SUPPLEMENTARY MATERIALS

Negative Cooperativity between Gemin2 and RNA Determines RNA Selection and Release of the SMN Complex in snRNP Assembly

Hongfei Yi, Li Mu, Congcong Shen, Xi Kong, Yingzhi Wang, Yan Hou and Rundong Zhang*

Containing

4 supplemental tables (Tables S1-4),

11 supplemental figures (Figures S1-11)

Table S1. Data Collection and Refinement Statistics

Variable protein components*				
	Complex A	Complex B	Complex C	Previous
Gemin2	FL(1-280)	ΔN39 (40-280)	ΔN39 (40-280)	FL(1-280)
SmD1	ΔC(1-82)	ΔC(1-82)	ΔC(1-82)	FL(1-119)
SmG	Yes	No	Yes	Yes
Data collection				
Wavelength (Å)	0.97853	0.97853	0.97853	0.99993
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell (Å):	83.32,	83.17,	82.96,	82.83,
a, b, c	115.76,	114.09,	114.12,	84.60,
	128.21	125.39	130.14	104.66
Highest resolution (Å)[$]:				
a*, b*, and c*	3.3, 3.5, 3.9	3.4, 3.4, 4.0	3.1, 3.2, 3.6	2.4, 3.2, 2.6
Unique reflections	15470	14569	18768	19831
Completeness (%)	98.2 (94.6)[a]	99.5(97.9)[c]	97.8(95.2)[b]	99.3(98.6)[d]
	79.2(15.9)[e]	84.2(35.5)[g]	83.1(28.8)[f]	76.0(20.0)[h]
R_{meas}	0.154 (0.905)	0.233(0.914)	0.166(0.885)	
R_{pim}	0.048 (0.261)	0.087(0.297)	0.056(0.243)	
Mean I/σ	13.6 (2.2)	12.9(2.0)	10.4(2.0)	13.7(3.8)
Redundancy	12.4 (12.3)	10.8 (8.7)	12.4(12.7)	6.3 (3.8)
Refinement Statistics				
Resolution range(Å)	48-3.3	47-3.4	47-3.1	40-2.5
R factor (%)	20.5	19.1	17.2	22.4
R_{free} factor (%)	26.5	24.5	24.0	29.7
Number of reflections	15470	14569	18768	19831
Number of atoms	4793	4346	4890	4826
Rmsd bond length (Å)	0.0112	0.0122	0.0129	0.0130
Rmsd bond angles (°)	1.576	1.624	1.477	1.579
Ramachandran plot (%)[#]:				
Favored,	93.9,	91.6,	94.7,	96.4,
additional allowed,	5.9,	7.5,	4.6,	3.0,
disallowed	0.2	0.9	0.7	0.6
PDB code	5XJQ	5XJS	5XJR	5XJL

* Complexes A-C and the previous 7S complex all contain SMN(residues 26-62), SmD2, SmF and SmE.

[$] Ellipsoidal truncation was performed on each data set due to serious anisotropic diffraction.

[a-d] For resolution range of 48-3.9, 47-3.6, 47-4.0 and 40-3.2 Å respectively. The number in the parenthesis corresponds to the highest-resolution shell of 4.04-3.90, 3.72-3.60, 4.12-4.00 and 3.35-3.20 Å respectively.

[e-h] For resolution range of 48-3.3, 47-3.1, 47-3.4 and 40-2.5 Å respectively. The number in the parenthesis corresponds to the highest-resolution shell of 3.4-3.3, 3.2-3.1, 3.52-3.40 and 2.62-2.50 Å respectively.

[#] gained from Coot program.

Table S2. RNA sequences used in this study.

Name	RNA sequence
9nt	AAUUUUUGA
3'Sm	GGGAAUUGAAAACUUUUCCCAAUACCCCAAUUUUUGA
U4	GGGAAUUGAAAACUUUUCCCAAUACCCCGCCGUGACGACUUGCAAUAU AGUCGGCACUGGCAAUUUUUGACAGUCUCUACGGAGACUG
U4ΔSm	GGGAAUUGAAAACUUUUCCCAAUACCCCGCCGUGACGACUUGCAAUAU AGTCGGCACUGGCAACCCCCGACAGUCUCUACGGAGACUG
U4-3'ss	GGGAAUUGAAAACUUUUCCCAAUACCCCGCCGUGACGACUUGCAAUAU AGUCGGCACUGGCAAUUUUUGACAGUCUCUACGCUCUGAC
U4-3'Δ	GGGAAUUGAAAACUUUUCCCAAUACCCCGCCGUGACGACUUGCAAUAU AGUCGGCACUGGCAAUUUUUGAC
U4-5'ss	GGGAAUUGAAAACUUUUCCCAAUACCCCAAUUUUUGACAGUCUCUACG GAGACUG
U4-5'Δ	GGCAAUUUUUGACAGUCUCUACGGAGACUG
U4-5'Δ-3'ss	GGCAAUUUUUGACAGUCUCUACGCUCUGAC
flU4	GGGCAGCUUUGCGCAGUGGCAGUAUCGUAGCCAAUGAGGUCUAUCCGA GGCGCGAUUAUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGUGAC GACUUGCAAUAUAGUCGGCACUGGCAAUUUUUGACAGUCUCUACGGAG ACUG
flU4ΔSm	GGGCAGCUUUGCGCAGUGGCAGUAUCGUAGCCAAUGAGGUCUAUCCGA GGCGCGAUUAUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGUGAC GACUUGCAAUAUAGUCGGCACUGGCAACCCCCGACAGUCUCUACGGAG ACUG
flU4-spacer	GGGCAGCUUUGCGCAGUGGCAGUAUCGUAGCCAAUGAGGUCUAUCCGA GGCGCGAUUAUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGUGAC GACUUGCAAUAUAGUCGGCACUCCGAAUUUUUGACAGUCUCUACGGAG ACUG
flU4-spacer-3'ss	GGGCAGCUUUGCGCAGUGGCAGUAUCGUAGCCAAUGAGGUCUAUCCGA GGCGCGAUUAUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGUGAC GACUUGCAAUAUAGUCGGCACUCCGAAUUUUUGACAGUCUCUACGCUC UGAC

Red, Sm site. Blue, ΔSm. Cyan, spacer.

Table S3. The components and amounts used for gel filtration chromatographic (GFC) assay.

Input components (RNA colored in blue)	Amount used (molar ratio in last brackets)
D1s/D2 + F/E/G (5Sm)	200μg + 200μg (1:1)
Gemin2ΔN39	200μg
F/E/G	1mg
3'Sm	10μg
5Sm + 3'Sm	400μg + 100μg (1:1)
5Sm + 9nt	400μg + 120μg (about 1:5)
Gemin2ΔN39/SMN_{Ge2BD} + D1s/D2 + F/E/G (7SΔN preparation)	0.6mg + 1mg + 1mg (1:2:2)
7SΔN + 3'Sm	160μg + 40μg (about 1:1.5)
7SΔN + U4	160μg + 60μg (1:1)
(7SΔN + D3/B) + U4	(160μg + 60μg) + 60μg (1:1.5:1)
7SΔN + U4ΔSm	160μg + 60μg (1:1)
7SΔN + U4-3'Δ	160μg + 50μg (1:1)
7SΔN + U4-5'Δ	160μg + 40μg (1:2)
U4-5'Δ	5μg
7SΔN + U4-3'ss	160μg + 60μg (1:1)
7SΔN + U4-5'ss	160μg + 40μg (1:1)
7SΔN + U4-5'Δ-3'ss	80μg + 40μg (1:4)
7SΔN + U4-5'Δ-3'ss + U4-5'Δ	80μg + 15μg + 15μg (1:1.5:1.5)
5Sm + U4-5'Δ	400μg + 40μg (2:1)
5Sm + U4-5'ss	400μg + 80μg (2:1)
flU4	10μg
7SΔN + flU4	120μg + 120μg (1:1.6)
7SΔN + flU4ΔSm	100μg + 100μg (1:1.6)
7SΔN + flU4-spacer	120μg + 120μg (1:1.6)
7SΔN + flU4-spacer-3'ss	80μg + 80μg (1:1.6)
(7SΔN + D3/B) + flU4	(120μg + 60μg) + 120μg（1:1.5:1.6）
(7SΔN + D3/B) + flU4-spacer-3'ss	(80μg + 40μg) + 80μg (1:2:1.6)
7SΔN + (flU4-spacer + U4-5'Δ-3'ss)	80μg + (80μg + 16μg) (1:1.6:1.6)
7SΔN + (flU4-spacer-3'ss + U4-5'Δ)	80μg + (80μg + 16μg) (1:1.6:1.6)
(7SΔN + D3/B) + (flU4-spacer-3'ss + U4-5'Δ)	(80μg + 40μg) + (80μg + 16μg) (1:2:1.6:1.6)
(7SΔN + D3/B) + (flU4 + U4-5'Δ)	(100μg + 50μg) + (100μg + 20μg) (1:2:1.6:1.6)

Table S4. The GFC elution positions of the RNAs, proteins, and their complexes studied.

Components	Elution position (ml)	+7SΔN39: Elution position (ml)	+5Sm: Elution position (ml)	+7Sm: Elution position (ml)
9nt	~19.6	Not formed	14.37	Not tested
3'Sm	16.64	Not formed	13.51	Not tested
U4	14.75	13.31	Not tested	13.43
U4ΔSm	14.91	Not formed	Not tested	Not tested
U4-3'ss	14.48	13.11	Not tested	Not tested
U4-3'Δ	15.66	Not formed	Not tested	Not tested
U4-5'ss	15.67	13.65(RNA+5Sm)	13.50	Not tested
U4-5'Δ	17.48	14.29(RNA+5Sm)	14.27	Not tested
U4-5'Δ-3'ss	16.62	13.97	Not tested	Not tested
flU4	13.08	12.26	Not tested	12.31
flU4ΔSm	12.93	Not formed	Not tested	Not tested
flU4-spacer	12.97	12.24(RNA+5Sm)	Not tested	Not tested
flU4-spacer-3'ss	12.90	~12.11	Not tested	12.14
7SΔN39	13.78	----------	----------	----------
7S	13.61	----------	----------	----------
D1s/D2	~16.5	----------	----------	----------
F/E/G	14.29(dimer), 16.62(monomer)	----------	----------	----------
D1s/D2/F/E/G(5Sm)	15.06	----------	----------	----------
Gemin2ΔN39/SMN_{Ge2BD}	15.66	----------	----------	----------

5Sm indicates D1s/D2 and F/E/G. 7Sm indicates 5Sm and D3(1-75)/B(1-91).

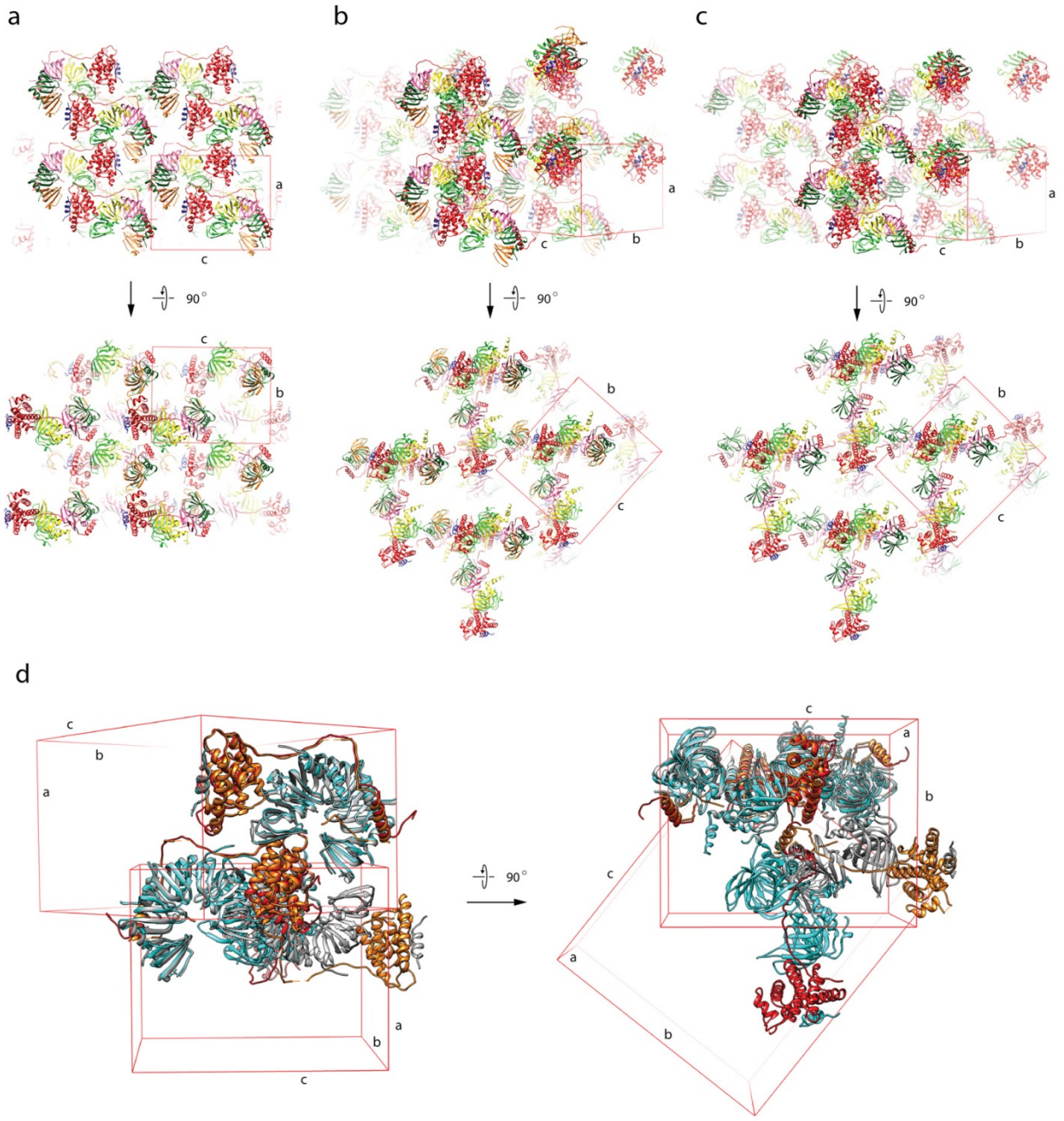


Fig. S1. Crystal packing comparison of three complexes. Two views of the crystal packing for each of 3 complexes: (a) the 7S complex from the previous study (3S6N), (b) Complex A and (c) Complex C. The five Sm proteins, D1, D2, F, E and G are colored in green, lemon, pink, dark green and orange respectively. Gemin2 and SMN_{Ge2BD} are colored in red and blue respectively. Unit cells and axis are showed. (d) Comparison of 3S6N (5Sm in light gray and Gemin2 in orange) with Complex A (5Sm in cyan and Gemin2 in red) in crystal packing. Two molecules of the complexes are superimposed in direction a.

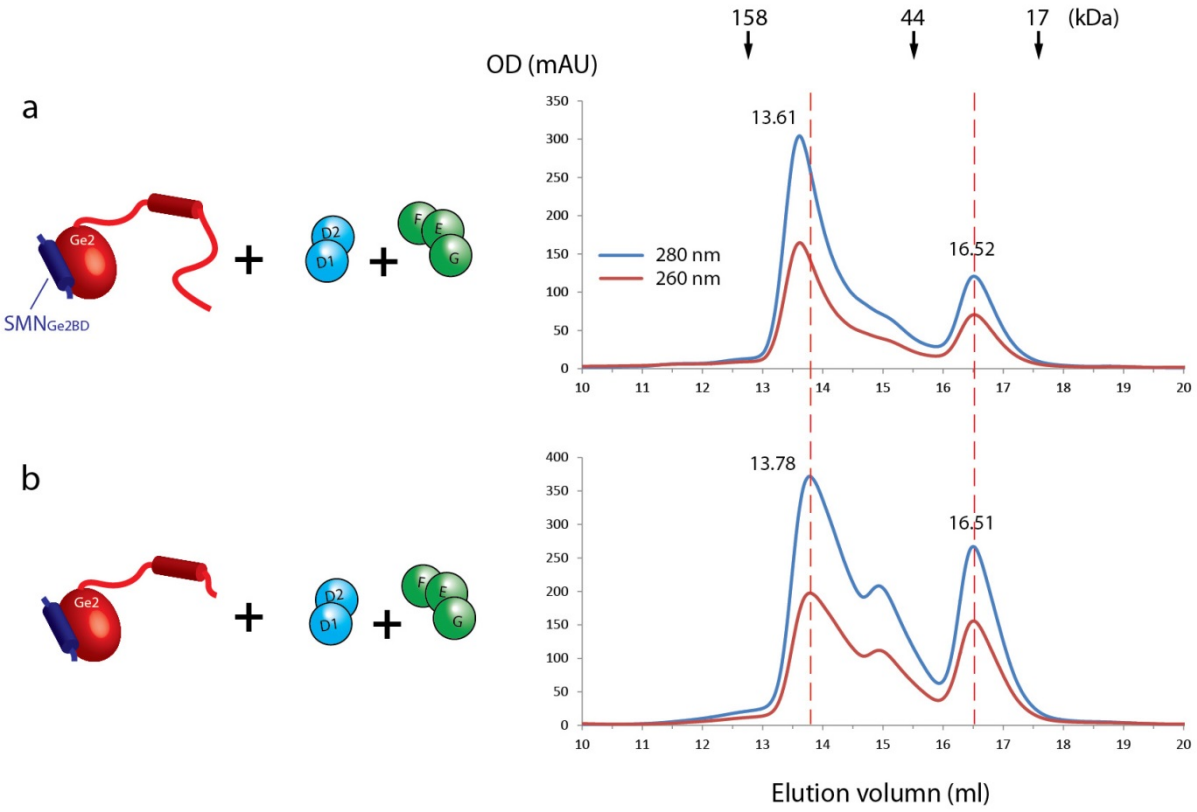


Fig. S2. The gel filtration profiles of 7S and 7SΔN39. Reconstitution of 7S or 7SΔN39 was made by mixing the 5 Sm proteins, D1s/D2 and F/E/G, with (a) Gemin2/SMN_{Ge2BD} or (b) Gemin2ΔN39/SMN_{Ge2BD} respectively followed by gel filtration chromatography separation (right panels). The input components are showed in cartoon (left panels). The positions of standard proteins are indicated at the top. Peak positions (ml) are indicated.

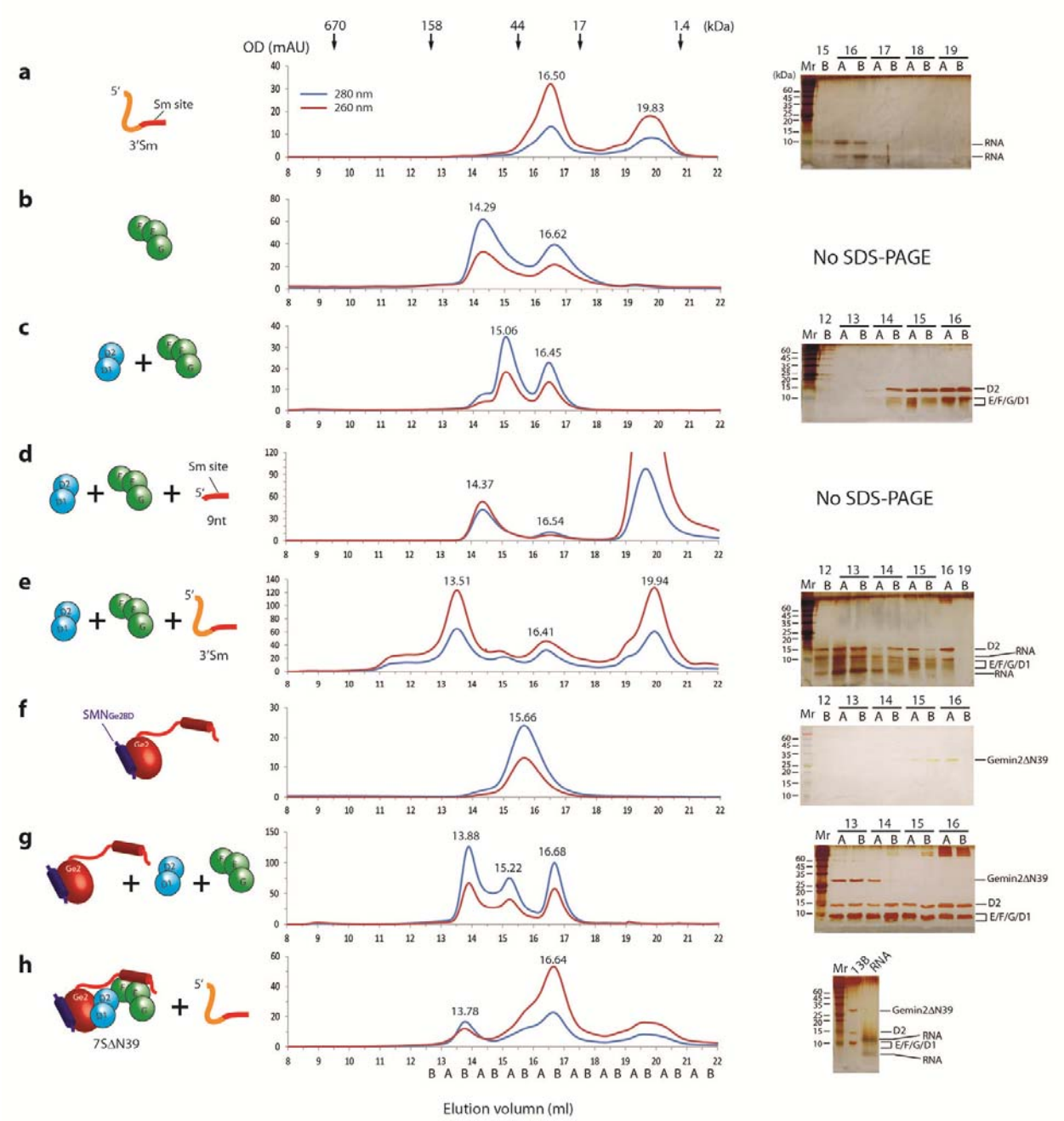


Fig. S3. The 7SΔN39 complex cannot bind RNA containing solely a Sm site at 3' end. Individual RNA containing Sm site at its 3' end, 3'Sm (a), SmF/E/G (b), mixture of SmD1/D2 and SmF/E/G (c), mixture of 9nt and the 5 Sm proteins (d), mixture of 3'Sm and the 5 Sm proteins (e), individual Gemin2ΔN39 (f), mixture of Gemin2ΔN39 and the 5 Sm proteins (g) or mixture of reconstituted 7SΔN39 and 3'Sm (h) was separated by gel filtration chromatography (middle panels) and individual fractions were analyzed by SDS-PAGE and silver staining (right panels). The input components are showed in cartoon (left panels). The positions of standard proteins are indicated at the top. Fractions (A &B) are named on the basis of volume positions (bottom). Peak positions (ml) are indicated. There were two oligomeric forms of F/E/G, monomer (16.45ml) and dimer (15.06 ml) in Panel (b).

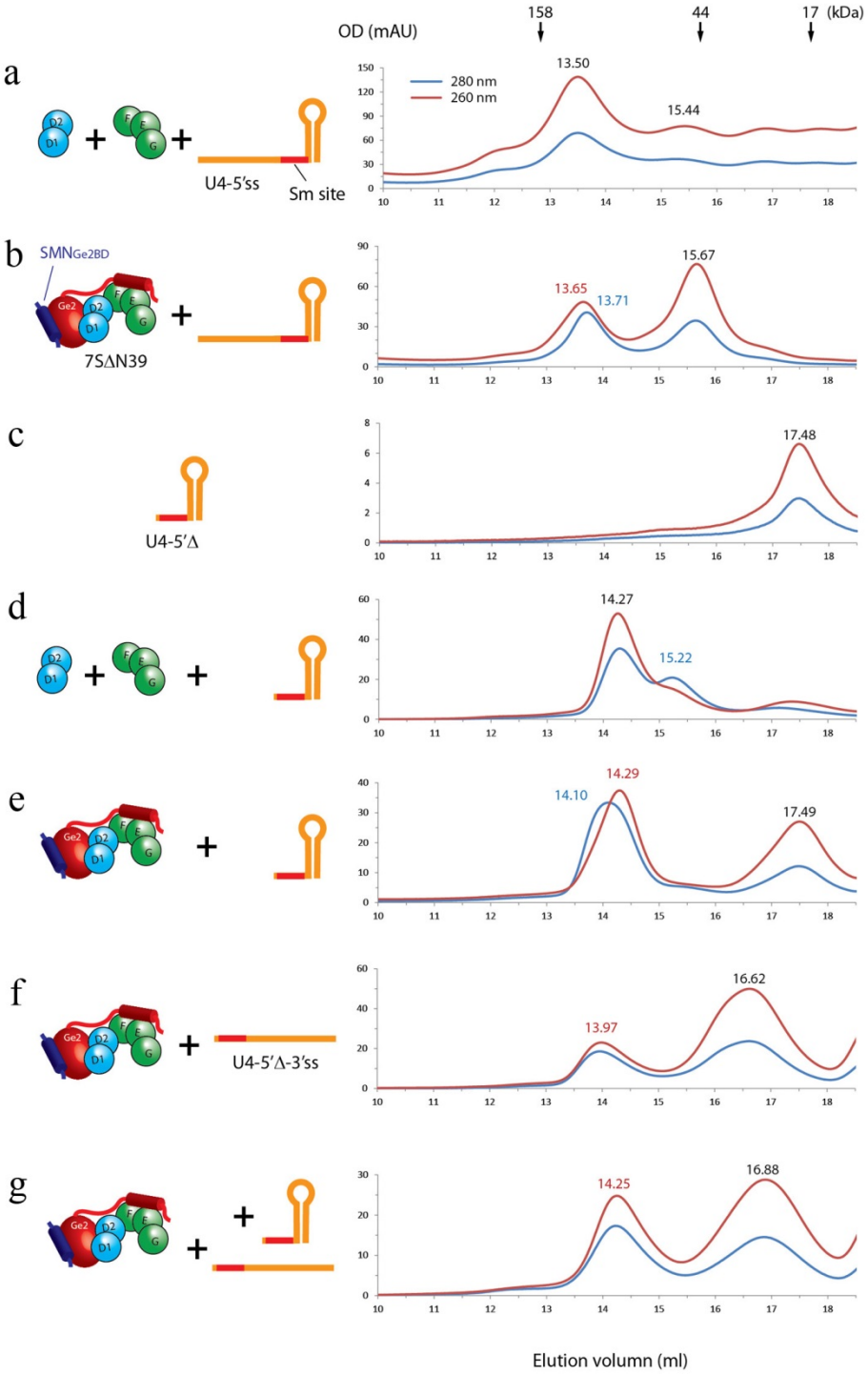


Fig. S4. The gel filtration profiles of RNA and Sm subcore formation. Mixture of 5Sm and U4-5'ss (a), mixture of 7SΔN39 and U4-5'ss (b), U4-5'Δ alone (c), mixture of 5Sm and U4-5'Δ (d), mixture of 7SΔN39 and U4-5'Δ (e), mixture of 7SΔN39 and U4-5'Δ-3'ss (f), or mixture of 7SΔN39 and equal molar amount of U4-5'Δ and U4-5'Δ-3'ss (g) was separated by gel filtration chromatography (right panels). (b) and (e) are the same as in Fig. 2 and are showed here for easy comparison. The input components are showed in cartoon (left panels). The positions of standard proteins are indicated at the top. Precise peak positions (ml) are indicated.

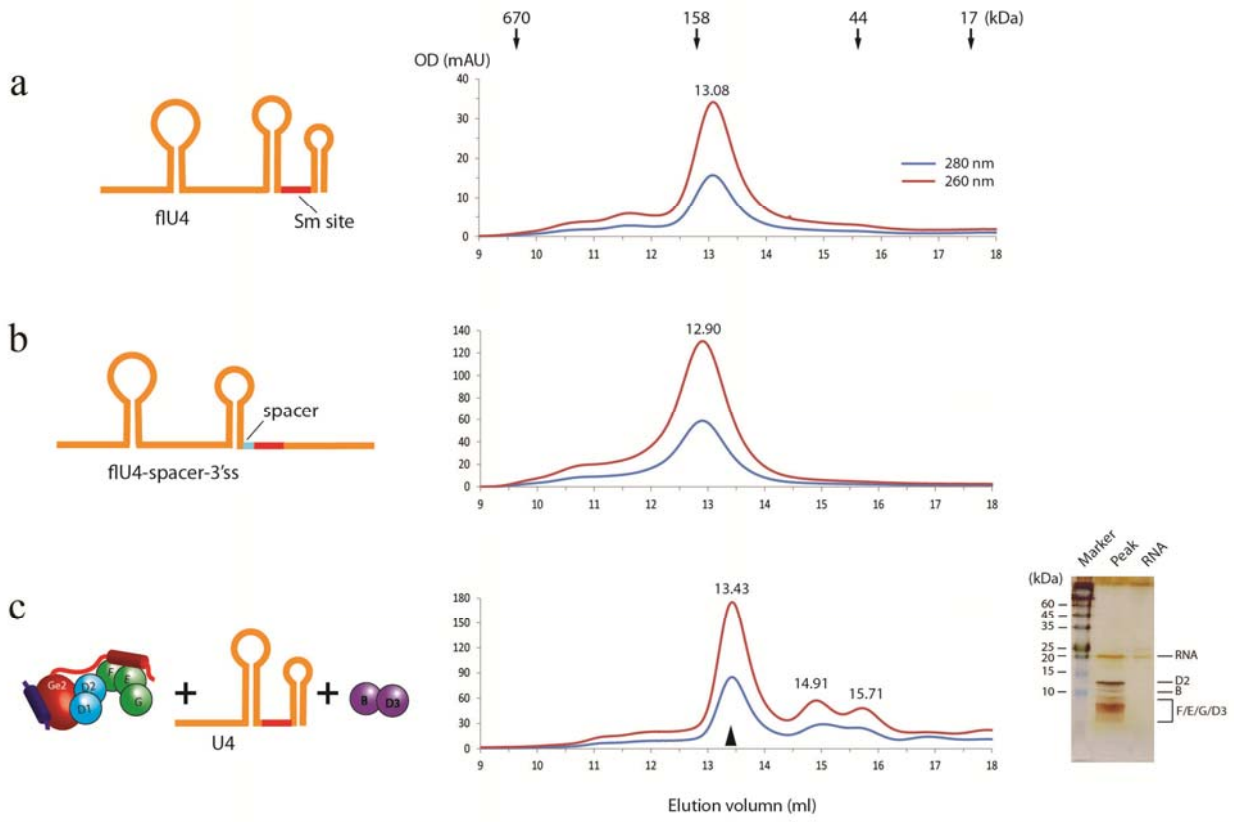


Fig. S5. The gel filtration profiles of flU4, flU4-spacer-3'ss and Sm core assembly of U4. flU4 (a) or flU4-spacer-3'ss (b) was separated by gel filtration chromatography (right panels). (c) Mixture of 7SΔN39, D3/B and U4 snRNA after incubation was separated by GFC (middle panel) and the peak (solid arrow head) and input RNA were subjected to SDS-PAGE followed by silver staining (right panel). Input components are showed in cartoon (left panels). The positions of standard proteins are indicated at the top. Precise peak positions (ml) are indicated.

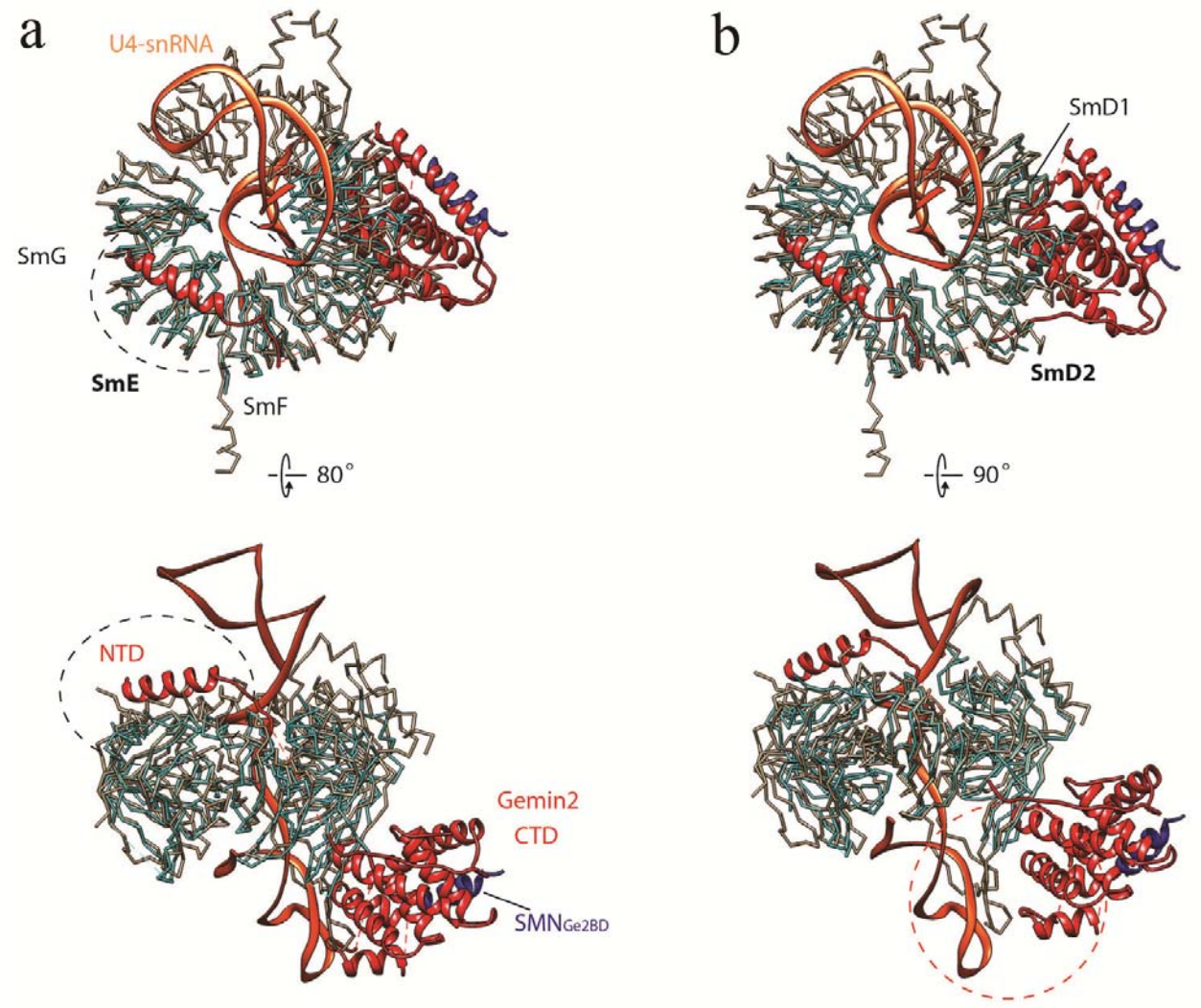


Fig. S6. The bindings of Gemin2 and U4 snRNA to 5Sm are compatible sterically. Superposition of SmE (a) or SmD2 (b) of the 7SΔN39 complex with that of U4 snRNP core (PDB code 4WZJ) reveals that Gemin2 is compatible with U4 snRNA in binding to 5Sm spatially. The five Sm proteins in the 7SΔN39 complex are colored in cyan, and in the U4 snRNP core in grey. SMN_{Ge2BD} and Gemin2 are colored in blue and red respectively. U4 snRNA is colored in orange.

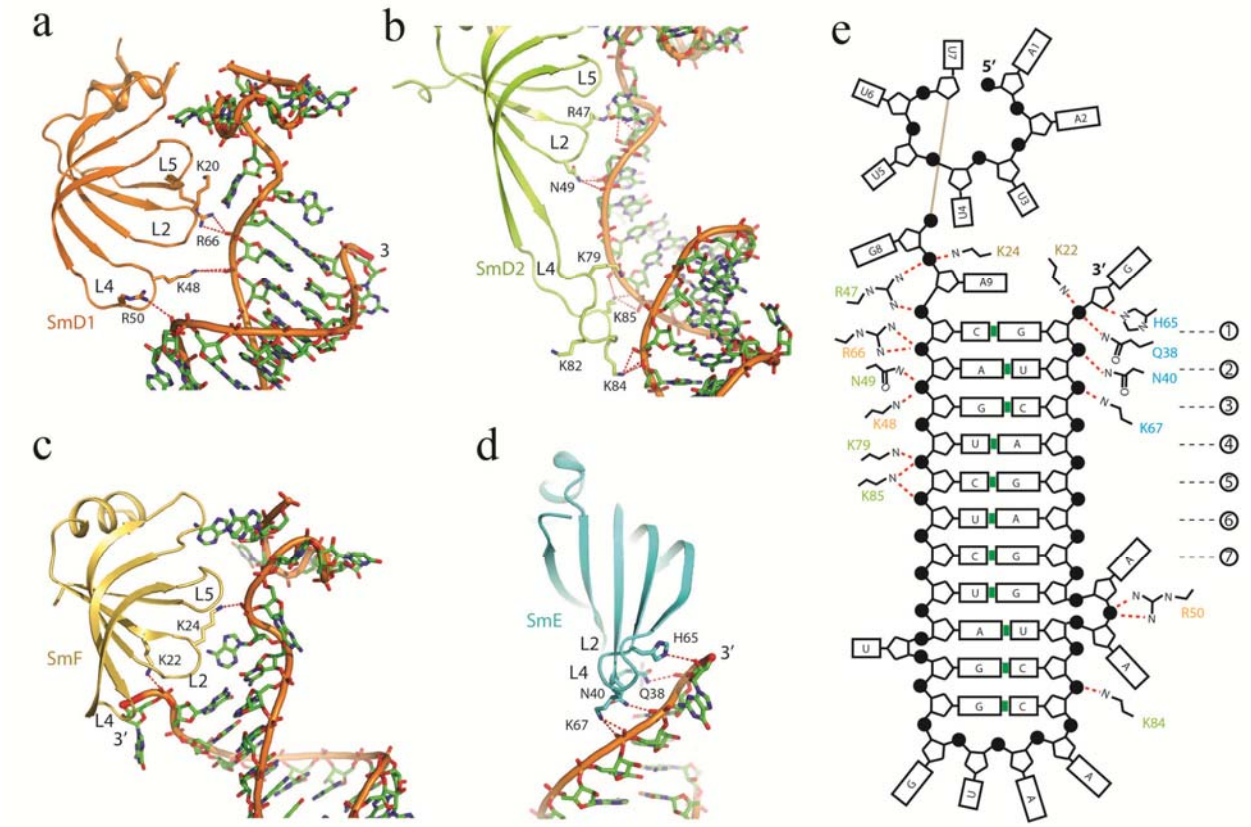


Fig. S7. Detailed interactions between the 3'-SL of U4 RNA and the Sm proteins. (a-d) Detailed interactions between each Sm protein and RNA 3'-SL (PDB code 4WZJ). (e) Schematic summary of the interactions between Sm proteins and RNA 3'-SL.

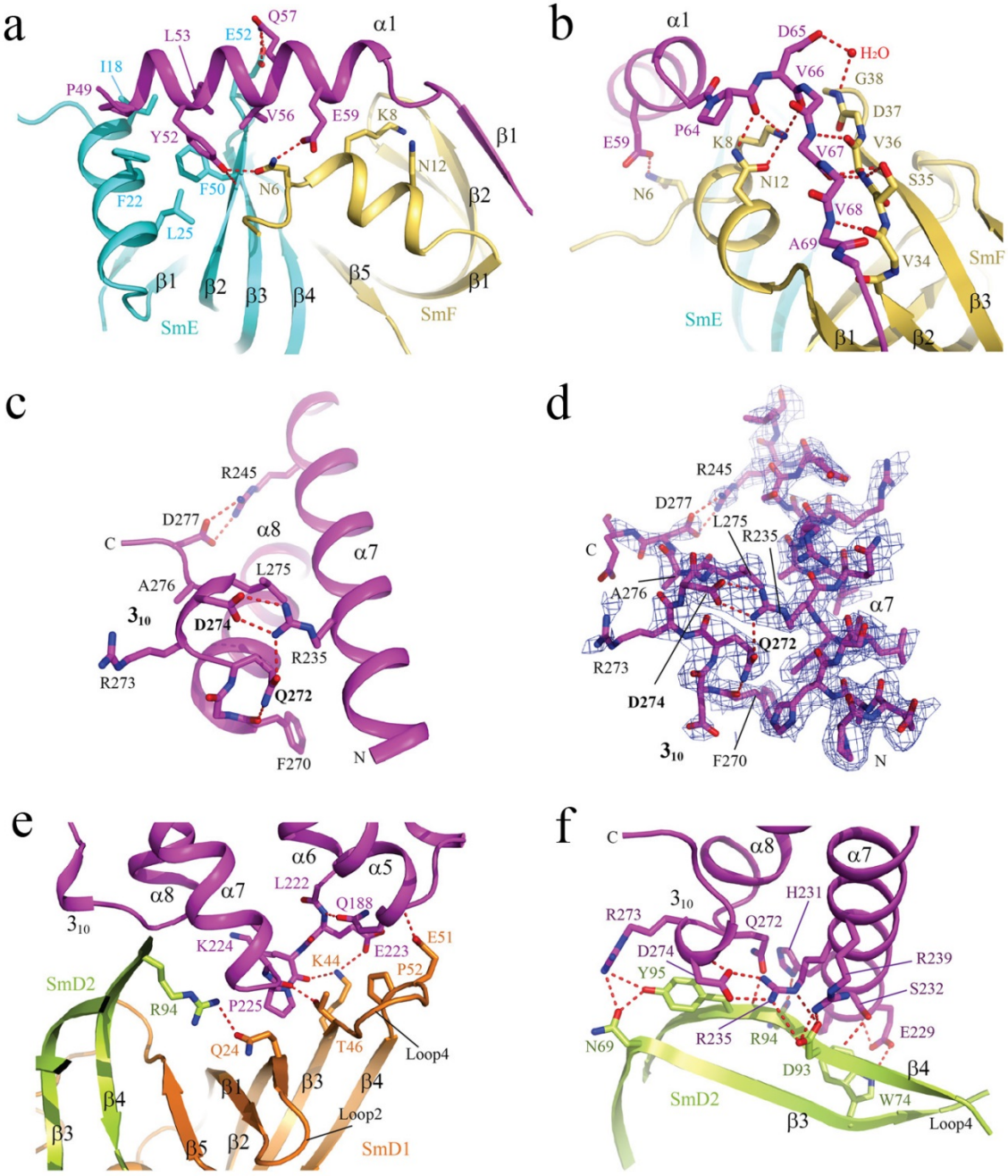
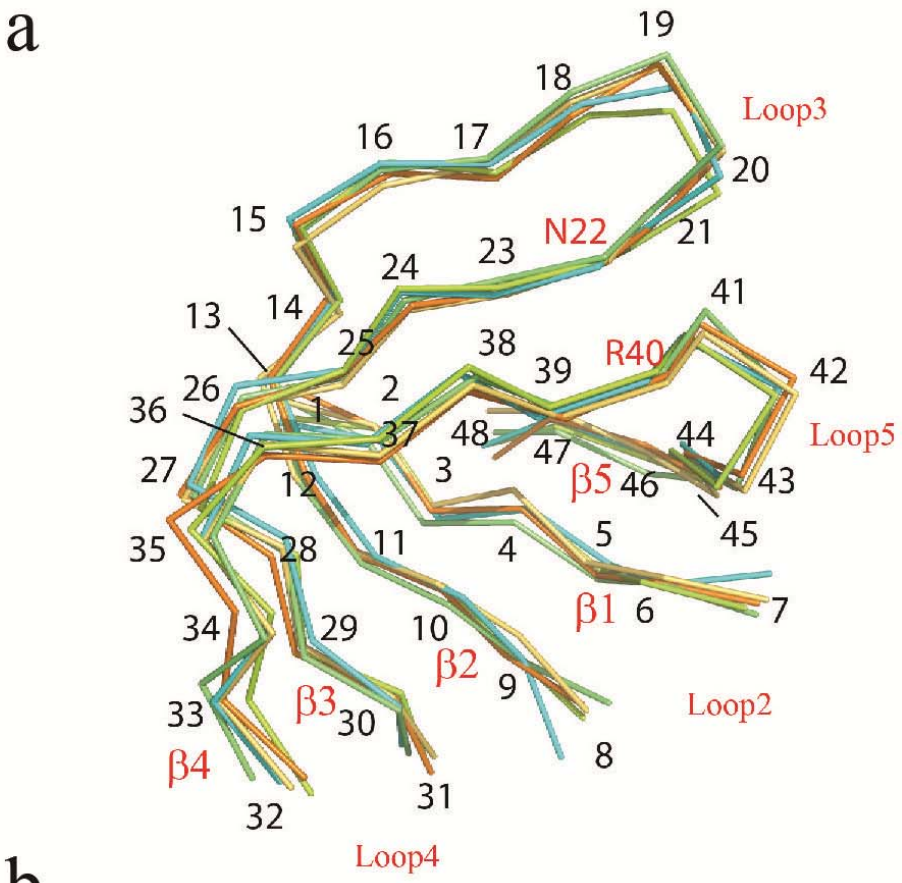


Fig. S8. Detailed views of interactions between Gemin2 and Sm proteins. (a-b) Gemin2's NTD interacts with SmF/E. (c-d) the C-terminal 3_{10} helix of Gemin2 and its interactions with helices α7-8. SigmaA-weighted 2Fo-Fc electron density maps (blue meshes) are contoured at 1.1σ in (D). (e-f) Gemin2's CTD interacts with SmD1/D2. Hydrogen bonds and salt bridges are shown as red dashed lines. The five Sm proteins, D1, D2, F, E and G, are colored in orange, lemon, yellow, cyan and green respectively. Gemin2 and SMN_{Ge2BD} are colored in purple and blue respectively. Oxygen and nitrogen are colored in red and blue respectively. Water is shown in red sphere.

a



b

RMSD (Å)	SmD1	SmD2	SmF	SmE	SmG
SmD1	**0.52**	0.78	4.11	6.84	7.86
SmD2	0.54	**0.39**	1.64	4.05	5.61
SmF	2.37	1.82	**0.48**	1.87	3.39
SmE	4.83	4.07	1.69	**0.51**	0.83
SmG	7.56	6.93	4.72	1.31	**0.80**

Fig. S9. Conformational changes of 5Sm from the 7S intermediate state to assembled Sm cores. (a) The Sm fold defined by 48 residues of each Sm protein used for superposition and analysis. Numbering residues in this system is shown at each position. (b) Position differences of each pair of Sm proteins in the 7S complex and U1 Sm core when one pair is superposed using the 48-residue Sm fold system.

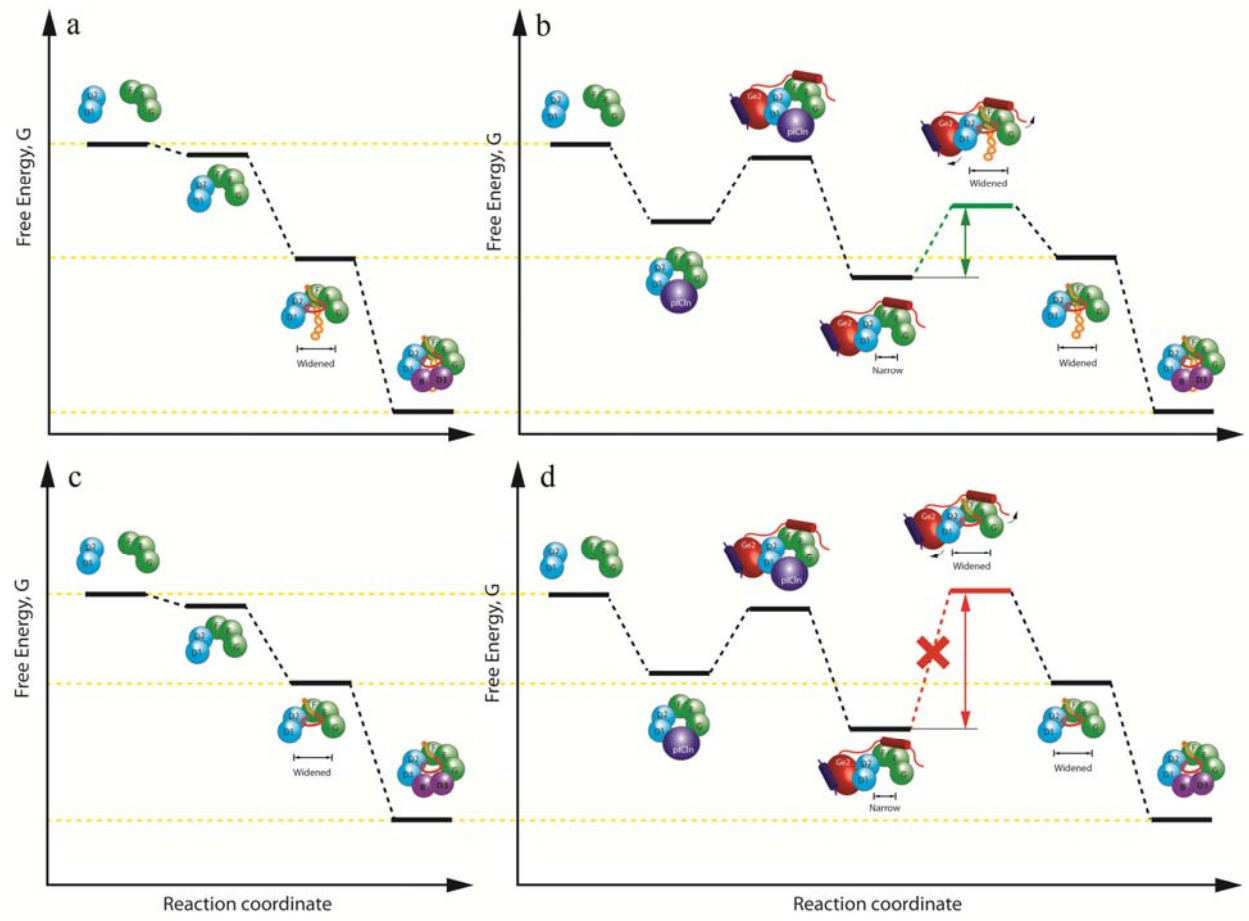


Fig. S10. Energy diagrams for Sm core assembly. Sm core assembles on the snRNP code without assembly chaperons, in vitro (a) and with assembly chaperons, in vivo (b), and on the Sm site alone without chaperons, in vitro (c) and with chaperons, in vivo (d).

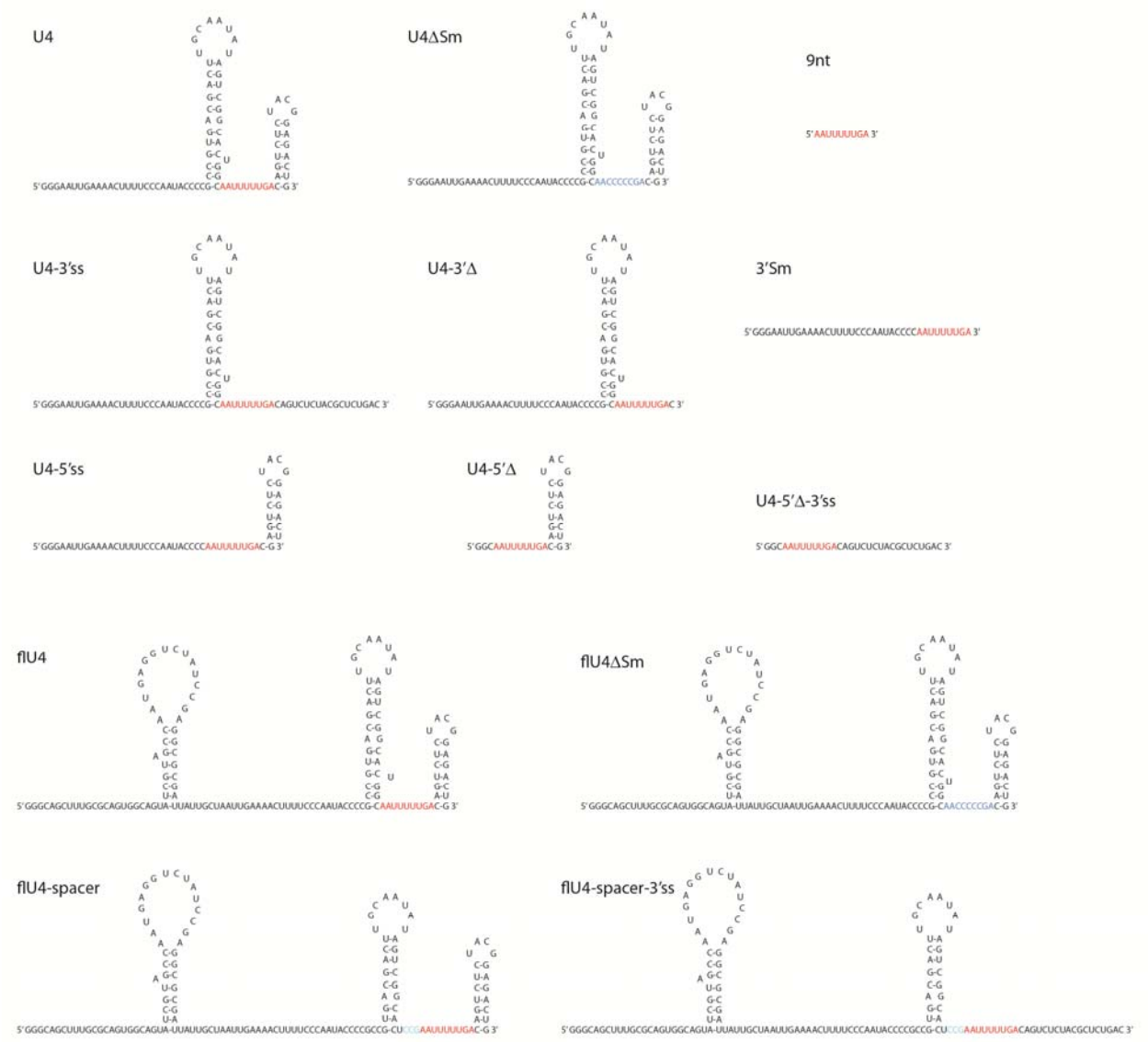


Fig. S11. The sequences and secondary structures of the RNAs used in this study. Sm site is colored in red, ΔSm in blue, and the 3nt spacer in cyan (same as in cartoon).