Full wwPDB X-ray Structure Validation Report (i)

May 4, 2024 – 06:14 pm BST

PDB ID	:	6SR7
Title	:	Structure of the U1A variant A1-98 Y31H/Q36R/K98W
Authors	:	Rosenbach, H.; Span, I.
Deposited on	:	2019-09-05
Resolution	:	1.86 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

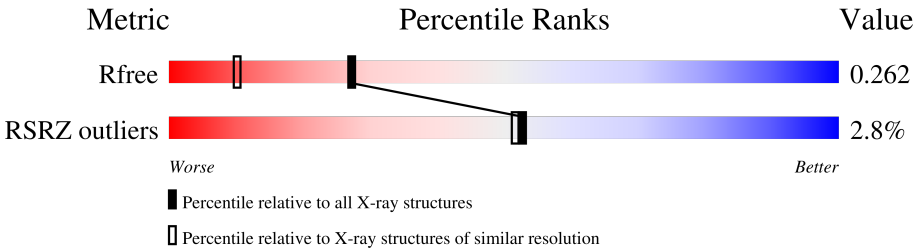
MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 1.86 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	2469 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6015 atoms, of which 3020 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.

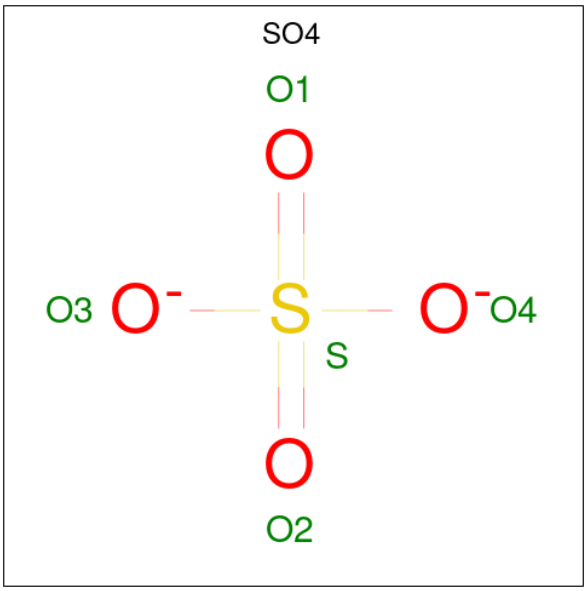
- Molecule 1 is a protein called U1 small nuclear ribonucleoprotein A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	91	1533	486	780	132	131	4	38	1	0
1	BBB	91	1519	481	773	130	131	4	36	0	0
1	CCC	91	1519	481	773	130	131	4	36	0	0
1	DDD	81	1358	429	694	118	114	3	34	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	31	HIS	TYR	engineered mutation	UNP P09012
AAA	36	ARG	GLN	engineered mutation	UNP P09012
AAA	98	TRP	LYS	engineered mutation	UNP P09012
BBB	31	HIS	TYR	engineered mutation	UNP P09012
BBB	36	ARG	GLN	engineered mutation	UNP P09012
BBB	98	TRP	LYS	engineered mutation	UNP P09012
CCC	31	HIS	TYR	engineered mutation	UNP P09012
CCC	36	ARG	GLN	engineered mutation	UNP P09012
CCC	98	TRP	LYS	engineered mutation	UNP P09012
DDD	31	HIS	TYR	engineered mutation	UNP P09012
DDD	36	ARG	GLN	engineered mutation	UNP P09012
DDD	98	TRP	LYS	engineered mutation	UNP P09012

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	CCC	1	5	4	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	AAA	24	24	24	0	0
			Total	O		
3	BBB	19	19	19	0	0
			Total	O		
3	CCC	23	23	23	0	0
			Total	O		
3	DDD	15	15	15	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	76.26Å 76.26Å 151.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.11 – 1.86 68.11 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.9 (68.11-1.86) 100.0 (68.11-1.86)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, R_{free}	0.221 , 0.256 0.229 , 0.262	Depositor DCC
R_{free} test set	1920 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	25.4	Xtriage
Anisotropy	0.103	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 30.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6015	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.04% 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.

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	CCC	101	-	4,4,4	0.21	0	6,6,6	0.16	0

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.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

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, 95th 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^2)	Q<0.9
1	AAA	91/98 (92%)	0.55	5 (5%)	25	24	20, 28, 47, 49	0
1	BBB	91/98 (92%)	0.40	1 (1%)	80	81	18, 27, 45, 64	0
1	CCC	91/98 (92%)	0.30	1 (1%)	80	81	17, 24, 43, 55	0
1	DDD	81/98 (82%)	0.43	3 (3%)	41	39	19, 27, 61, 82	0
All	All	354/392 (90%)	0.42	10 (2%)	53	52	17, 27, 48, 82	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	93	ILE	5.2
1	DDD	49	LEU	4.4
1	AAA	94	ILE	4.4
1	DDD	50	LYS	4.1
1	AAA	92	ASP	4.0
1	CCC	49	LEU	3.6
1	DDD	52	ARG	3.1
1	AAA	95	ALA	3.1
1	AAA	90	ASP	2.6
1	BBB	88	LYS	2.5

5.2 Non-standard residues in protein, DNA, RNA chains (i)

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

5.3 Carbohydrates (i)

There are no monosaccharides in this entry.

5.4 Ligands (i)

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. The B-factors column lists the minimum, median, 95th 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	CCC	101	5/5	0.95	0.11	37,39,40,42	0

5.5 Other polymers (i)

There are no such residues in this entry.