



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 07:19 am BST

PDB ID : 5K1B
Title : Crystal structure of the UAF1/USP12 complex in F222 space group
Authors : Li, H.; D'Andrea, A.D.; Zheng, N.
Deposited on : 2016-05-18
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ](#)) 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

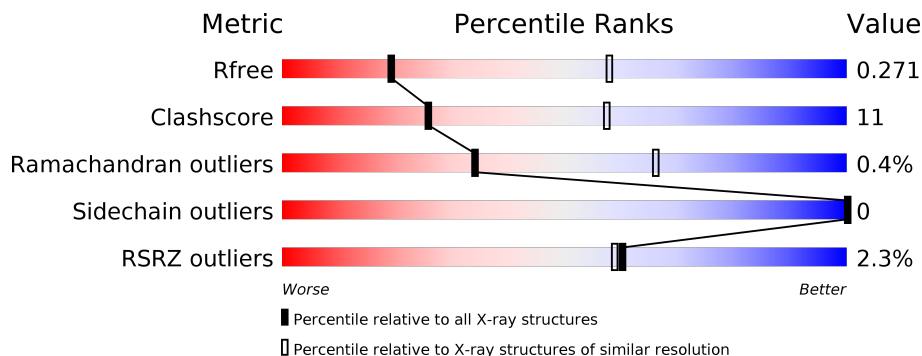
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	B	677	
2	A	367	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ZN	A	401	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6366 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 WD repeat-containing protein 48.

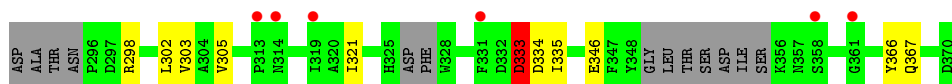
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	579	4396	2783	756	830	27	0	0	0

- Molecule 2 is a protein called Ubiquitin carboxyl-terminal hydrolase 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	264	1969	1266	322	367	14	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		



4 Data and refinement statistics

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	120.96Å 158.93Å 234.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 3.30 44.52 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.0 (50.01-3.30) 89.9 (44.52-3.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.49 (at 3.19Å)	Xtrriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.235 , 0.286 0.237 , 0.271	Depositor DCC
R_{free} test set	859 reflections (4.64%)	wwPDB-VP
Wilson B-factor (Å ²)	111.0	Xtrriage
Anisotropy	0.335	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 89.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6366	wwPDB-VP
Average B, all atoms (Å ²)	135.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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$
1	B	0.51	0/4482	0.75	1/6105 (0.0%)
2	A	0.72	6/2003 (0.3%)	0.77	3/2723 (0.1%)
All	All	0.59	6/6485 (0.1%)	0.76	4/8828 (0.0%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	237	ARG	CZ-NH2	11.94	1.48	1.33
2	A	237	ARG	CZ-NH1	10.20	1.46	1.33
2	A	334	ASP	C-O	7.73	1.38	1.23
2	A	237	ARG	NE-CZ	6.43	1.41	1.33
2	A	333	ASP	C-N	6.06	1.48	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	307	ASP	C-N-CD	5.73	140.43	128.40
2	A	226	CYS	N-CA-C	5.41	125.61	111.00
2	A	333	ASP	N-CA-C	5.41	125.60	111.00
2	A	334	ASP	CB-CG-OD2	5.18	122.96	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	B	4396	0	4199	92	0
2	A	1969	0	1785	38	0
3	A	1	0	0	2	0
All	All	6366	0	5984	130	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 11.

The worst 5 of 130 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)
1:B:250:VAL:HG12	1:B:277:TYR:CD2	1.97	1.00
1:B:93:ALA:HB2	1:B:123:LEU:HD11	1.46	0.95
1:B:307:ASP:O	1:B:309:PRO:N	2.04	0.90
1:B:93:ALA:CB	1:B:123:LEU:HD11	2.02	0.89
1:B:307:ASP:O	1:B:309:PRO:CD	2.25	0.84

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	B	559/677 (83%)	495 (89%)	62 (11%)	2 (0%)	34 66
2	A	246/367 (67%)	214 (87%)	31 (13%)	1 (0%)	34 66
All	All	805/1044 (77%)	709 (88%)	93 (12%)	3 (0%)	34 66

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	333	ASP
1	B	309	PRO
1	B	308	PRO

5.3.2 Protein sidechains [i](#)

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	B	463/597 (78%)	463 (100%)	0	100	100
2	A	195/335 (58%)	195 (100%)	0	100	100
All	All	658/932 (71%)	658 (100%)	0	100	100

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

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	43	ASN
1	B	60	ASN
1	B	468	GLN
1	B	478	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is 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.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 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	B	579/677 (85%)	-0.07	12 (2%) 63 62	91, 128, 161, 182	0
2	A	264/367 (71%)	-0.09	7 (2%) 54 52	81, 150, 185, 208	0
All	All	843/1044 (80%)	-0.08	19 (2%) 60 59	81, 135, 172, 208	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	312	ILE	4.9
2	A	319	ILE	4.8
1	B	313	TRP	4.3
1	B	569	LEU	3.6
1	B	311	ALA	3.5

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 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(\AA^2)	Q<0.9
3	ZN	A	401	1/1	0.97	0.07	218,218,218,218	0

6.5 Other polymers [i](#)

There are no such residues in this entry.