



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 10:47 PM GMT

PDB ID : 2Y9D  
Title : STRUCTURE OF THE SPLICEOSOMAL U4 SNRNP CORE DOMAIN  
Authors : Leung, A.K.W.; Nagai, K.; Li, J.  
Deposited on : 2011-02-13  
Resolution : 3.60 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

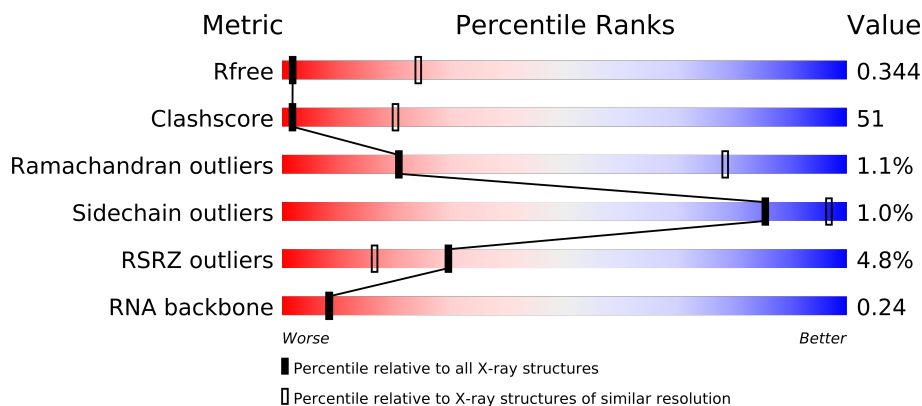
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



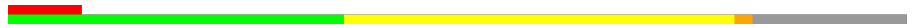


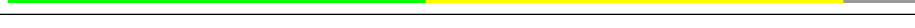
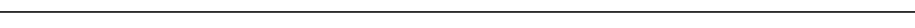
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1020 (3.86-3.34)
Clashscore	79885	1155 (3.80-3.40)
Ramachandran outliers	78287	1109 (3.80-3.40)
Sidechain outliers	78261	1108 (3.80-3.40)
RSRZ outliers	66119	1000 (3.84-3.36)
RNA backbone	1838	1012 (4.40-2.76)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	95	
1	H	95	
1	O	95	
2	B	119	
2	I	119	
2	P	119	
3	C	118	
3	J	118	
3	Q	118	
4	D	126	
4	K	126	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	R	126	
5	E	92	
5	L	92	
5	S	92	
6	F	86	
6	M	86	
6	T	86	
7	G	76	
7	N	76	
7	U	76	
8	V	68	
8	X	68	
8	Y	68	

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 18840 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN-ASSOCIATE DPROTEINS B AND B'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	90	Total	C	N	O	S	0	0	0
			724	455	133	129	7			
1	H	95	Total	C	N	O	S	0	0	0
			759	477	139	135	8			
1	O	91	Total	C	N	O	S	0	0	0
			733	461	135	130	7			

- Molecule 2 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	96	Total	C	N	O	S	0	0	1
			754	479	134	137	4			
2	I	94	Total	C	N	O	S	0	0	1
			737	468	131	135	3			
2	P	95	Total	C	N	O	S	0	0	1
			746	474	133	136	3			

- Molecule 3 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	105	Total	C	N	O	S	0	0	0
			845	529	157	154	5			
3	J	106	Total	C	N	O	S	0	0	0
			852	534	158	155	5			
3	Q	90	Total	C	N	O	S	0	0	1
			709	446	135	123	5			

- Molecule 4 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	89	Total	C	N	O	S	0	0	1
			696	433	126	130	7			
4	K	89	Total	C	N	O	S	0	0	1
			696	433	126	130	7			
4	R	89	Total	C	N	O	S	0	0	1
			696	433	126	130	7			

- Molecule 5 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	76	Total	C	N	O	S	0	0	0
			629	397	112	116	4			
5	L	75	Total	C	N	O	S	0	0	0
			622	392	111	115	4			
5	S	78	Total	C	N	O	S	0	0	0
			645	407	115	119	4			

- Molecule 6 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	83	Total	C	N	O	S	0	0	0
			657	416	107	128	6			
6	M	83	Total	C	N	O	S	0	0	0
			657	416	107	128	6			
6	T	83	Total	C	N	O	S	0	0	0
			657	416	107	128	6			

- Molecule 7 is a protein called SMALL NUCLEAR RIBONUCLEOPROTEIN G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	73	Total	C	N	O	S	0	0	0
			566	356	102	102	6			
7	N	73	Total	C	N	O	S	0	0	0
			568	358	102	102	6			
7	U	70	Total	C	N	O	S	0	0	1
			533	336	96	95	6			

- Molecule 8 is a RNA chain called HOMO SAPIENS U4A SNRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	V	68	Total	C	N	O	P	0	0	0
			1453	650	263	473	67			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	X	68	Total	C	N	O	P	0	0	0
			1453	650	263	473	67			
8	Y	68	Total	C	N	O	P	0	0	0
			1453	650	263	473	67			

There are 93 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	.	-	U	DELETION	GB 36174
V	.	-	G	DELETION	GB 36174
V	.	-	C	DELETION	GB 36174
V	.	-	A	DELETION	GB 36174
V	.	-	A	DELETION	GB 36174
V	.	-	U	DELETION	GB 36174
V	.	-	A	DELETION	GB 36174
V	.	-	U	DELETION	GB 36174
V	813	G	-	INSERTION	GB 36174
V	814	A	-	INSERTION	GB 36174
V	815	A	-	INSERTION	GB 36174
V	816	A	-	INSERTION	GB 36174
V	.	-	U	DELETION	GB 36174
V	.	-	A	DELETION	GB 36174
V	.	-	C	DELETION	GB 36174
V	.	-	G	DELETION	GB 36174
V	846	U	-	INSERTION	GB 36174
V	847	A	-	INSERTION	GB 36174
V	848	U	-	INSERTION	GB 36174
V	849	G	-	INSERTION	GB 36174
V	850	G	-	INSERTION	GB 36174
V	851	G	-	INSERTION	GB 36174
V	852	U	-	INSERTION	GB 36174
V	853	A	-	INSERTION	GB 36174
V	854	A	-	INSERTION	GB 36174
V	855	C	-	INSERTION	GB 36174
V	856	C	-	INSERTION	GB 36174
V	857	U	-	INSERTION	GB 36174
V	858	A	-	INSERTION	GB 36174
V	859	A	-	INSERTION	GB 36174
V	860	G	-	INSERTION	GB 36174
X	.	-	U	DELETION	GB 36174
X	.	-	G	DELETION	GB 36174
X	.	-	C	DELETION	GB 36174

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
X	.	-	A	DELETION	GB 36174
X	.	-	A	DELETION	GB 36174
X	.	-	U	DELETION	GB 36174
X	.	-	A	DELETION	GB 36174
X	.	-	U	DELETION	GB 36174
X	813	G	-	INSERTION	GB 36174
X	814	A	-	INSERTION	GB 36174
X	815	A	-	INSERTION	GB 36174
X	816	A	-	INSERTION	GB 36174
X	.	-	U	DELETION	GB 36174
X	.	-	A	DELETION	GB 36174
X	.	-	C	DELETION	GB 36174
X	.	-	G	DELETION	GB 36174
X	846	U	-	INSERTION	GB 36174
X	847	A	-	INSERTION	GB 36174
X	848	U	-	INSERTION	GB 36174
X	849	G	-	INSERTION	GB 36174
X	850	G	-	INSERTION	GB 36174
X	851	G	-	INSERTION	GB 36174
X	852	U	-	INSERTION	GB 36174
X	853	A	-	INSERTION	GB 36174
X	854	A	-	INSERTION	GB 36174
X	855	C	-	INSERTION	GB 36174
X	856	C	-	INSERTION	GB 36174
X	857	U	-	INSERTION	GB 36174
X	858	A	-	INSERTION	GB 36174
X	859	A	-	INSERTION	GB 36174
X	860	G	-	INSERTION	GB 36174
Y	.	-	U	DELETION	GB 36174
Y	.	-	G	DELETION	GB 36174
Y	.	-	C	DELETION	GB 36174
Y	.	-	A	DELETION	GB 36174
Y	.	-	A	DELETION	GB 36174
Y	.	-	U	DELETION	GB 36174
Y	.	-	A	DELETION	GB 36174
Y	.	-	U	DELETION	GB 36174
Y	813	G	-	INSERTION	GB 36174
Y	814	A	-	INSERTION	GB 36174
Y	815	A	-	INSERTION	GB 36174
Y	816	A	-	INSERTION	GB 36174
Y	.	-	U	DELETION	GB 36174
Y	.	-	A	DELETION	GB 36174

*Continued on next page...*

*Continued from previous page...*

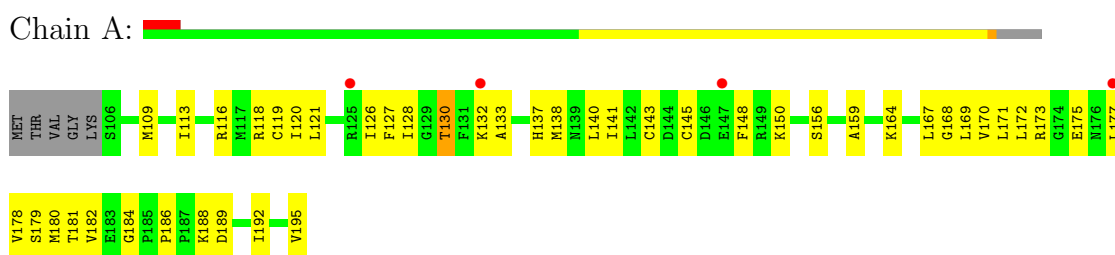
Chain	Residue	Modelled	Actual	Comment	Reference
Y	.	-	C	DELETION	GB 36174
Y	.	-	G	DELETION	GB 36174
Y	846	U	-	INSERTION	GB 36174
Y	847	A	-	INSERTION	GB 36174
Y	848	U	-	INSERTION	GB 36174
Y	849	G	-	INSERTION	GB 36174
Y	850	G	-	INSERTION	GB 36174
Y	851	G	-	INSERTION	GB 36174
Y	852	U	-	INSERTION	GB 36174
Y	853	A	-	INSERTION	GB 36174
Y	854	A	-	INSERTION	GB 36174
Y	855	C	-	INSERTION	GB 36174
Y	856	C	-	INSERTION	GB 36174
Y	857	U	-	INSERTION	GB 36174
Y	858	A	-	INSERTION	GB 36174
Y	859	A	-	INSERTION	GB 36174
Y	860	G	-	INSERTION	GB 36174



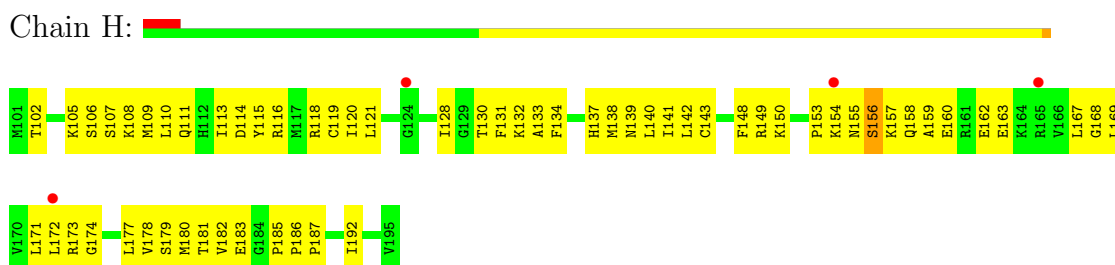
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

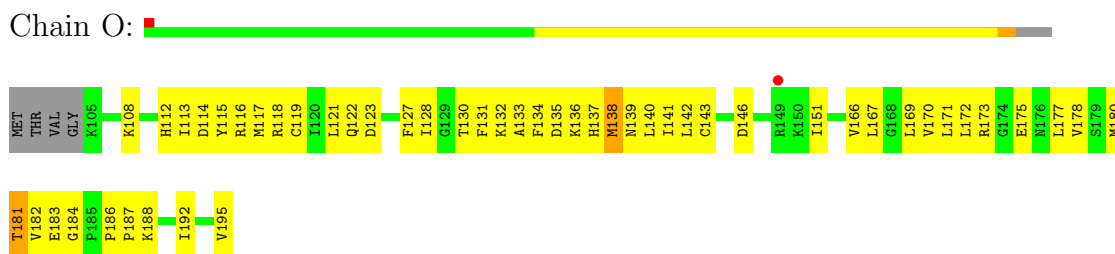
- Molecule 1: SMALL NUCLEAR RIBONUCLEOPROTEIN-ASSOCIATED PROTEINS B AND B'



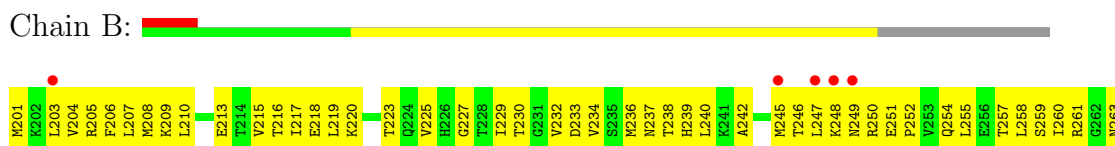
- Molecule 1: SMALL NUCLEAR RIBONUCLEOPROTEIN-ASSOCIATED PROTEINS B AND B'



- Molecule 1: SMALL NUCLEAR RIBONUCLEOPROTEIN-ASSOCIATED PROTEINS B AND B'



- Molecule 2: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1





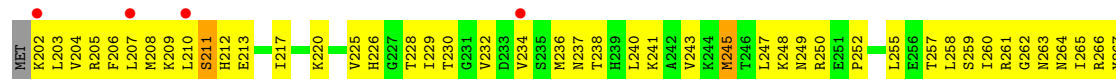
• Molecule 2: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1

Chain I:



• Molecule 2: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1

Chain P:



• Molecule 3: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2

Chain C:



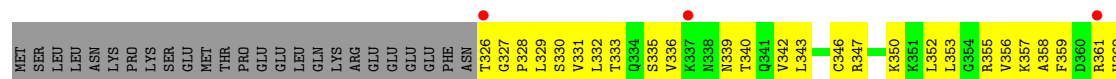
• Molecule 3: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2

Chain J:



• Molecule 3: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2

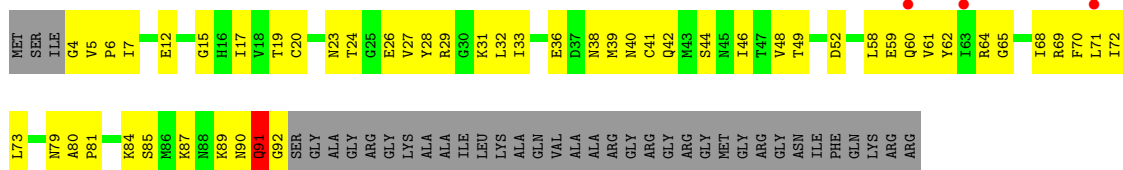
Chain Q:





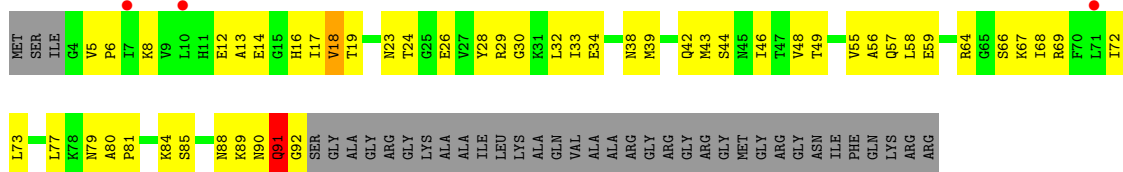
• Molecule 4: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3

Chain D:



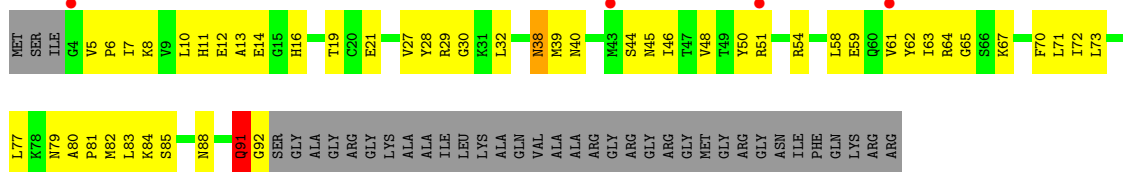
• Molecule 4: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3

Chain K:



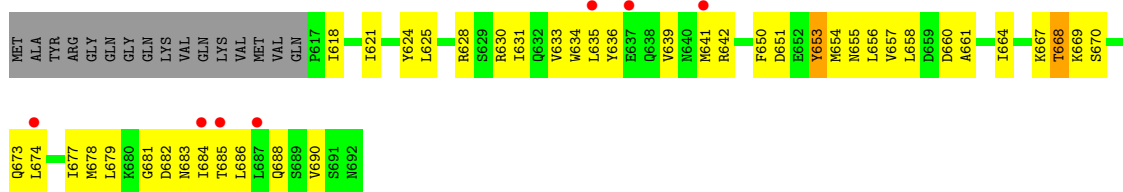
• Molecule 4: SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3

Chain R:



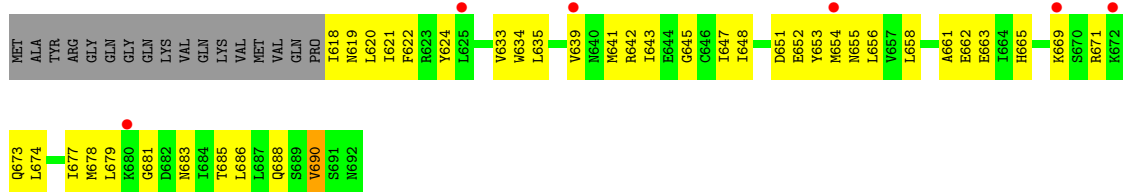
• Molecule 5: SMALL NUCLEAR RIBONUCLEOPROTEIN E

Chain E:



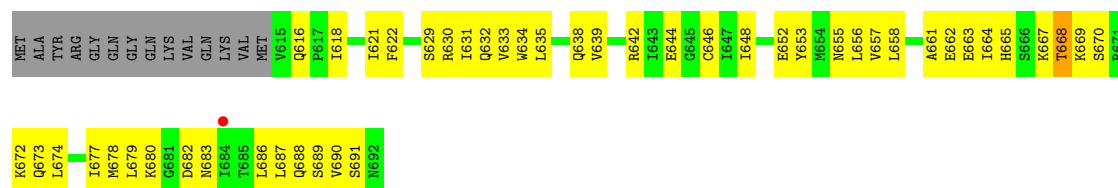
• Molecule 5: SMALL NUCLEAR RIBONUCLEOPROTEIN E

Chain L:



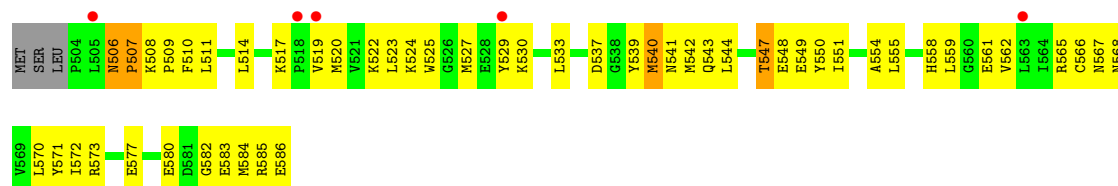
- Molecule 5: SMALL NUCLEAR RIBONUCLEOPROTEIN E

Chain S:



- Molecule 6: SMALL NUCLEAR RIBONUCLEOPROTEIN F

Chain F:



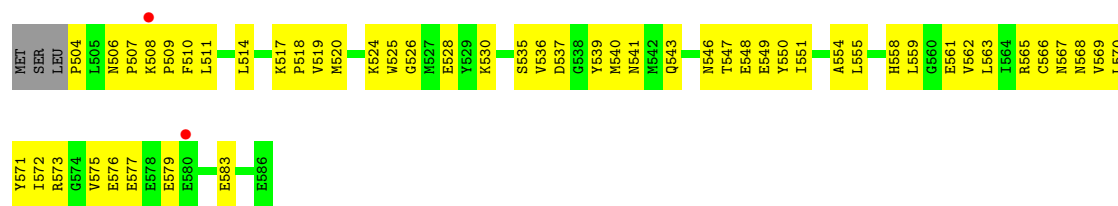
- Molecule 6: SMALL NUCLEAR RIBONUCLEOPROTEIN F

Chain M:



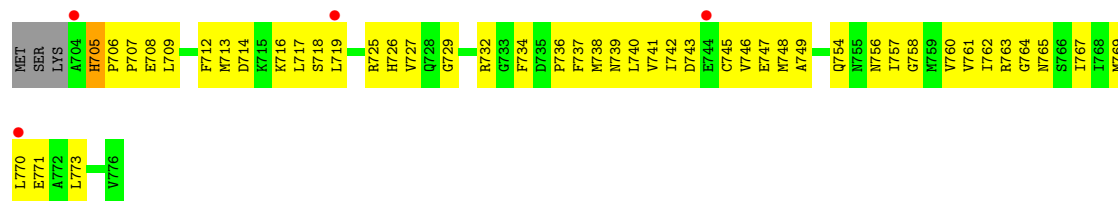
- Molecule 6: SMALL NUCLEAR RIBONUCLEOPROTEIN F

Chain T:



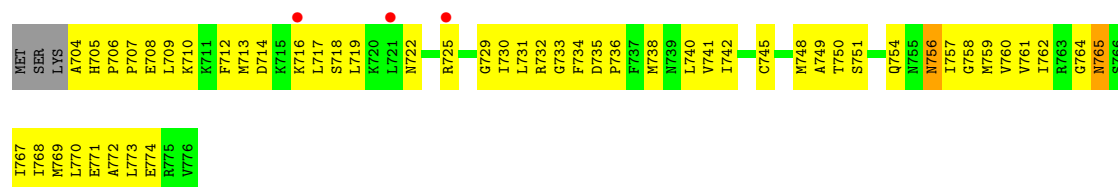
- Molecule 7: SMALL NUCLEAR RIBONUCLEOPROTEIN G

Chain G:



- Molecule 7: SMALL NUCLEAR RIBONUCLEOPROTEIN G

Chain N:



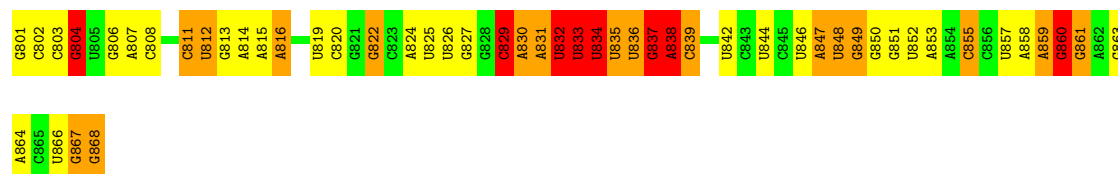
• Molecule 7: SMALL NUCLEAR RIBONUCLEOPROTEIN G

Chain U:



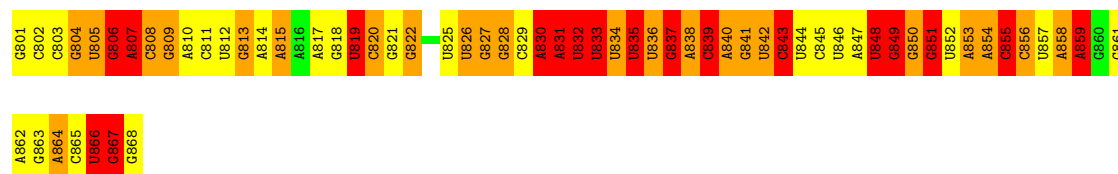
• Molecule 8: HOMO SAPIENS U4A SNRNA

Chain V:



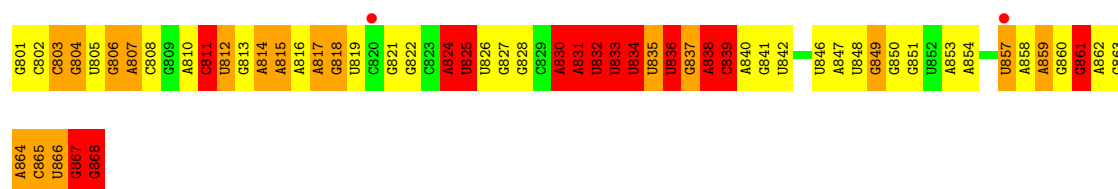
• Molecule 8: HOMO SAPIENS U4A SNRNA

Chain X:



• Molecule 8: HOMO SAPIENS U4A SNRNA

Chain Y:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	248.01Å 248.01Å 251.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	66.17 – 3.60 66.15 – 3.47	Depositor EDS
% Data completeness (in resolution range)	82.9 (66.17-3.60) 60.8 (66.15-3.47)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 3.49Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.277 , 0.321 0.349 , 0.344	Depositor DCC
$R_{free}$ test set	6801 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	78.5	Xtriage
Anisotropy	0.416	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , -20.0	EDS
Estimated twinning fraction	0.250 for H, K, L 0.252 for -H, H+K, -L 0.253 for K, H, -L 0.246 for -H, -K, L 0.357 for -h,-k,l 0.339 for h,-h-k,-l 0.297 for -k,-h,-l	Xtriage
Reported twinning fraction	0.250 for H, K, L 0.252 for -H, H+K, -L 0.253 for K, H, -L 0.246 for -H, -K, L	Depositor
L-test for twinning	$\langle  L  \rangle = 0.31$ , $\langle L^2 \rangle = 0.14$	Xtriage
Outliers	0 of 135989 reflections	Xtriage
$F_o, F_c$ correlation	0.67	EDS
Total number of atoms	18840	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.36% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.63	0/734	0.69	0/978
1	H	0.64	0/769	0.76	0/1024
1	O	0.63	0/743	0.76	0/989
2	B	0.67	0/763	0.77	0/1029
2	I	0.64	0/746	0.77	0/1008
2	P	0.67	0/755	0.78	0/1019
3	C	0.66	0/855	0.75	0/1141
3	J	0.68	0/863	0.77	0/1152
3	Q	0.65	0/718	0.85	1/962 (0.1%)
4	D	0.68	1/704 (0.1%)	0.73	0/946
4	K	0.67	1/704 (0.1%)	0.73	0/946
4	R	0.64	1/704 (0.1%)	0.71	0/946
5	E	0.64	0/637	0.75	0/853
5	L	0.58	0/629	0.71	0/842
5	S	0.59	0/653	0.63	0/876
6	F	0.69	0/669	0.71	0/899
6	M	0.65	0/669	0.75	0/899
6	T	0.62	0/669	0.71	0/899
7	G	0.60	0/573	0.74	0/765
7	N	0.58	0/575	0.74	0/768
7	U	0.68	0/540	0.74	0/723
8	V	1.45	9/1626 (0.6%)	1.86	48/2534 (1.9%)
8	X	1.35	6/1626 (0.4%)	2.03	75/2534 (3.0%)
8	Y	1.40	7/1626 (0.4%)	1.99	64/2534 (2.5%)
All	All	0.90	25/19550 (0.1%)	1.21	188/27266 (0.7%)

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	X	839	C	C3'-O3'	7.55	1.52	1.42
8	Y	832	U	C1'-N1	7.51	1.60	1.48
8	V	832	U	C1'-N1	6.67	1.58	1.48
8	V	827	G	C3'-O3'	6.27	1.50	1.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	V	844	U	C3'-O3'	6.07	1.50	1.42

The worst 5 of 188 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	V	850	G	N3-C2-N2	10.99	127.59	119.90
8	V	829	C	O4'-C1'-N1	9.87	116.09	108.20
8	X	858	A	C1'-O4'-C4'	-9.84	102.03	109.90
8	X	809	G	P-O3'-C3'	9.82	131.49	119.70
8	V	802	C	O4'-C1'-N1	9.59	115.87	108.20

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	724	0	755	72	0
1	H	759	0	796	97	0
1	O	733	0	768	91	0
2	B	754	0	811	132	0
2	I	737	0	789	127	0
2	P	746	0	802	135	0
3	C	845	0	892	126	0
3	J	852	0	900	184	0
3	Q	709	0	764	138	0
4	D	696	0	714	60	0
4	K	696	0	714	71	0
4	R	696	0	714	84	0
5	E	629	0	643	71	0
5	L	622	0	635	85	0
5	S	645	0	659	104	0
6	F	657	0	645	99	0
6	M	657	0	645	125	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	T	657	0	645	85	0
7	G	566	0	583	78	0
7	N	568	0	590	86	0
7	U	533	0	551	77	0
8	V	1453	0	733	45	0
8	X	1453	0	733	106	0
8	Y	1453	0	733	54	0
All	All	18840	0	17214	1844	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 51.

The worst 5 of 1844 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
6:F:547:THR:HG21	6:F:562:VAL:CG2	1.43	1.48
3:J:407:ILE:HG23	6:M:568:ASN:ND2	1.24	1.44
2:I:277:ASP:HB2	2:I:290:LYS:NZ	1.37	1.39
6:F:547:THR:CG2	6:F:562:VAL:HG23	1.62	1.29
3:C:407:ILE:HD13	6:F:525:TRP:CZ2	1.67	1.27

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	88/95 (93%)	84 (96%)	4 (4%)	0	100	100
1	H	93/95 (98%)	85 (91%)	8 (9%)	0	100	100
1	O	89/95 (94%)	79 (89%)	8 (9%)	2 (2%)	10	64
2	B	94/119 (79%)	93 (99%)	1 (1%)	0	100	100
2	I	92/119 (77%)	87 (95%)	4 (4%)	1 (1%)	21	78
2	P	93/119 (78%)	82 (88%)	11 (12%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	103/118 (87%)	92 (89%)	9 (9%)	2 (2%)	12	67
3	J	104/118 (88%)	95 (91%)	7 (7%)	2 (2%)	12	67
3	Q	88/118 (75%)	81 (92%)	7 (8%)	0	100	100
4	D	87/126 (69%)	79 (91%)	7 (8%)	1 (1%)	21	78
4	K	87/126 (69%)	83 (95%)	3 (3%)	1 (1%)	21	78
4	R	87/126 (69%)	83 (95%)	3 (3%)	1 (1%)	21	78
5	E	74/92 (80%)	65 (88%)	7 (10%)	2 (3%)	8	59
5	L	73/92 (79%)	66 (90%)	6 (8%)	1 (1%)	16	72
5	S	76/92 (83%)	69 (91%)	6 (8%)	1 (1%)	18	75
6	F	81/86 (94%)	70 (86%)	8 (10%)	3 (4%)	5	51
6	M	81/86 (94%)	77 (95%)	4 (5%)	0	100	100
6	T	81/86 (94%)	75 (93%)	5 (6%)	1 (1%)	19	77
7	G	71/76 (93%)	65 (92%)	5 (7%)	1 (1%)	16	72
7	N	71/76 (93%)	63 (89%)	8 (11%)	0	100	100
7	U	68/76 (90%)	64 (94%)	4 (6%)	0	100	100
All	All	1781/2136 (83%)	1637 (92%)	125 (7%)	19 (1%)	21	78

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	412	ASN
6	F	540	MET
2	I	295	VAL
4	D	91	GLN
5	E	653	TYR

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	81/85 (95%)	80 (99%)	1 (1%)	82	96
1	H	85/85 (100%)	84 (99%)	1 (1%)	82	96

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	O	82/85 (96%)	81 (99%)	1 (1%)	82	96
2	B	89/101 (88%)	88 (99%)	1 (1%)	84	96
2	I	87/101 (86%)	87 (100%)	0	100	100
2	P	88/101 (87%)	85 (97%)	3 (3%)	49	88
3	C	97/110 (88%)	96 (99%)	1 (1%)	85	97
3	J	98/110 (89%)	97 (99%)	1 (1%)	85	97
3	Q	83/110 (76%)	83 (100%)	0	100	100
4	D	78/101 (77%)	77 (99%)	1 (1%)	80	96
4	K	78/101 (77%)	77 (99%)	1 (1%)	80	96
4	R	78/101 (77%)	77 (99%)	1 (1%)	80	96
5	E	71/84 (84%)	71 (100%)	0	100	100
5	L	70/84 (83%)	70 (100%)	0	100	100
5	S	73/84 (87%)	73 (100%)	0	100	100
6	F	71/74 (96%)	70 (99%)	1 (1%)	78	96
6	M	71/74 (96%)	70 (99%)	1 (1%)	78	96
6	T	71/74 (96%)	71 (100%)	0	100	100
7	G	62/66 (94%)	61 (98%)	1 (2%)	75	95
7	N	63/66 (96%)	61 (97%)	2 (3%)	51	89
7	U	59/66 (89%)	59 (100%)	0	100	100
All	All	1635/1863 (88%)	1618 (99%)	17 (1%)	85	97

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	332	LEU
4	K	18	VAL
2	P	211	SER
1	H	156	SER
2	P	245	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	264	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	L	683	ASN
7	U	726	HIS
4	K	11	HIS
4	K	60	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	V	67/68 (98%)	25 (37%)	5 (7%)
8	X	67/68 (98%)	31 (46%)	7 (10%)
8	Y	67/68 (98%)	33 (49%)	6 (8%)
All	All	201/204 (98%)	89 (44%)	18 (8%)

5 of 89 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	V	804	G
8	V	806	G
8	V	807	A
8	V	808	C
8	V	822	G

5 of 18 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	X	830	A
8	X	836	U
8	Y	830	A
8	X	813	G
8	X	826	U

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	90/95 (94%)	0.48	4 (4%)	33	18	62, 95, 108, 120	0
1	H	95/95 (100%)	0.52	4 (4%)	35	19	65, 97, 111, 116	0
1	O	91/95 (95%)	0.50	1 (1%)	77	50	69, 96, 110, 118	0
2	B	96/119 (80%)	0.61	7 (7%)	15	9	68, 91, 103, 111	0
2	I	94/119 (78%)	0.64	10 (10%)	7	6	64, 92, 102, 123	0
2	P	95/119 (79%)	0.60	5 (5%)	25	14	75, 91, 104, 111	0
3	C	105/118 (88%)	0.51	7 (6%)	17	10	73, 91, 107, 112	0
3	J	106/118 (89%)	0.52	4 (3%)	38	22	67, 92, 107, 124	0
3	Q	90/118 (76%)	0.52	4 (4%)	33	18	65, 87, 105, 119	0
4	D	89/126 (70%)	0.43	3 (3%)	43	24	70, 94, 107, 124	0
4	K	89/126 (70%)	0.48	3 (3%)	43	24	76, 96, 106, 115	0
4	R	89/126 (70%)	0.48	4 (4%)	32	18	72, 96, 106, 120	0
5	E	76/92 (82%)	0.79	7 (9%)	9	7	70, 95, 114, 120	0
5	L	75/92 (81%)	0.65	6 (8%)	12	8	72, 93, 106, 113	0
5	S	78/92 (84%)	0.52	1 (1%)	74	47	72, 96, 112, 124	0
6	F	83/86 (96%)	0.60	5 (6%)	21	12	67, 90, 103, 110	0
6	M	83/86 (96%)	0.52	5 (6%)	21	12	60, 91, 108, 123	0
6	T	83/86 (96%)	0.72	2 (2%)	56	32	66, 88, 112, 120	0
7	G	73/76 (96%)	0.49	4 (5%)	24	13	73, 98, 112, 120	0
7	N	73/76 (96%)	0.39	3 (4%)	35	20	75, 96, 105, 107	0
7	U	70/76 (92%)	0.79	7 (10%)	8	6	76, 98, 108, 117	0
8	V	68/68 (100%)	0.25	0	100	100	57, 79, 97, 106	0
8	X	68/68 (100%)	0.30	0	100	100	50, 79, 94, 110	0
8	Y	68/68 (100%)	0.47	2 (2%)	49	28	53, 82, 101, 110	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
All	All	2027/2340 (86%)	0.54	98 (4%)	29 16	50, 93, 108, 124	0

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	202	LYS	7.3
7	U	716	LYS	6.8
2	B	281	VAL	5.1
7	G	704	ALA	4.7
7	U	740	LEU	4.4

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.