



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 19, 2020 – 09:08 AM BST

PDB ID : 6XJL
Title : Structure of the SM protein Vps45
Authors : Jeffrey, P.D.; Shimamura, G.S.; Allen, F.; Hughson, F.M.
Deposited on : 2020-06-24
Resolution : 2.00 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.14.6
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) : 2.14.6

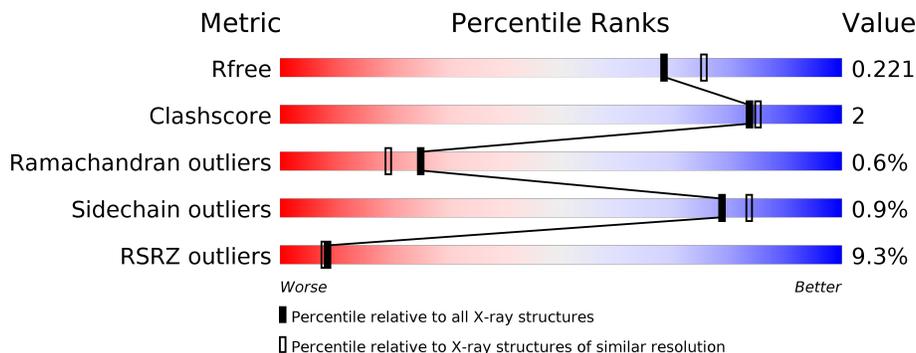
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 $\leq 5\%$. 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	A	649	

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 4485 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 Vps45.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	536	4215	2676	742	780	17	0	2	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	HIS	-	expression tag	UNP G0S539
A	-23	HIS	-	expression tag	UNP G0S539
A	-22	HIS	-	expression tag	UNP G0S539
A	-21	HIS	-	expression tag	UNP G0S539
A	-20	HIS	-	expression tag	UNP G0S539
A	-19	HIS	-	expression tag	UNP G0S539
A	-18	HIS	-	expression tag	UNP G0S539
A	-17	HIS	-	expression tag	UNP G0S539
A	-16	GLY	-	expression tag	UNP G0S539
A	-15	SER	-	expression tag	UNP G0S539
A	-14	ASP	-	expression tag	UNP G0S539
A	-13	TYR	-	expression tag	UNP G0S539
A	-12	ASP	-	expression tag	UNP G0S539
A	-11	ILE	-	expression tag	UNP G0S539
A	-10	PRO	-	expression tag	UNP G0S539
A	-9	THR	-	expression tag	UNP G0S539
A	-8	THR	-	expression tag	UNP G0S539
A	-7	GLU	-	expression tag	UNP G0S539
A	-6	ASN	-	expression tag	UNP G0S539
A	-5	LEU	-	expression tag	UNP G0S539
A	-4	TYR	-	expression tag	UNP G0S539
A	-3	PHE	-	expression tag	UNP G0S539
A	-2	GLN	-	expression tag	UNP G0S539
A	-1	GLY	-	expression tag	UNP G0S539
A	0	SER	-	expression tag	UNP G0S539

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	270	Total 270	O 270	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	68.49Å 93.96Å 102.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.65 – 2.00 34.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.4 (34.65-2.00) 99.5 (34.65-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.25 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.17_3644	Depositor
R, R_{free}	0.179 , 0.221 0.179 , 0.221	Depositor DCC
R_{free} test set	2268 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	29.4	Xtrriage
Anisotropy	0.629	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 58.3	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	0.96	EDS
Total number of atoms	4485	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.68% 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](#)

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	0.37	0/4300	0.53	0/5820

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 [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	A	4215	0	4260	20	0
2	A	270	0	0	2	0
All	All	4485	0	4260	20	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 2.

All (20) 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:A:493:ASN:O	1:A:495:TYR:N	2.22	0.73
1:A:156:ARG:NH1	1:A:597:ASP:OD1	2.29	0.65
1:A:277:ILE:HG23	1:A:344:LYS:HD2	1.77	0.64
1:A:302:ILE:HD11	1:A:345:HIS:HB3	1.91	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:484:ARG:HG3	1:A:503:GLU:OE2	2.09	0.53
1:A:335:PHE:CE1	1:A:338:LEU:HD23	2.44	0.53
1:A:187:ARG:NH1	2:A:712:HOH:O	2.44	0.51
1:A:493:ASN:HD21	1:A:497:GLN:HG2	1.76	0.49
1:A:299:GLY:O	1:A:303:LYS:HD2	2.14	0.47
1:A:145:LEU:HA	1:A:580:LEU:O	2.15	0.46
1:A:146:PHE:CZ	1:A:581:GLY:HA3	2.51	0.46
1:A:303:LYS:O	1:A:306:VAL:HG12	2.16	0.45
1:A:220:ARG:NE	2:A:703:HOH:O	2.34	0.45
1:A:207:ARG:HD2	1:A:522:VAL:HG11	1.99	0.44
1:A:499:SER:OG	1:A:503:GLU:OE1	2.25	0.43
1:A:275:LYS:HB3	1:A:275:LYS:HE2	1.82	0.43
1:A:113:LYS:HD3	1:A:113:LYS:HA	1.84	0.42
1:A:76:ARG:HD3	1:A:104:HIS:CE1	2.56	0.41
1:A:379:SER:HA	1:A:382:LYS:HE2	2.03	0.40
1:A:207:ARG:HD2	1:A:522:VAL:CG1	2.52	0.40

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	A	528/649 (81%)	520 (98%)	5 (1%)	3 (1%)	25 19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	493	ASN
1	A	494	VAL
1	A	492	GLU

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	A	468/557 (84%)	464 (99%)	4 (1%)	78 83

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-2	GLN
1	A	85	LEU
1	A	261	ARG
1	A	496	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

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, 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	A	536/649 (82%)	0.37	50 (9%) 8 8	21, 35, 82, 115	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	309	TYR	6.9
1	A	335	PHE	6.9
1	A	313	THR	6.5
1	A	337	LYS	5.7
1	A	311	SER	5.6
1	A	312	ARG	5.6
1	A	308	GLN	5.4
1	A	609	GLY	5.2
1	A	314	LYS	5.0
1	A	315	SER	4.9
1	A	333	PRO	4.9
1	A	430	VAL	4.7
1	A	339	SER	3.8
1	A	426	ALA	3.8
1	A	68	ASN	3.8
1	A	334	GLU	3.8
1	A	433	ARG	3.8
1	A	488	LEU	3.7
1	A	124	ASP	3.6
1	A	487	GLY	3.5
1	A	122	ALA	3.4
1	A	493	ASN	3.4
1	A	338	LEU	3.4
1	A	513	ARG	3.3
1	A	449	HIS	3.3
1	A	27	PRO	3.2
1	A	306	VAL	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	67	GLN	3.0
1	A	310	GLN	3.0
1	A	494	VAL	2.9
1	A	304	ASP	2.9
1	A	547	PRO	2.7
1	A	342	VAL	2.7
1	A	427	ALA	2.6
1	A	305	TYR	2.6
1	A	121	GLU	2.5
1	A	-3	PHE	2.5
1	A	316	THR	2.5
1	A	495	TYR	2.4
1	A	446[A]	ARG	2.4
1	A	117	GLU	2.3
1	A	482	GLY	2.3
1	A	307	GLU	2.3
1	A	496	THR	2.2
1	A	85	LEU	2.2
1	A	491	VAL	2.2
1	A	517	SER	2.1
1	A	271	ARG	2.1
1	A	515	ARG	2.1
1	A	285	PRO	2.0

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.