



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 07:16 am BST

PDB ID : 5JW7  
Title : Crystal structure of SopA-Trim56 complex  
Authors : Bhogaraju, S.; Dikic, I.  
Deposited on : 2016-05-11  
Resolution : 2.85 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](mailto: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

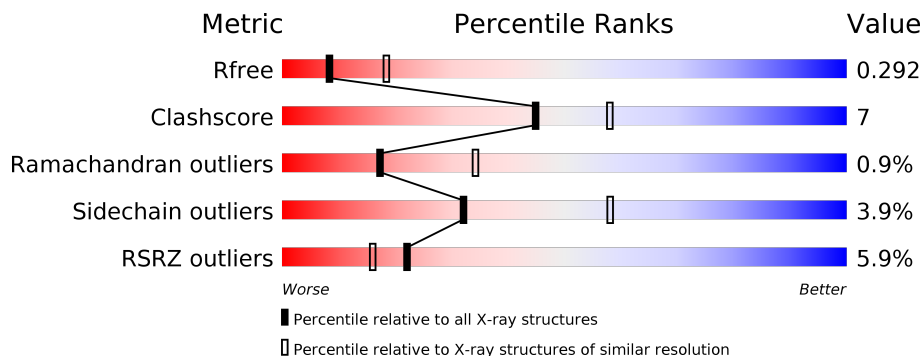
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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	1031 (2.86-2.82)
Clashscore	141614	1078 (2.86-2.82)
Ramachandran outliers	138981	1050 (2.86-2.82)
Sidechain outliers	138945	1051 (2.86-2.82)
RSRZ outliers	127900	1019 (2.86-2.82)

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 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	282	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 80%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 12%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 7%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 19%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">2%      80%      12%      • 7%</p>
2	B	93	<div style="display: flex; align-items: center;"> <div style="width: 13%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 48%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 24%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">13%      48%      13%      • •      34%</p>

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 4610 atoms, of which 2215 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 E3 ubiquitin-protein ligase SopA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	262	3845	1254	1856	331	396	8	0	1	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	148	GLY	-	expression tag	UNP Q8ZNR3
A	149	SER	-	expression tag	UNP Q8ZNR3
A	150	GLY	-	expression tag	UNP Q8ZNR3
A	151	SER	-	expression tag	UNP Q8ZNR3
A	152	GLU	-	expression tag	UNP Q8ZNR3
A	153	ASN	-	expression tag	UNP Q8ZNR3
A	154	LEU	-	expression tag	UNP Q8ZNR3
A	155	TYR	-	expression tag	UNP Q8ZNR3
A	156	PHE	-	expression tag	UNP Q8ZNR3
A	157	GLN	-	expression tag	UNP Q8ZNR3
A	158	GLY	-	expression tag	UNP Q8ZNR3
A	159	GLY	-	expression tag	UNP Q8ZNR3
A	160	SER	-	expression tag	UNP Q8ZNR3
A	161	GLY	-	expression tag	UNP Q8ZNR3
A	162	SER	-	expression tag	UNP Q8ZNR3
A	426	LEU	-	expression tag	UNP Q8ZNR3
A	427	GLU	-	expression tag	UNP Q8ZNR3
A	428	ALA	-	expression tag	UNP Q8ZNR3
A	429	LEU	-	expression tag	UNP Q8ZNR3

- Molecule 2 is a protein called E3 ubiquitin-protein ligase TRIM56.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	61	762	251	359	67	78	7	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	Zn 2	0	0


- Molecule 4 is water.

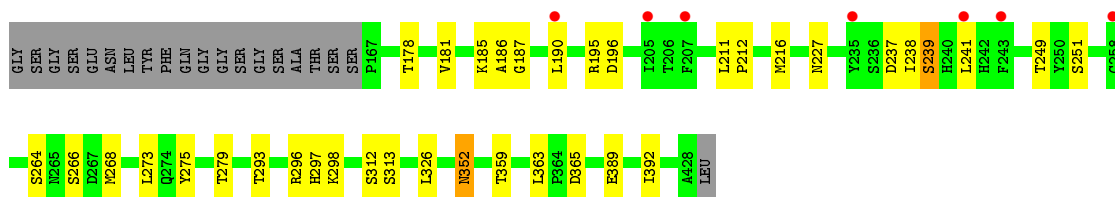
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	O 1	0	0

### 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 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.

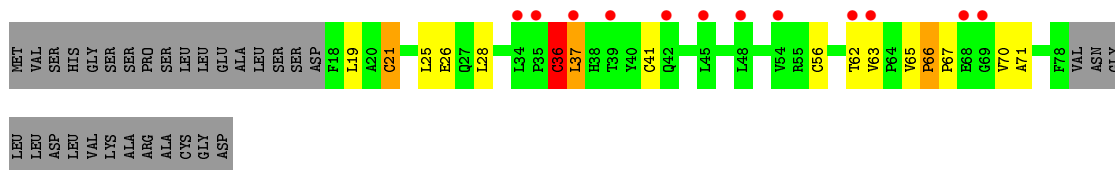
- Molecule 1: E3 ubiquitin-protein ligase SopA

Chain A: 



- Molecule 2: E3 ubiquitin-protein ligase TRIM56

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.02Å 71.02Å 122.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.48 – 2.85 43.48 – 2.85	Depositor EDS
% Data completeness (in resolution range)	98.0 (43.48-2.85) 98.4 (43.48-2.85)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 2.86Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.223 , 0.278 0.238 , 0.292	Depositor DCC
$R_{free}$ test set	418 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.1	Xtrriage
Anisotropy	0.469	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 64.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.066 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4610	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.79% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	A	0.23	0/2035	0.42	0/2776
2	B	0.30	0/410	0.69	0/567
All	All	0.25	0/2445	0.48	0/3343

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	36	CYS	Peptide

### 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	1989	1856	1855	19	0
2	B	403	359	359	13	0
3	B	2	0	0	0	0
4	A	1	0	0	0	0
All	All	2395	2215	2214	30	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 7.

All (30) 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)
2:B:36:CYS:SG	2:B:37:LEU:N	2.55	0.80
1:A:275:TYR:OH	1:A:298:LYS:O	2.10	0.69
1:A:268:MET:O	1:A:293:THR:OG1	2.13	0.61
2:B:70:VAL:HG12	2:B:71:ALA:H	1.69	0.57
2:B:66:PRO:HB2	2:B:67:PRO:CD	2.38	0.54
2:B:65:VAL:HG13	2:B:67:PRO:O	2.11	0.51
1:A:264:SER:O	1:A:266:SER:OG	2.24	0.50
1:A:297:HIS:NE2	2:B:26:GLU:OE2	2.38	0.49
1:A:196[B]:ASP:OD1	1:A:227:ASN:ND2	2.45	0.49
1:A:185:LYS:N	1:A:186:ALA:HA	2.28	0.48
1:A:190:LEU:O	1:A:195:ARG:NH1	2.45	0.48
1:A:211:LEU:HB3	1:A:212:PRO:HD2	1.95	0.48
1:A:389:GLU:O	1:A:392:ILE:N	2.47	0.48
1:A:297:HIS:O	1:A:359:THR:OG1	2.31	0.48
1:A:312:SER:OG	1:A:313:SER:N	2.47	0.47
2:B:21:CYS:HB2	2:B:41:CYS:H	1.78	0.47
2:B:36:CYS:O	2:B:37:LEU:HB3	2.15	0.47
2:B:65:VAL:CG2	2:B:66:PRO:HD2	2.46	0.45
1:A:326:LEU:HD21	1:A:363:LEU:HD11	1.97	0.45
1:A:352:ASN:O	1:A:352:ASN:ND2	2.46	0.45
2:B:28:LEU:H	2:B:28:LEU:HD12	1.83	0.43
1:A:178:THR:HA	1:A:181:VAL:HG12	2.01	0.43
2:B:62:THR:HG22	2:B:63:VAL:N	2.34	0.42
1:A:238:ILE:HG13	1:A:241:LEU:HD12	2.02	0.42
1:A:186:ALA:HB1	1:A:187:GLY:C	2.40	0.42
1:A:365:ASP:N	1:A:365:ASP:OD1	2.53	0.42
1:A:237:ASP:OD1	1:A:239:SER:OG	2.32	0.41
1:A:296:ARG:NH2	2:B:25:LEU:O	2.45	0.41
2:B:65:VAL:HG23	2:B:66:PRO:HD2	2.03	0.41
2:B:70:VAL:O	2:B:71:ALA:HB3	2.20	0.41



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	261/282 (93%)	241 (92%)	20 (8%)	0	100	100
2	B	59/93 (63%)	52 (88%)	4 (7%)	3 (5%)	2	3
All	All	320/375 (85%)	293 (92%)	24 (8%)	3 (1%)	17	34

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	37	LEU
2	B	66	PRO
2	B	36	CYS

### 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	217/249 (87%)	210 (97%)	7 (3%)	39	63
2	B	40/79 (51%)	37 (92%)	3 (8%)	13	28
All	All	257/328 (78%)	247 (96%)	10 (4%)	32	58

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	216	MET
1	A	239	SER
1	A	249	THR
1	A	251	SER
1	A	273	LEU
1	A	279	THR
1	A	352	ASN
2	B	19	LEU
2	B	21	CYS
2	B	56	CYS

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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.

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, 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	A	262/282 (92%)	0.19	7 (2%) 54 49	44, 77, 134, 156	0
2	B	61/93 (65%)	0.48	12 (19%) 1 0	67, 111, 147, 157	0
All	All	323/375 (86%)	0.24	19 (5%) 22 16	44, 83, 142, 157	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	45	LEU	3.5
2	B	37	LEU	3.4
2	B	35	PRO	3.1
1	A	190	LEU	3.1
1	A	205	ILE	3.0
2	B	54	VAL	2.8
1	A	207	PHE	2.8
2	B	34	LEU	2.7
1	A	241	LEU	2.5
2	B	68	GLU	2.3
2	B	62	THR	2.3
1	A	243	PHE	2.2
2	B	48	LEU	2.2
2	B	42	GLN	2.2
2	B	69	GLY	2.2
1	A	235	TYR	2.1
1	A	258	CYS	2.1
2	B	63	VAL	2.0
2	B	39	THR	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 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, 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	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	ZN	B	102	1/1	0.90	0.22	104,104,104,104	0
3	ZN	B	101	1/1	0.99	0.17	73,73,73,73	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.