



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 10:36 PM EDT

PDB ID : 5XGS
Title : Crystal structure of human WBSCR16
Authors : Koyama, M.; Sasaki, N.; Matsuura, Y.
Deposited on : 2017-04-16
Resolution : 2.00 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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	:	2.37.1
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)	:	2.37.1

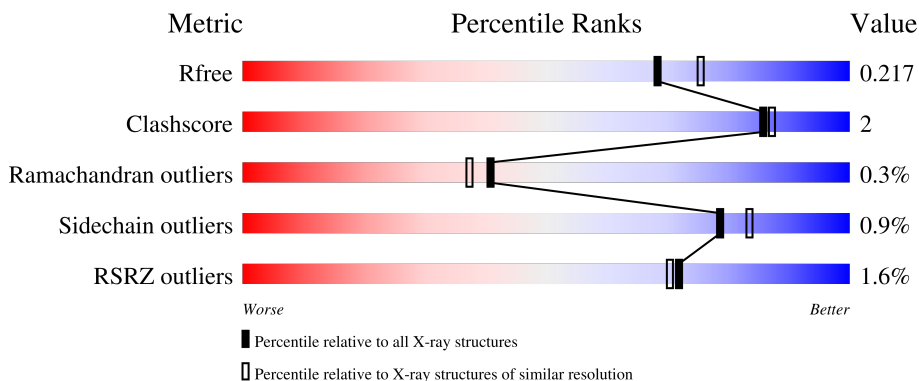
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

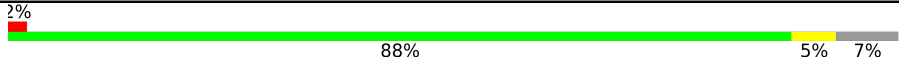
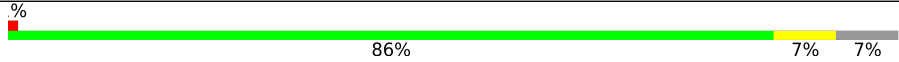
The reported resolution of this entry is 2.00 Å.

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 of 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	427	
1	B	427	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6614 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 RCC1-like G exchanging factor-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	0	0
			2999	1890	530	560	19			
1	B	398	Total	C	N	O	S	0	0	0
			3017	1904	533	561	19			

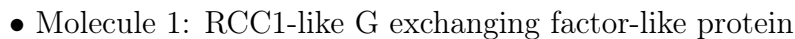
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ARG	see sequence details	UNP Q96I51
A	?	-	SER	see sequence details	UNP Q96I51
A	?	-	ARG	see sequence details	UNP Q96I51
A	?	-	LYS	see sequence details	UNP Q96I51
A	?	-	ASP	see sequence details	UNP Q96I51
A	?	-	LYS	see sequence details	UNP Q96I51
A	152	GLY	THR	see sequence details	UNP Q96I51
A	153	GLY	ARG	see sequence details	UNP Q96I51
B	?	-	ARG	see sequence details	UNP Q96I51
B	?	-	SER	see sequence details	UNP Q96I51
B	?	-	ARG	see sequence details	UNP Q96I51
B	?	-	LYS	see sequence details	UNP Q96I51
B	?	-	ASP	see sequence details	UNP Q96I51
B	?	-	LYS	see sequence details	UNP Q96I51
B	152	GLY	THR	see sequence details	UNP Q96I51
B	153	GLY	ARG	see sequence details	UNP Q96I51

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	301	Total	O	0	0
			301	301		
2	B	297	Total	O	0	0
			297	297		

- Molecule 1: RCC1-like G exchanging factor-like protein



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	111.02Å 54.55Å 140.47Å 90.00° 95.82° 90.00°	Depositor
Resolution (Å)	34.94 – 2.00 34.94 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.1 (34.94-2.00) 99.1 (34.94-2.00)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.11.1 _2575	Depositor
R, R_{free}	0.180 , 0.218 0.179 , 0.217	Depositor DCC
R_{free} test set	2822 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	18.4	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6614	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.39% 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	0.26	0/3072	0.47	0/4170
1	B	0.26	0/3088	0.52	2/4188 (0.0%)
All	All	0.26	0/6160	0.49	2/8358 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	91	ARG	NE-CZ-NH1	-12.37	114.12	120.30
1	B	91	ARG	NE-CZ-NH2	8.96	124.78	120.30

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	A	2999	0	2915	10	0
1	B	3017	0	2937	17	0
2	A	301	0	0	2	0
2	B	297	0	0	3	1
All	All	6614	0	5852	26	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 2.

The worst 5 of 26 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:90:ALA:O	2:B:501:HOH:O	2.01	0.76
1:B:57:ARG:NH2	1:B:445:GLU:OE1	2.23	0.72
1:A:212:GLU:OE1	2:A:501:HOH:O	2.11	0.68
1:A:187:LEU:HD11	1:A:242:SER:HB3	1.82	0.61
1:A:383:THR:HG21	1:B:443:PRO:HB3	1.85	0.58

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)
2:B:511:HOH:O	2:B:563:HOH:O[4_655]	2.08	0.12

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	396/427 (93%)	392 (99%)	3 (1%)	1 (0%)	41	37
1	B	392/427 (92%)	386 (98%)	5 (1%)	1 (0%)	41	37
All	All	788/854 (92%)	778 (99%)	8 (1%)	2 (0%)	41	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	291	TYR
1	B	291	TYR

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	320/342 (94%)	316 (99%)	4 (1%)	69	74
1	B	323/342 (94%)	321 (99%)	2 (1%)	86	90
All	All	643/684 (94%)	637 (99%)	6 (1%)	78	83

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	412	LYS
1	B	68	PHE
1	B	134	MET
1	A	68	PHE
1	A	60	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	231	GLN
1	B	411	ASN

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 [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	A	398/427 (93%)	-0.26	8 (2%) 65 63	9, 16, 34, 47	0
1	B	398/427 (93%)	-0.24	5 (1%) 77 76	11, 18, 35, 50	0
All	All	796/854 (93%)	-0.25	13 (1%) 72 70	9, 17, 34, 50	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	153	GLY	4.9
1	A	155	TYR	4.6
1	A	154	GLY	3.8
1	B	390	PHE	3.7
1	A	212	GLU	3.4

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.