



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:27 PM UTC

PDB ID : 5Z6P / pdb_00005z6p
Title : The crystal structure of an agarase, AgWH50C
Authors : Mao, X.; Zhou, J.; Zhang, P.; Zhang, L.; Zhang, J.; Li, Y.
Deposited on : 2018-01-24
Resolution : 2.06 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

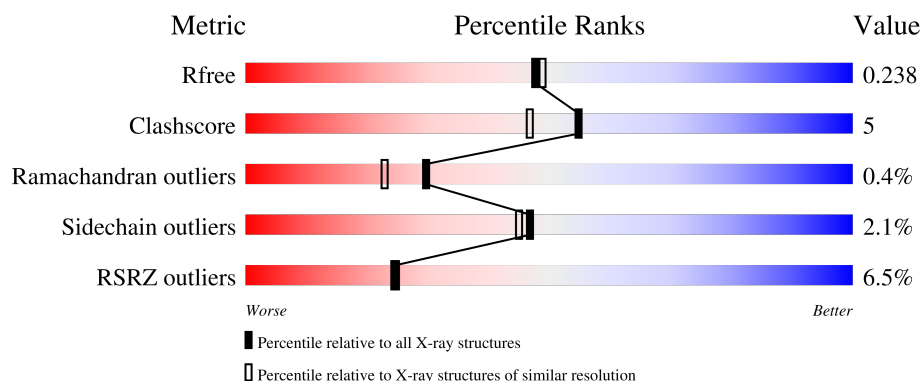
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.06 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}	180053	3774 (2.08-2.04)
Clashscore	190562	3883 (2.08-2.04)
Ramachandran outliers	187476	3860 (2.08-2.04)
Sidechain outliers	187428	3860 (2.08-2.04)
RSRZ outliers	180081	3775 (2.08-2.04)

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	765	 3% 84% 7% • 8%
1	B	765	 9% 72% 12% • 16%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11777 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 B-agarase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	702	Total	C	N	O	S	0	0	0
			5621	3598	952	1045	26			
1	B	646	Total	C	N	O	S	0	0	0
			5181	3312	881	962	26			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	HIS	-	expression tag	UNP A0A023LLI1
A	-23	HIS	-	expression tag	UNP A0A023LLI1
A	-22	HIS	-	expression tag	UNP A0A023LLI1
A	-21	HIS	-	expression tag	UNP A0A023LLI1
A	-20	HIS	-	expression tag	UNP A0A023LLI1
A	-19	HIS	-	expression tag	UNP A0A023LLI1
A	-18	HIS	-	expression tag	UNP A0A023LLI1
A	-17	HIS	-	expression tag	UNP A0A023LLI1
A	-16	GLY	-	expression tag	UNP A0A023LLI1
A	-15	SER	-	expression tag	UNP A0A023LLI1
A	-14	ASP	-	expression tag	UNP A0A023LLI1
A	-13	TYR	-	expression tag	UNP A0A023LLI1
A	-12	ASP	-	expression tag	UNP A0A023LLI1
A	-11	ILE	-	expression tag	UNP A0A023LLI1
A	-10	PRO	-	expression tag	UNP A0A023LLI1
A	-9	THR	-	expression tag	UNP A0A023LLI1
A	-8	THR	-	expression tag	UNP A0A023LLI1
A	-7	GLU	-	expression tag	UNP A0A023LLI1
A	-6	ASN	-	expression tag	UNP A0A023LLI1
A	-5	LEU	-	expression tag	UNP A0A023LLI1
A	-4	TYR	-	expression tag	UNP A0A023LLI1
A	-3	PHE	-	expression tag	UNP A0A023LLI1
A	-2	GLN	-	expression tag	UNP A0A023LLI1
A	-1	GLY	-	expression tag	UNP A0A023LLI1
A	0	SER	-	expression tag	UNP A0A023LLI1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	ASP	GLY	see sequence details	UNP A0A023LLI1
B	-24	HIS	-	expression tag	UNP A0A023LLI1
B	-23	HIS	-	expression tag	UNP A0A023LLI1
B	-22	HIS	-	expression tag	UNP A0A023LLI1
B	-21	HIS	-	expression tag	UNP A0A023LLI1
B	-20	HIS	-	expression tag	UNP A0A023LLI1
B	-19	HIS	-	expression tag	UNP A0A023LLI1
B	-18	HIS	-	expression tag	UNP A0A023LLI1
B	-17	HIS	-	expression tag	UNP A0A023LLI1
B	-16	GLY	-	expression tag	UNP A0A023LLI1
B	-15	SER	-	expression tag	UNP A0A023LLI1
B	-14	ASP	-	expression tag	UNP A0A023LLI1
B	-13	TYR	-	expression tag	UNP A0A023LLI1
B	-12	ASP	-	expression tag	UNP A0A023LLI1
B	-11	ILE	-	expression tag	UNP A0A023LLI1
B	-10	PRO	-	expression tag	UNP A0A023LLI1
B	-9	THR	-	expression tag	UNP A0A023LLI1
B	-8	THR	-	expression tag	UNP A0A023LLI1
B	-7	GLU	-	expression tag	UNP A0A023LLI1
B	-6	ASN	-	expression tag	UNP A0A023LLI1
B	-5	LEU	-	expression tag	UNP A0A023LLI1
B	-4	TYR	-	expression tag	UNP A0A023LLI1
B	-3	PHE	-	expression tag	UNP A0A023LLI1
B	-2	GLN	-	expression tag	UNP A0A023LLI1
B	-1	GLY	-	expression tag	UNP A0A023LLI1
B	0	SER	-	expression tag	UNP A0A023LLI1
B	307	ASP	GLY	see sequence details	UNP A0A023LLI1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	592	Total O 592 592	0	0
2	B	383	Total O 383 383	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	215.80Å 93.65Å 85.05Å 90.00° 105.66° 90.00°	Depositor
Resolution (Å)	46.82 – 2.06 46.82 – 2.06	Depositor EDS
% Data completeness (in resolution range)	92.2 (46.82-2.06) 92.2 (46.82-2.06)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.15 (at 2.07Å)	Xtriage
Refinement program	PHENIX 1.8.3_1479	Depositor
R, R_{free}	0.195 , 0.245 (Not available) , 0.238	Depositor DCC
R_{free} test set	4573 reflections (4.54%)	wwPDB-VP
Wilson B-factor (Å ²)	26.6	Xtriage
Anisotropy	0.153	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 38.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11777	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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	0.63	0/5782	0.80	4/7842 (0.1%)
1	B	0.55	0/5326	0.82	5/7214 (0.1%)
All	All	0.59	0/11108	0.81	9/15056 (0.1%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	511	CYS	CA-C-N	7.03	127.33	119.32
1	B	511	CYS	C-N-CA	7.03	127.33	119.32
1	B	627	GLU	CA-C-N	-5.96	113.84	120.45
1	B	627	GLU	C-N-CA	-5.96	113.84	120.45
1	A	738	PHE	CA-C-N	5.29	131.22	121.70

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	5621	0	5340	41	0
1	B	5181	0	4901	60	0
2	A	592	0	0	11	5
2	B	383	0	0	6	2
All	All	11777	0	10241	101	5

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 5.

The worst 5 of 101 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:624:LEU:H	1:B:677:THR:HG21	1.29	0.96
1:A:567:GLU:OE1	1:A:606:ARG:NH1	2.04	0.90
1:B:552:LEU:HD23	1:B:554:HIS:H	1.37	0.89
1:A:355:TYR:OH	1:A:366:GLN:OE1	1.92	0.85
1:A:192:ASN:ND2	2:A:801:HOH:O	2.07	0.85

All (5) 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:A:963:HOH:O	2:B:933:HOH:O[2_1169]	2.14	0.06
2:A:1222:HOH:O	2:A:1304:HOH:O[2_11510]	2.16	0.04
2:A:925:HOH:O	2:A:1258:HOH:O[2_11510]	2.19	0.01
2:A:1245:HOH:O	2:A:1355:HOH:O[2_11510]	2.19	0.01
2:A:1320:HOH:O	2:B:1109:HOH:O[1_566]	2.19	0.01

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	696/765 (91%)	677 (97%)	19 (3%)	0	100	100
1	B	632/765 (83%)	600 (95%)	27 (4%)	5 (1%)	16	8
All	All	1328/1530 (87%)	1277 (96%)	46 (4%)	5 (0%)	30	23

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	125	ALA
1	B	18	SER
1	B	508	ALA
1	B	551	GLY
1	B	507	SER

5.3.2 Protein sidechains ⓘ

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	587/644 (91%)	572 (97%)	15 (3%)	40	37
1	B	542/644 (84%)	533 (98%)	9 (2%)	53	53
All	All	1129/1288 (88%)	1105 (98%)	24 (2%)	47	45

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	677	THR
1	B	194	LEU
1	B	108	LYS
1	B	333	GLU
1	A	330	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	61	ASN
1	B	94	GLN
1	B	558	GLN
1	B	366	GLN
1	A	192	ASN

5.3.3 RNA ⓘ

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 oligosaccharides 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	702/765 (91%)	-0.16	21 (2%)	52 53	13, 20, 41, 70	0
1	B	646/765 (84%)	0.23	66 (10%)	12 12	12, 24, 54, 75	0
All	All	1348/1530 (88%)	0.03	87 (6%)	25 24	12, 21, 49, 75	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	554	HIS	6.0
1	A	444	TRP	5.0
1	B	552	LEU	4.9
1	B	520	LEU	4.6
1	B	558	GLN	4.4

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 oligosaccharides 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.