Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 06:56 am BST

PDB ID	:	3PGW
Title	:	Crystal structure of human U1 snRNP
Authors	:	Weber, G.; Trowitzsch, S.; Kastner, B.; Luehrmann, R.; Wahl, M.C.
Deposited on	:	2010-11-02
Resolution	:	4.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ①) were used in the production of this report:

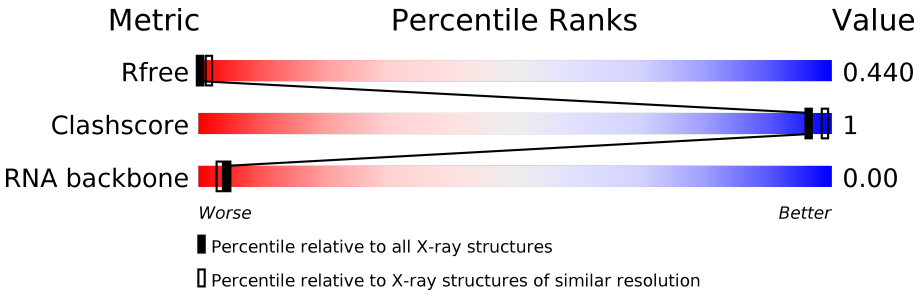
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.40 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}	130704	1043 (5.00-3.80)
Clashscore	141614	1111 (5.00-3.80)
RNA backbone	3102	1058 (5.60-3.00)

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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	282	38% 62%
1	P	282	40% 60%
2	L	437	34% 66%
2	S	437	34% 66%
3	W	126	73% 27%
3	Z	126	63% 37%
4	B	231	35% 65%
4	Q	231	33% 67%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	U	119	71% 29%
5	X	119	71% 29%
6	V	118	86% 13%
6	Y	118	92% 6%
7	F	86	99%
7	I	86	86% 14%
8	E	92	91% 9%
8	H	92	97%
9	G	76	100%
9	J	76	100%
10	N	164	99%
10	R	164	99%
11	D	9	89% 11%
11	M	9	89% 11%

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 2058 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 U1-A.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
1	A	108	Total C 108 108	0	0	108
1	P	113	Total C 113 113	0	0	113

- Molecule 2 is a protein called U1-70K.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	S	150	Total C 150 150	0	0	150
2	L	148	Total C 148 148	0	0	148

- Molecule 3 is a protein called Sm-D3.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	Z	80	Total C 80 80	0	0	80
3	W	92	Total C 92 92	0	0	92

- Molecule 4 is a protein called Sm B.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	B	81	Total C 81 81	0	0	81
4	Q	76	Total C 76 76	0	0	76

- Molecule 5 is a protein called Sm-D1.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	X	85	Total C 85 85	0	0	85
5	U	85	Total C 85 85	0	0	85

- Molecule 6 is a protein called Sm-D2.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
6	Y	111	Total C 111 111	7	0	111
6	V	103	Total C 103 103	0	0	103

- Molecule 7 is a protein called Sm-F.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
7	F	85	Total C 85 85	0	0	85
7	I	74	Total C 74 74	0	0	74

- Molecule 8 is a protein called Sm-E.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
8	E	84	Total C 84 84	0	0	84
8	H	89	Total C 89 89	0	0	89

- Molecule 9 is a protein called Sm G.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
9	G	76	Total C 76 76	0	0	76
9	J	76	Total C 76 76	0	0	76

- Molecule 10 is a RNA chain called U1 snRNA.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
10	R	163	Total P 163 163	0	0	163

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
10	N	163	Total P 163 163	0	0	163

- Molecule 11 is a DNA chain called DNA 5'-D(*AP*GP*GP*TP*AP*AP*GP*TP*A)-3'.

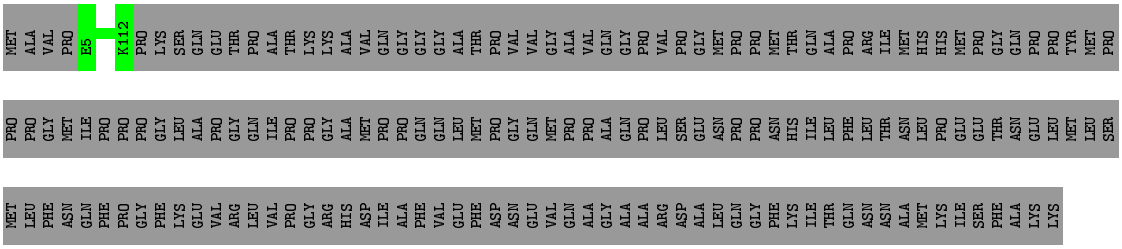
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
11	D	8	Total P 8 8	0	0	8
11	M	8	Total P 8 8	0	0	8

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

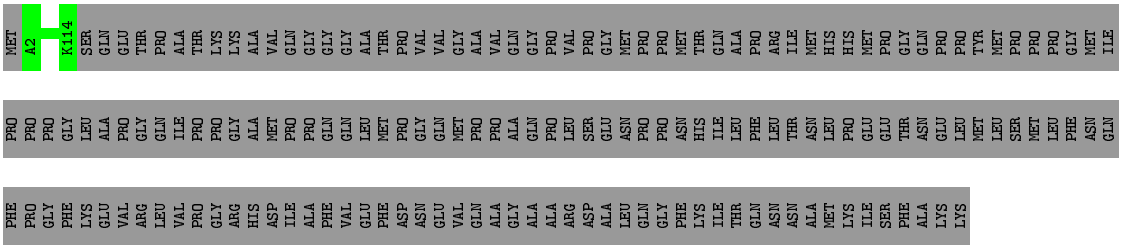
- Molecule 1: U1-A

Chain A:



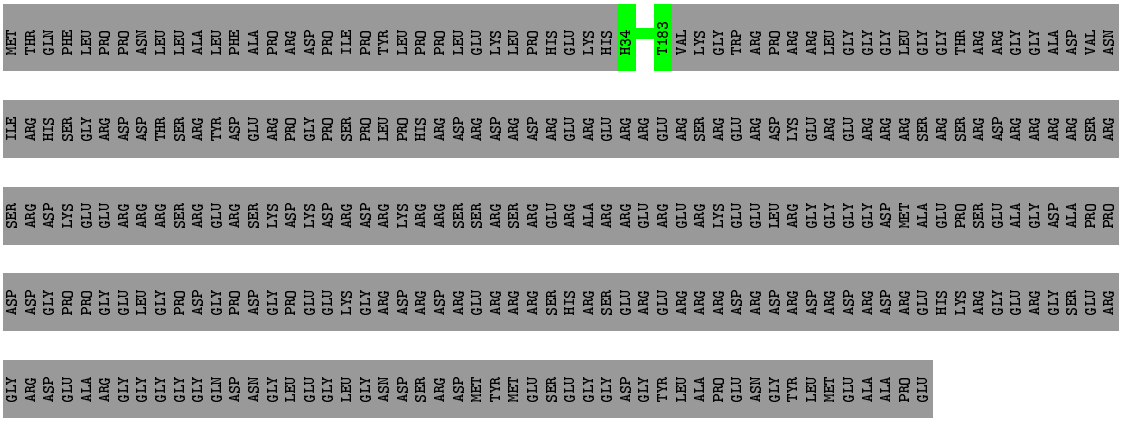
- Molecule 1: U1-A

Chain P:



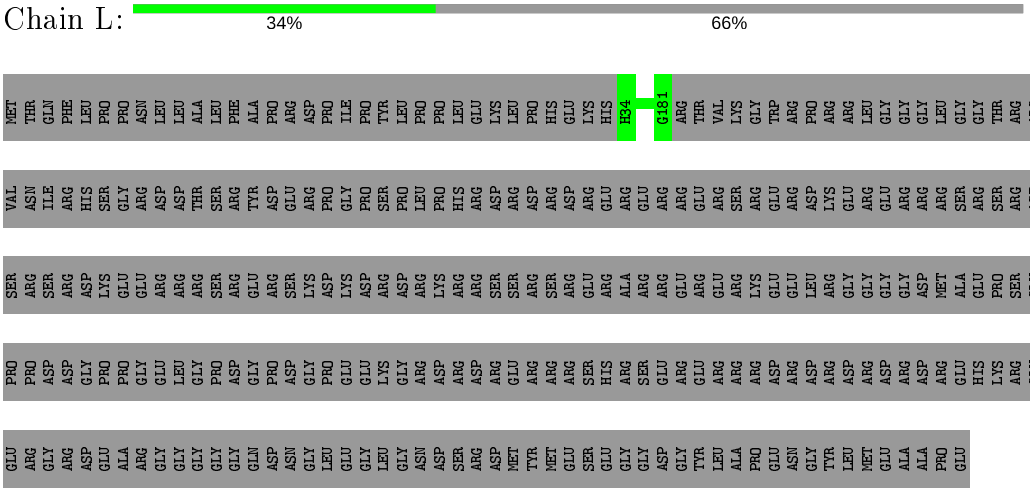
- Molecule 2: U1-70K

Chain S:



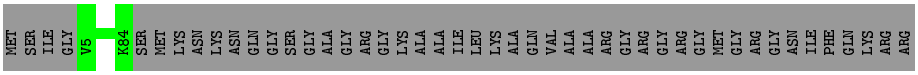
- Molecule 2: U1-70K

Chain L:



- Molecule 3: Sm-D3

Chain Z:

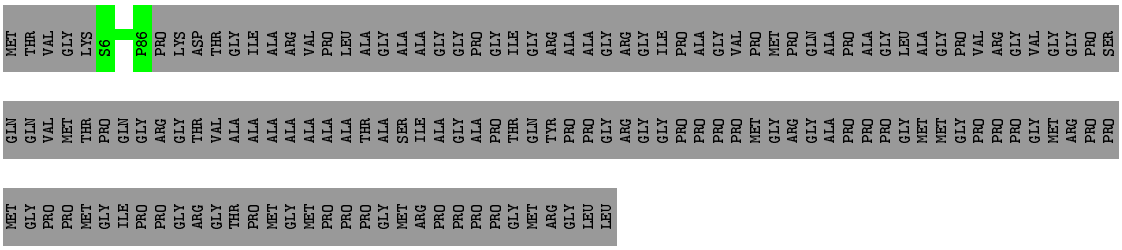


- Molecule 3: Sm-D3

Chain W:

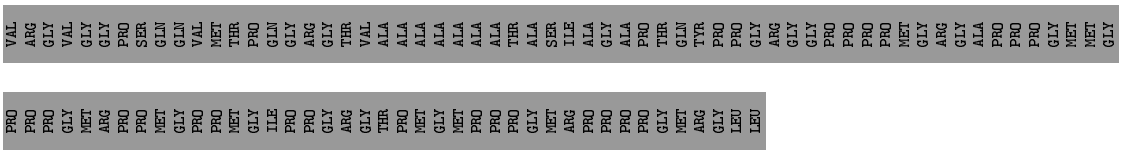
- Molecule 4: Sm B

Chain B:



- Molecule 4: Sm B

Chain Q:



- Molecule 5: Sm-D1

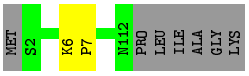
Chain X:

- Molecule 5: Sm-D1

Chain U:

- Molecule 6: Sm-D2

Chain Y:

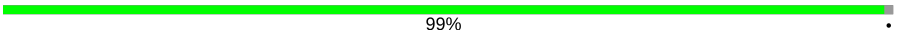


- Molecule 6: Sm-D2

Chain V:

- Molecule 7: Sm-F

Chain F:

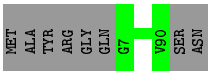


- Molecule 7: Sm-F

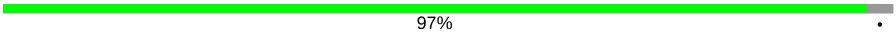
Chain I:

- Molecule 8: Sm-E

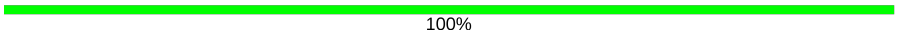
Chain E:



• Molecule 8: Sm-E

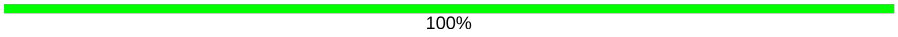
Chain H: 

• Molecule 9: Sm G

Chain G: 

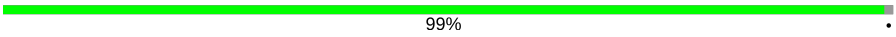
There are no outlier residues recorded for this chain.

• Molecule 9: Sm G

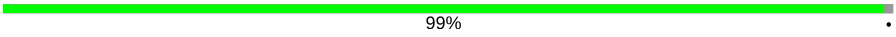
Chain J: 

There are no outlier residues recorded for this chain.

• Molecule 10: U1 snRNA

Chain R: 

• Molecule 10: U1 snRNA

Chain N: 

• Molecule 11: DNA 5'-D(*AP*GP*GP*TP*AP*AP*GP*TP*A)-3'

Chain D: 

• Molecule 11: DNA 5'-D(*AP*GP*GP*TP*AP*AP*GP*TP*A)-3'

Chain M: 

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	358.42Å 88.22Å 150.90Å 90.00° 111.88° 90.00°	Depositor
Resolution (Å)	100.00 – 4.40 140.03 – 4.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (100.00-4.40) 99.5 (140.03-4.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 4.47Å)	Xtriage
Refinement program	PHENIX, CNS 1.3	Depositor
R, R_{free}	0.299 , 0.348 0.440 , 0.440	Depositor DCC
R_{free} test set	1404 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	222.6	Xtriage
Anisotropy	0.203	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.54 , 0.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.037 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.70	EDS
Total number of atoms	2058	wwPDB-VP
Average B, all atoms (Å^2)	255.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	108	0	0	0	0
1	P	113	0	0	0	0
2	L	148	0	0	0	0
2	S	150	0	0	0	0
3	W	92	0	0	0	0
3	Z	80	0	0	0	0
4	B	81	0	0	0	0
4	Q	76	0	0	0	0
5	U	85	0	0	0	0
5	X	85	0	0	0	0
6	V	103	0	0	1	0
6	Y	111	0	0	1	0
7	F	85	0	0	0	0
7	I	74	0	0	0	0
8	E	84	0	0	0	0
8	H	89	0	0	0	0
9	G	76	0	0	0	0
9	J	76	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	N	163	0	0	0	0
10	R	163	0	0	0	0
11	D	8	0	0	0	0
11	M	8	0	0	0	0
All	All	2058	0	0	2	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (2) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:V:6:LYS:CA	6:V:7:PRO:CA	2.81	0.58
6:Y:6:LYS:CA	6:Y:7:PRO:CA	2.83	0.56

There are no symmetry-related clashes.

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
10	N	0/164	-	-
10	R	0/164	-	-
All	All	0/328	-	-

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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.