



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2017 – 10:58 AM EDT

PDB ID : 4OWW
Title : Structural basis of SOSS1 in complex with a 35nt ssDNA
Authors : Ren, W.; Sun, Q.; Tang, X.; Song, H.
Deposited on : unknown
Resolution : 2.30 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

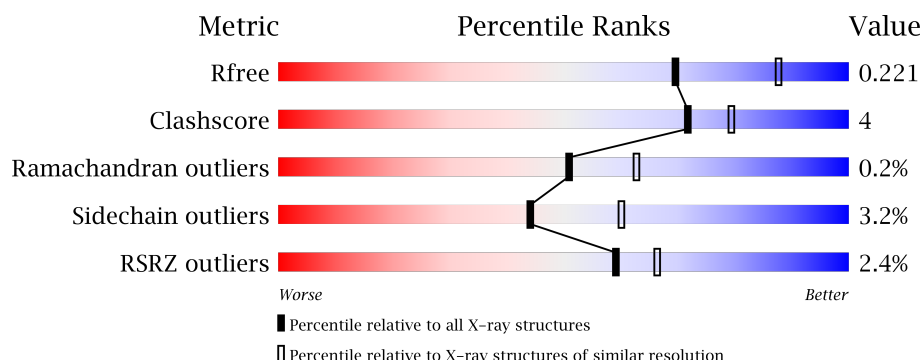
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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. 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	L	35	<div> <div>14%</div> <div>83%</div> </div>
2	A	500	<div> <div>2%</div> <div>82%</div> <div>10%</div> <div>7%</div> </div>
3	B	211	<div> <div>3%</div> <div>45%</div> <div>5%</div> <div>49%</div> </div>
4	C	104	<div> <div>35%</div> <div>63%</div> </div>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5184 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 DNA chain called DNA (5'-D(P*TP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	6	Total	C	N	O	P	0	0	0
			120	60	12	42	6			

- Molecule 2 is a protein called Integrator complex subunit 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	463	Total	C	N	O	S	0	0	0
			3724	2384	646	664	30			

- Molecule 3 is a protein called SOSS complex subunit B1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	107	Total	C	N	O	S	0	0	0
			827	529	139	155	4			

- Molecule 4 is a protein called SOSS complex subunit C.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	C	39	Total	C	N	O	0	0	0
			300	191	57	52			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	L	7	Total	O	0	0
			7	7		
5	A	151	Total	O	0	0
			151	151		
5	B	35	Total	O	0	0
			35	35		
5	C	20	Total	O	0	0
			20	20		

MET	ALA	ASN	SER	GLY	GLN	PHE	GLN	ASN	LYS	ASN	ARG	VAL	ALA	ILE	LEU	ALA	GLU	LEU	ASP	LYS	GLU	LYS	ARG	LYS	LYS	LEU	LEU	MET	GLN	ASN	GLN	SER	SER	THR	ASN	HIS	PRO	GLY	ALA	SER	ILE	ALA	LEU	SER	ARG	PRO	SER	LEU	ASN	LYS	ASP	PHE	ARG	ASP	HIS	ALA	GLU
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

GLN	Q62	Q73	L95	P99	R100	LEU	ASP	PRO	GLU
-----	-----	-----	-----	-----	------	-----	-----	-----	-----

GLOBAL-STATISTICS INFOmissingINFO

4 Model quality [i](#)

4.1 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	L	0.58	0/131	1.71	1/200 (0.5%)
2	A	0.27	0/3798	0.45	0/5142
3	B	0.29	0/841	0.49	0/1135
4	C	0.25	0/308	0.43	0/418
All	All	0.29	0/5078	0.54	1/6895 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	6	DT	O4'-C4'-C3'	-6.89	101.75	104.50

There are no chirality outliers.

There are no planarity outliers.

4.2 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	L	120	0	73	1	0
2	A	3724	0	3817	31	0
3	B	827	0	848	8	0
4	C	300	0	293	3	0
5	A	151	0	0	7	1
5	B	35	0	0	2	1
5	C	20	0	0	1	0
5	L	7	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5184	0	5031	40	1

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 4.

The worst 5 of 40 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:A:150:ARG:NH2	5:A:656:HOH:O	2.07	0.71
3:B:34:ASP:N	3:B:34:ASP:OD2	2.19	0.71
2:A:51:ARG:NH2	2:A:87:GLU:OE1	2.21	0.67
2:A:90:CYS:HB2	2:A:121:ILE:HD11	1.79	0.65
2:A:216:ILE:HG23	2:A:229:ARG:HD2	1.79	0.64

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:623:HOH:O	5:B:303:HOH:O[2_574]	2.09	0.11

4.3 Torsion angles [i](#)

4.3.1 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	
2	A	461/500 (92%)	455 (99%)	6 (1%)	0	100	100
3	B	105/211 (50%)	98 (93%)	6 (6%)	1 (1%)	18	20
4	C	37/104 (36%)	35 (95%)	2 (5%)	0	100	100
All	All	603/815 (74%)	588 (98%)	14 (2%)	1 (0%)	51	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	32	THR

4.3.2 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	
2	A	416/427 (97%)	404 (97%)	12 (3%)	48	64
3	B	93/179 (52%)	88 (95%)	5 (5%)	26	35
4	C	30/85 (35%)	30 (100%)	0	100	100
All	All	539/691 (78%)	522 (97%)	17 (3%)	44	60

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	229	ARG
2	A	253	ASP
3	B	32	THR
2	A	167	PHE
3	B	34	ASP

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

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 carbohydrates in this entry.

4.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	L	6/35 (17%)	0.74	0 100 100	60, 63, 73, 86	0
2	A	463/500 (92%)	0.13	9 (1%) 67 73	12, 32, 62, 84	0
3	B	107/211 (50%)	-0.02	6 (5%) 25 32	13, 35, 71, 101	0
4	C	39/104 (37%)	0.05	0 100 100	16, 32, 56, 61	0
All	All	615/850 (72%)	0.10	15 (2%) 59 66	12, 33, 64, 101	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	28	ARG	3.7
2	A	113	LEU	3.1
3	B	32	THR	2.9
2	A	497	CYS	2.9
3	B	31	LYS	2.9

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 carbohydrates in this entry.

5.4 Ligands [i](#)

There are no ligands in this entry.

5.5 Other polymers

There are no such residues in this entry.