



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 04:25 AM GMT

PDB ID : 3CW1  
Title : Crystal Structure of Human Spliceosomal U1 snRNP  
Authors : Pomeranz Krummel, D.A.; Oubridge, C.; Leung, A.K.; Li, J.; Nagai, K.  
Deposited on : 2008-04-21  
Resolution : 5.49 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>

---

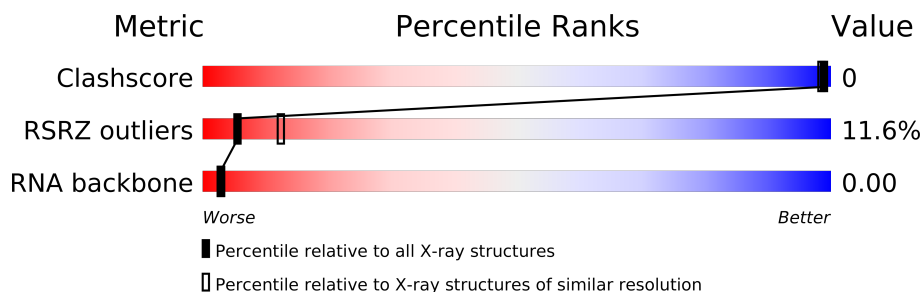
The following versions of software and data (see [references](#)) were used in the production of this report:

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






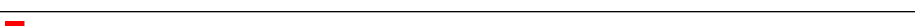









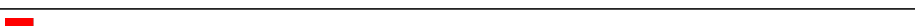
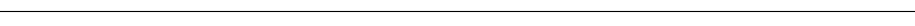
The reported resolution of this entry is 5.49 Å.

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))
Clashscore	79885	1008 (7.40-3.52)
RSRZ outliers	66119	1082 (7.50-3.50)
RNA backbone	1838	1040 (7.50-2.80)

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	V	138	
1	v	138	
1	w	138	
1	x	138	
2	D	126	
2	S	126	
2	T	126	
2	U	126	
3	A	174	
3	H	174	
3	I	174	
3	J	174	
4	B	119	
4	M	119	
4	N	119	
4	O	119	
5	C	118	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	P	118	
5	Q	118	
5	R	118	
6	1	86	
6	2	86	
6	F	86	
6	Z	86	
7	E	92	
7	W	92	
7	X	92	
7	Y	92	
8	3	76	
8	4	76	
8	5	76	
8	G	76	
9	6	216	
9	7	216	
9	8	216	
9	K	216	
10	0	77	
10	9	77	
10	L	77	
10	1	77	

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 3365 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 RNA chain called U1 snRNA.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
1	V	138	Total	P	0	0	138
			138	138			
1	v	138	Total	P	0	0	138
			138	138			
1	w	138	Total	P	0	0	138
			138	138			
1	x	138	Total	P	0	0	138
			138	138			

- Molecule 2 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
2	D	76	Total	C	0	0	76
			76	76			
2	S	76	Total	C	0	0	76
			76	76			
2	T	76	Total	C	0	0	76
			76	76			
2	U	76	Total	C	0	0	76
			76	76			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	66	CYS	SER	CONFLICT	UNP P62318
S	266	CYS	SER	CONFLICT	UNP P62318
T	466	CYS	SER	CONFLICT	UNP P62318
U	666	CYS	SER	CONFLICT	UNP P62318

- Molecule 3 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	A	64	Total C 64 64	0	0	64
3	H	64	Total C 64 64	0	0	64
3	I	63	Total C 63 63	0	0	63
3	J	63	Total C 63 63	0	0	63

- Molecule 4 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	B	77	Total C 77 77	0	0	77
4	M	77	Total C 77 77	0	0	77
4	N	76	Total C 76 76	0	0	76
4	O	77	Total C 77 77	0	0	77

- Molecule 5 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	C	88	Total C 88 88	0	0	88
5	P	86	Total C 86 86	0	0	86
5	Q	89	Total C 89 89	0	0	89
5	R	85	Total C 85 85	0	0	85

- Molecule 6 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
6	F	76	Total C 76 76	0	0	76
6	Z	70	Total C 70 70	0	0	70
6	1	75	Total C 75 75	0	0	75

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
6	2	76	Total C 76 76	0	0	76

- Molecule 7 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
7	E	75	Total C 75 75	0	0	75
7	W	75	Total C 75 75	0	0	75
7	X	75	Total C 75 75	0	0	75
7	Y	75	Total C 75 75	0	0	75

- Molecule 8 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
8	G	73	Total C 73 73	0	0	73
8	3	73	Total C 73 73	0	0	73
8	4	73	Total C 73 73	0	0	73
8	5	73	Total C 73 73	0	0	73

- Molecule 9 is a protein called U1 small nuclear ribonucleoprotein 70 kDa.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
9	K	120	Total C 120 120	0	0	120
9	6	120	Total C 120 120	0	0	120
9	7	120	Total C 120 120	0	0	120
9	8	120	Total C 120 120	0	0	120

- Molecule 10 is a protein called U1 small nuclear ribonucleoprotein C.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
10	L	57	Total C 57 57	0	0	57
10	9	58	Total C 58 58	0	0	58
10	0	54	Total C 54 54	0	0	54
10	1	58	Total C 58 58	0	0	58

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	39	CYS	GLN	CONFLICT	UNP P09234
9	239	CYS	GLN	CONFLICT	UNP P09234
0	439	CYS	GLN	CONFLICT	UNP P09234
1	639	CYS	GLN	CONFLICT	UNP P09234

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

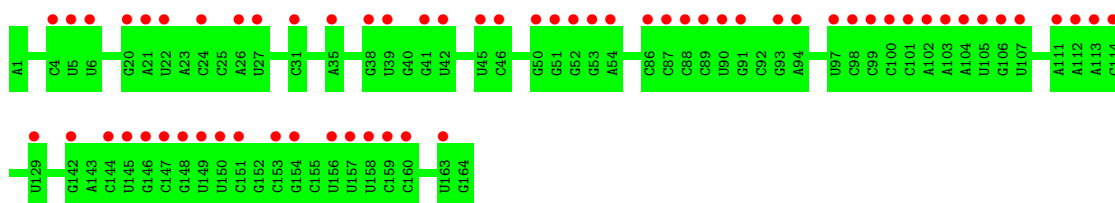
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	0	1	Total Zn 1 1	0	0
11	1	1	Total Zn 1 1	0	0
11	9	1	Total Zn 1 1	0	0
11	L	1	Total Zn 1 1	0	0

### 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.

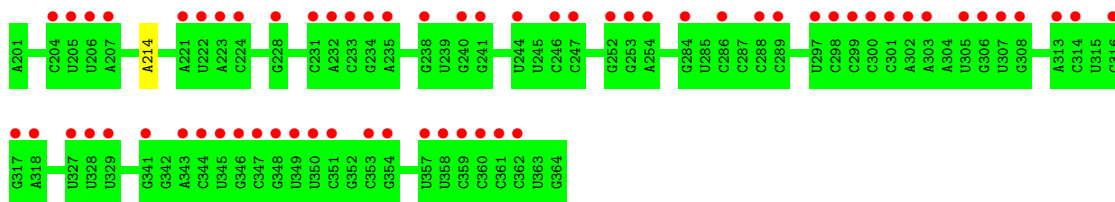
#### • Molecule 1: U1 snRNA

Chain V: 



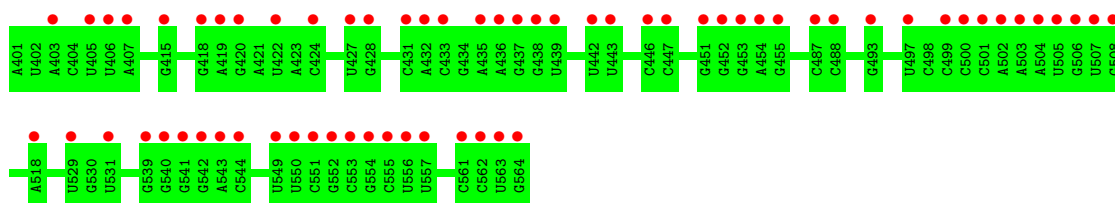
#### • Molecule 1: U1 snRNA

Chain v: 



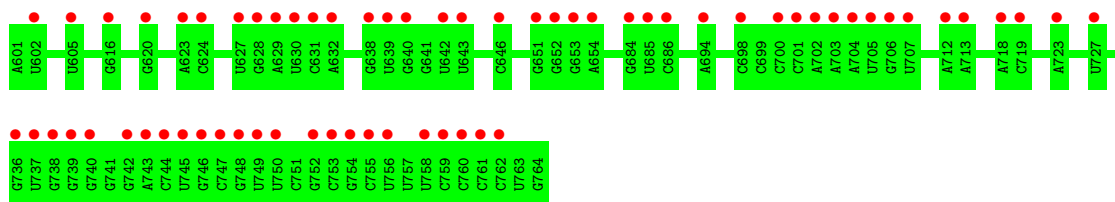
#### • Molecule 1: U1 snRNA

Chain w: 



#### • Molecule 1: U1 snRNA

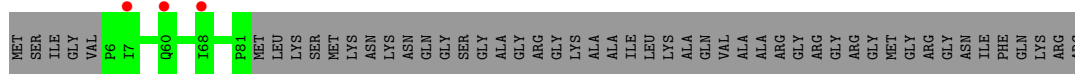
Chain x: 





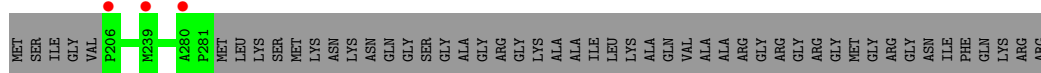
- Molecule 2: Small nuclear ribonucleoprotein Sm D3

Chain D:



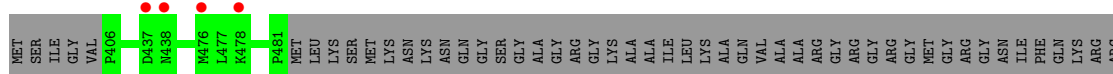
- Molecule 2: Small nuclear ribonucleoprotein Sm D3

Chain S:



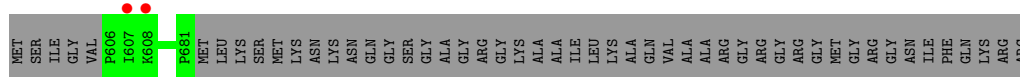
- Molecule 2: Small nuclear ribonucleoprotein Sm D3

Chain T:



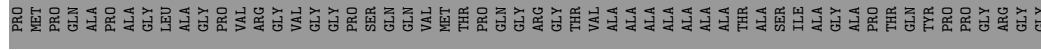
- Molecule 2: Small nuclear ribonucleoprotein Sm D3

Chain U:



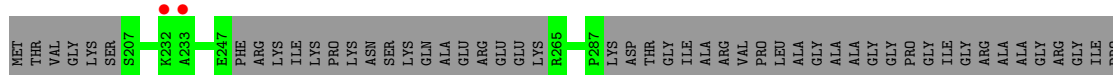
- Molecule 3: Small nuclear ribonucleoprotein-associated proteins B and B

Chain A:



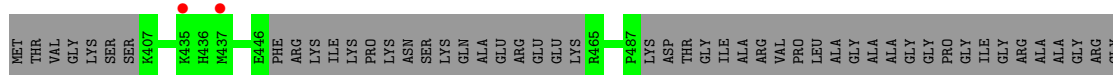
- Molecule 3: Small nuclear ribonucleoprotein-associated proteins B and B

Chain H:



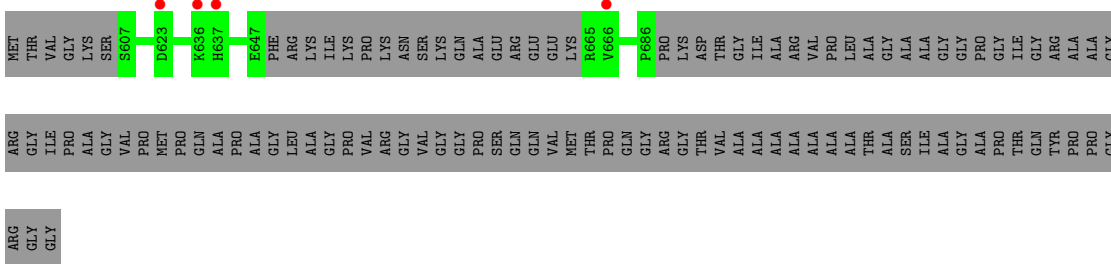
- Molecule 3: Small nuclear ribonucleoprotein-associated proteins B and B

Chain I:



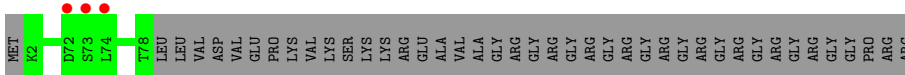
- Molecule 3: Small nuclear ribonucleoprotein-associated proteins B and B

Chain J:



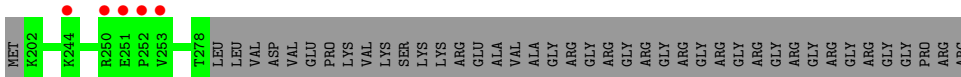
- Molecule 4: Small nuclear ribonucleoprotein Sm D1

Chain B:



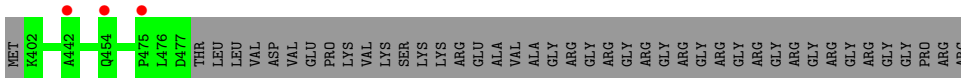
- Molecule 4: Small nuclear ribonucleoprotein Sm D1

Chain M:



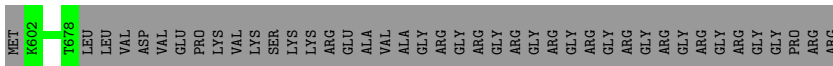
- Molecule 4: Small nuclear ribonucleoprotein Sm D1

Chain N:



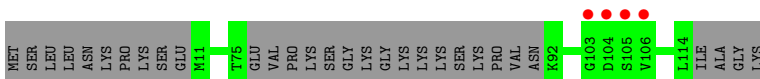
- Molecule 4: Small nuclear ribonucleoprotein Sm D1

Chain 0:



- Molecule 5: Small nuclear ribonucleoprotein Sm D2

Chain C:



- Molecule 5: Small nuclear ribonucleoprotein Sm D2

MET	SER	LEU	LEU	ASN	LYS	PRO	PRO	LYS	SER	GLU	<b>#211</b>	<b>7233</b>	<b>7275</b>	GLU	VAL	PRO	LYS	SER	GLY	GLY	LYS	LYS	LYS	SER	LYS	PRO	VAL	ASN	<b>#292</b>	<b>#312</b>	PRO	LEU	ILE	ALA	ALA	GLY	LYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------------	-------------	-------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------------	-------------	-----	-----	-----	-----	-----	-----	-----

- | Category | Value | Significance |
|----------|-------|--------------|
| MET      |       |              |
| SER      |       |              |
| LEU      |       |              |
| LEU      |       |              |
| ASN      |       |              |
| LYS      |       |              |
| PRO      |       |              |
| LYS      |       |              |
| SER      |       |              |
| GLU      |       |              |
| M411     |       |              |
| S430     |       | Significant  |
| T433     |       | Significant  |
| Q434     |       | Significant  |
| E476     |       |              |
| VAL      |       |              |
| PRO      |       |              |
| LYS      |       |              |
| SER      |       |              |
| GLY      |       |              |
| LYS      |       |              |
| GLY      |       |              |
| LYS      |       |              |
| LYS      |       |              |
| SER      |       |              |
| LYS      |       |              |
| PRO      |       |              |
| VAL      |       |              |
| ASN      |       |              |
| K492     |       |              |
| V506     |       | Significant  |
| N512     |       | Significant  |
| P513     |       | Significant  |
| L514     |       | Significant  |
| ILE      |       |              |
| ALA      |       |              |
| GLY      |       |              |
| LYS      |       |              |

- 
- | Amino Acid | Count |
|------------|-------|
| MET        | 1     |
| SER        | 1     |
| LEU        | 3     |
| LEU        | 1     |
| LEU        | 1     |
| ASN        | 1     |
| LYS        | 3     |
| PRO        | 2     |
| LYS        | 1     |
| LYS        | 1     |
| SER        | 3     |
| SER        | 1     |
| LYS        | 3     |
| PRO        | 2     |
| VAL        | 2     |
| VAL        | 1     |
| ASN        | 1     |
| MET        | 1     |
| P652       | 1     |
| P713       | 1     |
| LEU        | 1     |
| ILE        | 1     |
| ALA        | 1     |
| GLY        | 1     |
| VAL        | 1     |

- MET  
 SER  
 LEU  
 PRO  
 LEU  
 N6  
 E80  
 D81  
 GLY  
 GLU  
 MET  
 ARG  
 GLU

- | Method | LEU | PRO | SER | MET |
|--------|-----|-----|-----|-----|
| M206   | ✓   |     |     |     |
| P218   |     | ✓   | ✓   |     |
| V219   | ✓   |     |     |     |
| M225   | ✓   | ✓   | ✓   |     |
| G226   | ✓   | ✓   | ✓   |     |
| M227   | ✓   |     |     |     |
| G231   | ✓   |     | ✓   |     |
| V232   |     |     |     |     |
| V239   |     |     | ✓   |     |
| M242   | ✓   |     | ✓   |     |
| G243   | ✓   |     | ✓   |     |
| L244   |     |     | ✓   |     |
| G253   |     |     | ✓   |     |
| V262   | ✓   |     |     |     |
| L263   | ✓   |     |     |     |
| L264   | ✓   |     |     |     |
| G274   | ✓   |     | ✓   |     |
| V275   | ✓   |     | ✓   |     |
| GLU    |     |     |     |     |
| GLU    |     |     |     |     |
| GLU    |     |     |     |     |
| GLU    |     |     |     |     |
| GLU    |     |     |     |     |
| ASP    |     |     |     |     |
| GLY    |     |     |     |     |
| GLU    |     |     |     |     |
| MET    |     |     |     |     |
| ARG    |     |     |     |     |
| GLU    |     |     |     |     |

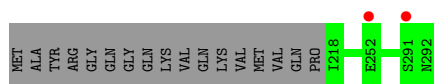
- 

- |     |  |      |      |     |
|-----|--|------|------|-----|
| MET |  | N606 | D681 | GLY |
| SER |  |      |      | GLU |
| LEU |  |      |      | MET |
| PRO |  |      |      | ARG |
| LEU |  |      |      | GLU |

- MET  
ALA  
TYR  
ARG  
GLY  
GLN  
GLY  
GLN  
LYS  
VAL  
GLN  
LYS  
VAL  
MET  
VAL  
GLN  
PRO  
I18  
N92

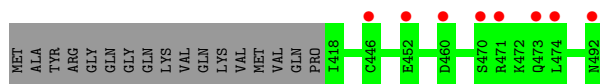
- Molecule 7: Small nuclear ribonucleoprotein E

Chain W: 



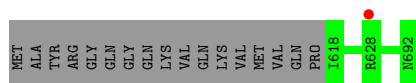
- Molecule 7: Small nuclear ribonucleoprotein E

Chain X: 



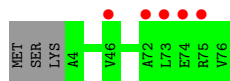
- Molecule 7: Small nuclear ribonucleoprotein E

Chain Y: 



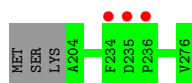
- Molecule 8: Small nuclear ribonucleoprotein G

Chain G: 



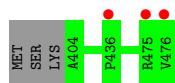
- Molecule 8: Small nuclear ribonucleoprotein G

Chain 3: 



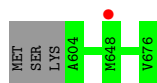
- Molecule 8: Small nuclear ribonucleoprotein G

Chain 4: 



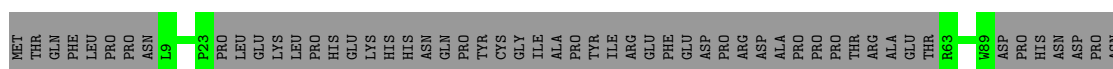
- Molecule 8: Small nuclear ribonucleoprotein G

Chain 5: 



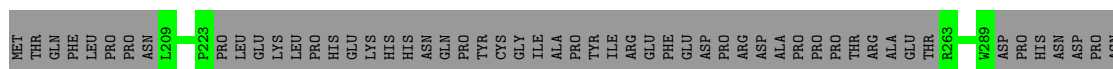
- Molecule 9: U1 small nuclear ribonucleoprotein 70 kDa

Chain K: 



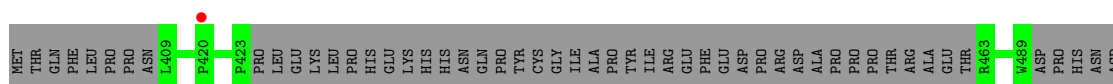
- Molecule 9: U1 small nuclear ribonucleoprotein 70 kDa

Chain 6:



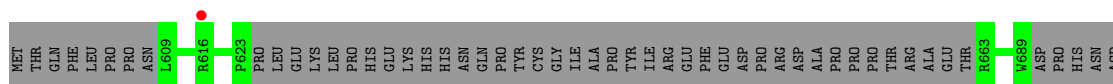
- Molecule 9: U1 small nuclear ribonucleoprotein 70 kDa

Chain 7:



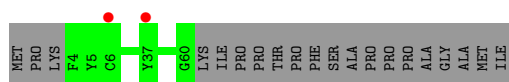
- Molecule 9: U1 small nuclear ribonucleoprotein 70 kDa

Chain 8:



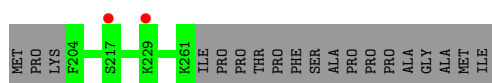
- Molecule 10: U1 small nuclear ribonucleoprotein C

Chain L:



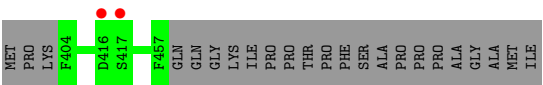
- Molecule 10: U1 small nuclear ribonucleoprotein C

Chain 9:



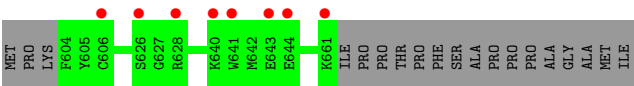
- Molecule 10: U1 small nuclear ribonucleoprotein C

Chain 0: 



● Molecule 10: U1 small nuclear ribonucleoprotein C

Chain 1: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.47Å 127.08Å 152.02Å 95.42° 105.92° 101.80°	Depositor
Resolution (Å)	123.09 – 5.49 122.79 – 5.49	Depositor EDS
% Data completeness (in resolution range)	(Not available) (123.09-5.49) 99.0 (122.79-5.49)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 5.42Å)	Xtriage
Refinement program	O, version 9.0.7	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.495 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	251.7	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	-3.36 , 71.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 28397 reflections	Xtriage
$F_o, F_c$ correlation	0.73	EDS
Total number of atoms	3365	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	231.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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

Bond lengths and bond angles in the following residue types are not validated in this section:  
ZN

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	$\# Z  > 5$

There are no bond length outliers.

There are no bond angle outliers.

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	V	138	0	0	0	0
1	v	138	0	0	0	1
1	w	138	0	0	0	0
1	x	138	0	0	0	0
2	D	76	0	0	0	0
2	S	76	0	0	0	0
2	T	76	0	0	0	0
2	U	76	0	0	0	0
3	A	64	0	0	0	0
3	H	64	0	0	0	0
3	I	63	0	0	0	0
3	J	63	0	0	0	0
4	B	77	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	M	77	0	0	0	0
4	N	76	0	0	0	0
4	O	77	0	0	0	0
5	C	88	0	0	0	0
5	P	86	0	0	0	0
5	Q	89	0	0	0	0
5	R	85	0	0	0	0
6	1	75	0	0	0	0
6	2	76	0	0	0	0
6	F	76	0	0	0	0
6	Z	70	0	0	0	0
7	E	75	0	0	0	0
7	W	75	0	0	0	0
7	X	75	0	0	0	0
7	Y	75	0	0	0	0
8	3	73	0	0	0	0
8	4	73	0	0	0	0
8	5	73	0	0	0	0
8	G	73	0	0	0	0
9	6	120	0	0	0	0
9	7	120	0	0	0	0
9	8	120	0	0	0	0
9	K	120	0	0	0	1
10	0	54	0	0	0	0
10	9	58	0	0	0	0
10	L	57	0	0	0	0
10	1	58	0	0	0	0
11	0	1	0	0	0	0
11	9	1	0	0	0	0
11	L	1	0	0	0	0
11	1	1	0	0	0	0
All	All	3365	0	0	0	1

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 0.

There are no clashes within the asymmetric unit.

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:v:214:A:P	9:K:153:HIS:CA[1_545]	2.05	0.15

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

There are no protein backbone outliers to report in this entry.

### 5.3.2 Protein sidechains ⓘ

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	V	0/138	-	-
1	v	0/138	-	-
1	w	0/138	-	-
1	x	0/138	-	-
All	All	0/552	-	-

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

## 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 ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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(Å <sup>2</sup> )	Q<0.9
1	V	138/138 (100%)	3.04	63 (45%) 1 3	232, 232, 232, 232	0
1	v	138/138 (100%)	2.73	64 (46%) 1 3	232, 232, 232, 232	0
1	w	138/138 (100%)	3.17	65 (47%) 1 3	232, 232, 232, 232	0
1	x	138/138 (100%)	2.70	65 (47%) 1 3	232, 232, 232, 232	0
2	D	76/126 (60%)	0.06	3 (3%) 37 35	232, 232, 232, 232	0
2	S	76/126 (60%)	0.03	3 (3%) 37 35	232, 232, 232, 232	0
2	T	76/126 (60%)	0.08	4 (5%) 25 28	232, 232, 232, 232	0
2	U	76/126 (60%)	0.01	2 (2%) 53 45	232, 232, 232, 232	0
3	A	64/174 (36%)	-0.28	0 100 100	232, 232, 232, 232	0
3	H	64/174 (36%)	-0.01	2 (3%) 47 41	232, 232, 232, 232	0
3	I	63/174 (36%)	-0.08	2 (3%) 45 40	232, 232, 232, 232	0
3	J	63/174 (36%)	0.07	4 (6%) 19 24	232, 232, 232, 232	0
4	B	77/119 (64%)	-0.01	3 (3%) 37 35	232, 232, 232, 232	0
4	M	77/119 (64%)	0.24	5 (6%) 18 24	232, 232, 232, 232	0
4	N	76/119 (63%)	0.04	3 (3%) 37 35	232, 232, 232, 232	0
4	O	77/119 (64%)	0.03	0 100 100	232, 232, 232, 232	0
5	C	88/118 (74%)	0.03	4 (4%) 32 32	232, 232, 232, 232	0
5	P	86/118 (72%)	-0.03	1 (1%) 75 63	232, 232, 232, 232	0
5	Q	89/118 (75%)	0.48	7 (7%) 13 19	232, 232, 232, 232	0
5	R	85/118 (72%)	0.07	5 (5%) 22 26	232, 232, 232, 232	0
6	1	75/86 (87%)	0.22	6 (8%) 12 19	232, 232, 232, 232	0
6	2	76/86 (88%)	-0.23	0 100 100	232, 232, 232, 232	0
6	F	76/86 (88%)	-0.07	1 (1%) 74 62	232, 232, 232, 232	0
6	Z	70/86 (81%)	0.93	16 (22%) 1 5	232, 232, 232, 232	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
7	E	75/92 (81%)	-0.16	0 100 100	232, 232, 232, 232	0
7	W	75/92 (81%)	-0.00	2 (2%) 52 44	232, 232, 232, 232	0
7	X	75/92 (81%)	0.26	8 (10%) 6 14	232, 232, 232, 232	0
7	Y	75/92 (81%)	-0.30	1 (1%) 74 62	232, 232, 232, 232	0
8	3	73/76 (96%)	-0.10	3 (4%) 35 34	232, 232, 232, 232	0
8	4	73/76 (96%)	0.08	3 (4%) 35 34	232, 232, 232, 232	0
8	5	73/76 (96%)	-0.22	1 (1%) 72 60	232, 232, 232, 232	0
8	G	73/76 (96%)	0.17	5 (6%) 17 23	232, 232, 232, 232	0
9	6	120/216 (55%)	0.07	8 (6%) 17 23	232, 232, 232, 232	0
9	7	120/216 (55%)	0.06	6 (5%) 28 29	232, 232, 232, 232	0
9	8	120/216 (55%)	-0.09	4 (3%) 44 40	232, 232, 232, 232	0
9	K	120/216 (55%)	0.08	6 (5%) 28 29	232, 232, 232, 232	0
10	0	54/77 (70%)	0.25	2 (3%) 39 36	232, 232, 232, 232	0
10	9	58/77 (75%)	0.07	2 (3%) 43 39	232, 232, 232, 232	0
10	L	57/77 (74%)	0.21	2 (3%) 42 38	232, 232, 232, 232	0
10	l	58/77 (75%)	0.33	8 (13%) 4 9	232, 232, 232, 232	0
All	All	3361/4888 (68%)	0.53	389 (11%) 5 12	232, 232, 232, 232	0

The worst 5 of 389 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	w	553	C	37.4
1	V	51	G	34.1
1	w	552	G	21.8
1	V	149	U	21.2
5	Q	433	THR	20.2

## 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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
11	ZN	9	278	1/1	0.27	-	231,231,231,231	0
11	ZN	L	78	1/1	0.64	-	231,231,231,231	0
11	ZN	1	678	1/1	0.57	-	231,231,231,231	0
11	ZN	0	478	1/1	0.24	-	231,231,231,231	0

## 6.5 Other polymers

There are no such residues in this entry.