46
ß4
ß5
ß6
Dm ELAEGISIEAKQVSLHGISSN----PRKCIYFMLDHKVEWNGVYGDPPQQ 93
Sk TSGKGFSLQNPAISLHAVSRDLTAFPHECLYMVVNANLSNEP-------- 88
Dr GSGMGFCLENPTISLHAISRDLSAFPEEHLYMVVNAKLDDEGEAAP---- 92
Ls -EGNGFSLENPHIALHAVSKDLSAFPEECLYLMIDVRLMESEDP------ 89
Am DTQQGFSLENPHISLHAISRDEQVHPRQCLYIMVDAKVDLPDVS------ 91
Ce AGTKGFSVANPAIVLHAISTDVSVFPSEHIFMVVDQRKSVRRRRRAPVLR 97
Hs GSGLGFSLENPTISLHALSRDRSDCLGEHLYMVVNAKFEEES-------- 90
Xl GSGLGFSLENPSISLHAISRDTAAYPEEHLYMVVNSKLADKED------- 89
L5 Acidic Region
ß7
Dm AVNGRNGGGSEAEVDEGNGSDEHDEDDNFEDAVDEQFGEVTECMLMPEDI 143
Sk ----------DGSDDET--QEPTITEMRFVPADK 109
Dr ----------LEKDPDEEDEDEEDSDSEG--SGVTITEIRFVPSDK 125
Ls ----------TPASSVDGEDDNDENDDQSSSGMTEMRFVPQDK 122
Am ----------LSPASDSGSENEFEDAD---TPITEMRFAPDNT 121
Ce TIQ---------EDDEQRGLELAAAELEDEESDDDEEEPALEIRFVPDDK 138
Hs ----------KEP-VADEEEEDSDDD---VEPITEFRFVPSDK 119
Xl ----------KEAHMADQEEEESEDDDDDEEPITEIRFVPGEK 122
Dm HTVDTMYSAMTTCQALHPDSANSDS------EDSDPMQDAGG-------- 179
Sk GELDAMFDAMAKCQALHPDEDDNTN--SEDSENYEDALEVGGD-----GD 152
Dr AALEHMFSAMCDCQALHPDPDDADSEDDDDYEGEEYDVEEAEQEQAQAHG 175
Ls SKLDTIIFRVLSQCQSLHPDPEDVSS--GDELNGNDEDNEDNED------- 163
Am NNLEAMFQAMNQCQALHPDPQDSFS------DAEEDIYEDAEE-------- 158
Ce DSLSQIYHQIAVCQEENPEEDDPMY----DDEEEEMEEEMGDD------- 177
Hs SALEAMFTAMCEQQALHPDPEDEDS---DDYDGEEYDVEAHEQ----GQG 162
Xl SDLGEMFSAMCDCQALHPDPEDADS--DDDYEGDEYDVEAHEG------- 163
Dm ----------LEDEAMEEDDALTLGRNGVQNL 201
Sk CNDEWYTEAEGIDHLTEQGQSNLHRIENLLLEGANNPQNTIDIAHSNGQV 202
Dr DIPSFYTYEEGLSHLTAEGQATLERLEGMLAQSVAQQYHMAGVRTEEPDA 225
Ls ----------FINEGGVYDDAE-- 175
Am ----------DDFEHYD-VGAGDAPYILPTEQIGTNHNGT 187
Ce ----------GQGQSGQWFTADNIDHMQMSEEGLANMQRIF 208
Hs DIPTFYTYEEGLSHLTAEGQATLERLEGMLSQSVSSQYNMAGVRTEDSIR 212
Xl ----------QATLERLENMLSNSIGNQHTMAGVRTEGPAL 194
pI net charge
Dm SLDD----------DEERFEDADE--- 215 3.96 39
Sk ILADTAVDSDAMETDVSGGQFDDASEADH 231 3.95 40
Dr QFED---GMEVDSN-TVAGQFDDA-DVDH 249 3.84 53
Ls --ED----------EPMDER----- 183 3.95 34
Am EADD---------AMDIEAGQFEDA-EEDL 207 3.86 41
Ce GRGD----------QHQHHHNEDESME- 225 4.19 37
Hs DYED---GMEVDTTPTVAGQFEDA-DVDH 237 3.97 44
Xl EPED---GMDVENTQTVAGQFEDA-DVDH 219 3.98 43
SmD1:plCln interface
SmG:plCln interface
beta stand
alpha helix
Dm Drosophila melanogaster
Dr Danio rerio
Am Apis mellifera
Hs Homo sapiens
Sk Saccoglossus kowalevskii
Ls Lepeotheirus salmonis
Ce Caenorhabditis elegans
Xl Xenopus laevis
High
Higher
Strict
CONSURF Conservation Score

Figure S4. Multiple sequence alignment of pICln sequences from different organisms calculated with the program ConSurf, Related to Figure 4

Dm – *Drosophila melanogaster*, Dr – *Danio rerio*, Am – *Apis mellifera*, Hs – *Homo sapiens*, Sk – *Saccogloccus kowalevskii*, Ls – *Lepeotheirus salmonis*, Ce –*Caenorhabditis elegans*, Xl – *Xenopus laevis.* Amino acids involved in the formation of the interfaces to SmD1 or SmG are indicated by boxes (fat lines: interface to SmD1, duplex lines: interface to SmG). Conserved residues are colored according to their calculated conservation score. The color code is as in Figure 4D. The acidic region located within the cover loop is marked in red. pI is the theoretical isoelectric point calculated from the sequence information.

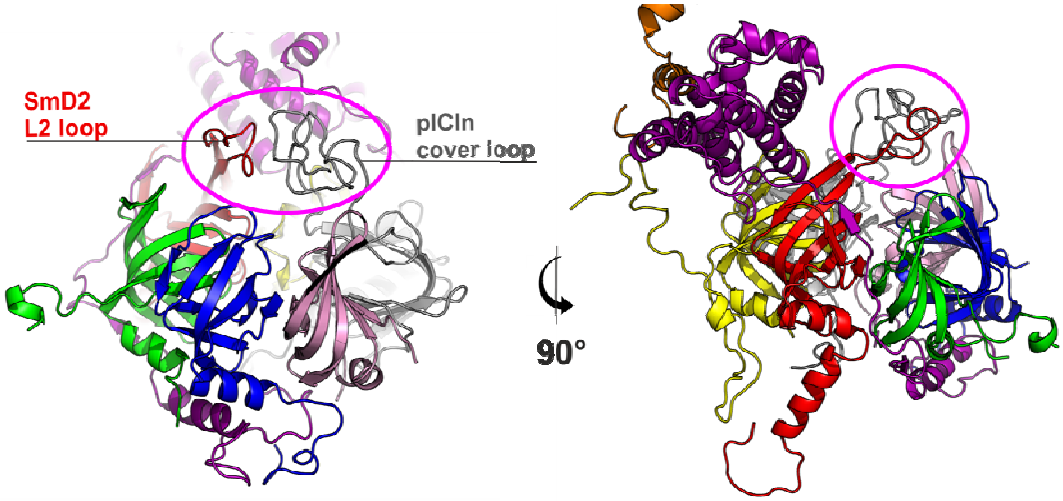

Figure S5. Structure of the reconstructed cover loop after 2ns of MD simulation, Related to Figure 5

An 8S model with reconstructed cover and SmD2-L2 loop after 2ns simulation time in two perpendicular views along the ring plane. The color coding corresponds to Figure 2 and the cover loop region is encircled in pink. During the simulation, the cover loop (modeled according to the NMR structure 1ZYI by Furst et al) quickly approaches the conformation depicted here, where it is stabilized in a compact conformation next to the central hole mainly by strong electrostatic interactions with the predominantly positive charged L2 loop of SmD2 and the ring surface. Also the smaller SmD1 L2 loop (yellow, in the background) contributes to the stabilization. Note the excellent agreement of the two loops in the observed conformation and location with the extra EM density shown in Figure 3B.

Table S1. RMSD values between the 8S, 6S complexes and the structure of the late assembly intermediate (3S6N), Related to Figure 5

	8S/6S	**8S/3S6N**	**6S/3S6N**
overall	0.57	1.68	1.08
Sm pentamer	0.49	1.17	1.08
SmF	0.10	0.36	0.41
SmE	0.16	0.59	0.70
SmG	0.11	0.69	0.70
SmD1	0.17	0.44	0.54
SmD2	0.17	0.50	0.52
pICln	0.25	N/A	N/A
SMN	N/A	0.36	N/A
Gemin2	N/A	1.63	N/A

Values given for superposition of the whole molecules (overall) as well as for individual subunits from the respective complexes.

Table S2. Plasmids used in this study, Related to the Experimental Procedures

Plasmid name	Comments
pETM13	EMBL Heidelberg
D1D2:pRK172	This work and Chari et al. (2008), kind gift from Dr. Christian Kambach, PSI Villigen (Kambach et al., 1999).
EFG:pET15b	This work and Chari et al. (2008), kind gift from Dr. Christian Kambach, PSI Villigen (Kambach et al., 1999).
dSMN(1-122):pETM30	This work, Chari et al. (2008) and Kroiss et al. (2008), the *Drosophila melanogaster* SMN cDNA was obtained from the DGRC, Bloomington, Indiana. The nucleotide sequence encoding for dSMN(1-122) was PCR amplified as an *Nco*I-*Not*I fragment and cloned into the *Nco*I and *Not*I sites of pETM30.
dSMN(1-40):pETM30	This work, the nucleotide sequence encoding for dSMN(1-40) was PCR amplified as an *Nco*I-*Acc65I* fragment and cloned into the *Nco*I and *Acc65I* sites of pETM30. This construct was used for EM.
dGemin2:pETM30	This work, Chari et al. (2008) and Kroiss et al. (2008), the *Drosophila melanogaster* Gemin2 cDNA was obtained from the DGRC, Bloomington, Indiana. The dGemin2 ORF was PCR amplified as an *Nco*I-*Not*I fragment and cloned into *Nco*I and *Not*I sites of pETM30.
dpICln:pET28a	This work and Chari et al. (2008), the *Drosophila melanogaster* pICln cDNA was obtained from the Drosophila Genomics Resource Center (DGRC), Bloomington, Indiana. The dpICln ORF was PCR amplified as an *Nco*I-*Not*I fragment and cloned into *Nco*I and *Not*I sites of pET28a.
dpICln(Δ181-215):pETM13	This work, the dpIClnΔ181-215-His_6 fragment was PCR amplified from dpICln:pET28a as a NcoI-His_6-Acc65I fragment and cloned into the NcoI and Acc65I sites of pETM13. This pICln construct was used for crystallization of the 8S complex.
dpICln(Δ90-125 Δ160-215):pETM13	This work, the dpIClnΔ160-215 fragment was PCR amplified from dpIClnΔ181-215:pETm13 as a NcoI-His_6Acc65I fragment and cloned into the NcoI and Acc65I sites of pETm13. Subsequent deletion of residues 90-

	125 was conducted according to the Quick-Change site-directed mutagenesis protocol. This construct was used for EM.
dpICln(Δ90-125 H144A Δ160-215):pETM13	This work, mutagenesis of residue 144 was conducted according to the Quick-Change site-directed mutagenesis protocol. This construct was used for crystallization of the 6S complex.

Supplemental References

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291-294.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*, 10037-10041.

Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. Rna *7*, 1531-1542.

Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999). Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. Human molecular genetics *8*, 2351-2357.

Chari, A., and Fischer, U. (2010). Cellular strategies for the assembly of molecular machines. Trends in biochemical sciences *35*, 676-683.

Chari, A., Golas, M.M., Klingenhager, M., Neuenkirchen, N., Sander, B., Englbrecht, C., Sickmann, A., Stark, H., and Fischer, U. (2008). An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. Cell *135*, 497-509.

Chari, A., Paknia, E., and Fischer, U. (2009). The role of RNP biogenesis in spinal muscular atrophy. Current opinion in cell biology *21*, 387-393.

Derewenda, Z.S., and Vekilov, P.G. (2006). Entropy and surface engineering in protein crystallization. Acta crystallographica Section D, Biological crystallography *62*, 116-124.

Ellis, R.J. (2001). Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci *26*, 597-604.

Ellis, R.J. (2006). Molecular chaperones: assisting assembly in addition to folding. Trends Biochem Sci *31*, 395-401.

Felder, C.E., Prilusky, J., Silman, I., and Sussman, J.L. (2007). A server and database for dipole moments of proteins. Nucleic Acids Res *35*, W512-521.

Fischer, U., Englbrecht, C., and Chari, A. (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. Wiley Interdiscip Rev RNA *2*, 718-731.

Fisher, D.E., Conner, G.E., Reeves, W.H., Wisniewolski, R., and Blobel, G. (1985). Small nuclear ribonucleoprotein particle assembly in vivo: demonstration of a 6S RNA-free core precursor and posttranslational modification. Cell *42*, 751-758.

Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001a). SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. MolCell *7*, 1111-1117.

Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001b). The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. Mol Cell Biol *21*, 8289-8300.

Furst, J., Schedlbauer, A., Gandini, R., Garavaglia, M.L., Saino, S., Gschwentner, M., Sarg, B., Lindner, H., Jakab, M., Ritter, M., *et al.* (2005a). ICln159 folds into a pleckstrin homology domain-like structure. Interaction with kinases and the splicing factor LSm4. The Journal of biological chemistry *280*, 31276-31282.

Furst, J., Schedlbauer, A., Gandini, R., Garavaglia, M.L., Saino, S., Gschwentner, M., Sarg, B., Lindner, H., Jakab, M., Ritter, M., *et al.* (2005b). ICln159 folds into a pleckstrin homology domain-like structure. Interaction with kinases and the splicing factor LSm4. J Biol Chem *280*, 31276-31282.

Gruber, C.C., and Pleiss, J. (2011). Systematic benchmarking of large molecular dynamics simulations employing GROMACS on massive multiprocessing facilities. J Comput Chem *32*, 600-606.

Gubitz, A.K., Feng, W., and Dreyfuss, G. (2004). The SMN complex. Exp Cell Res *296*, 51-56.

Harauz, G., Boekema, E., and van Heel, M. (1988). Statistical image analysis of electron micrographs of ribosomal subunits. Methods Enzymol *164*, 35-49.

Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. Nature *402*, C47-52.

Huber, S., Braun, G., Schroppel, B., and Horster, M. (1998). Chloride channels ClC-2 and ICln mRNA expression differs in renal epithelial ontogeny. Kidney international Supplement *67*, S149-151.

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

Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J., and Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell *96*, 375-387.

Karplus, P.A., and Diederichs, K. (2012). In Press. Science.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., *et al.* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods *5*, 53-55.

Kelch, B.A., Makino, D.L., O'Donnell, M., and Kuriyan, J. (2011). How a DNA polymerase clamp loader opens a sliding clamp. Science *334*, 1675-1680.

Kondrashov, D.A., Van Wynsberghe, A.W., Bannen, R.M., Cui, Q., and Phillips, G.N., Jr. (2007). Protein structural variation in computational models and crystallographic data. Structure *15*, 169-177.

Kroiss, M., Schultz, J., Wiesner, J., Chari, A., Sickmann, A., and Fischer, U. (2008). Evolution of an RNP assembly system: A minimal SMN complex facilitates formation of UsnRNPs in *Drosophila melanogaster*. PNAS *105*, 10045–10050.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic acids research *33*, W299-302.

Leung, A.K., Nagai, K., and Li, J. (2011). Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. Nature *473*, 536-539.

Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol *128*, 82-97.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Meister, G., Bühler, D., Pillai, R., Lottspeich, F., and Fischer, U. (2001a). A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. Nat Cell Biol *3*, 945-949.

Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001b). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. Curr Biol *11*, 1990-1994.

Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. EMBO J *21*, 5853-5863.

Nishiyama, M., Ishikawa, T., Rechsteiner, H., and Glockshuber, R. (2008). Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. Science *320*, 376-379.

Palfi, Z., Jae, N., Preusser, C., Kaminska, K.H., Bujnicki, J.M., Lee, J.H., Gunzl, A., Kambach, C., Urlaub, H., and Bindereif, A. (2009). SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes. Genes Dev *23*, 1650-1664.

Pellizzoni, L., Yong, J., and Dreyfuss, G. (2002). Essential role for the SMN complex in the specificity of snRNP assembly. Science *298*, 1775-1779.

Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K., Li, J., and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature *458*, 475-480.

Raker, V.A., Plessel, G., and Luhrmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J *15*, 2256-2269.

Sander, B., Golas, M.M., and Stark, H. (2003a). Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra. J Struct Biol *142*, 392-401.

Sander, B., Golas, M.M., and Stark, H. (2003b). Corrim-based alignment for improved speed in single-particle image processing. J Struct Biol *143*, 219-228.

Sander, B., Golas, M.M., and Stark, H. (2005). Advantages of CCD detectors for de novo three-dimensional structure determination in single-particle electron microscopy. J Struct Biol *151*, 92-105.

Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proc Natl Acad Sci U S A *103*, 8060-8065.

Suhre, K., and Sanejouand, Y.H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic acids research *32*, W610-614.

Tarn, W.Y., and Steitz, J.A. (1997). Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. Trends Biochem Sci *22*, 132-137.

Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. Nature structural & molecular biology *18*, 1414-1420.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J Comput Chem *26*, 1701-1718.

Van Heel, M. (1987). Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. Ultramicroscopy *21*, 111-123.

van Heel, M., and Frank, J. (1981). Use of multivariate statistics in analysing the images of biological macromolecules. Ultramicroscopy *6*, 187-194.

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

Weber, G., Trowitzsch, S., Kastner, B., Luhrmann, R., and Wahl, M.C. (2010). Functional organization of the Sm core in the crystal structure of human U1 snRNP. The EMBO journal *29*, 4172-4184.

Yong, J., Wan, L., and Dreyfuss, G. (2004). Why do cells need an assembly machine for RNA-protein complexes? Trends Cell Biol *14*, 226-232.

Zhang, R., So, B.R., Li, P., Yong, J., Glisovic, T., Wan, L., and Dreyfuss, G. (2011). Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. Cell *146*, 384-395.