



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2025 – 09:58 PM EST

PDB ID : 9EH6
Title : Crystal Structure of AroC
Authors : Hu, Z.; Klupt, K.; Zechel, D.; Jia, Z.; Howe, G.
Deposited on : 2024-11-22
Resolution : 2.25 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 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)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.3

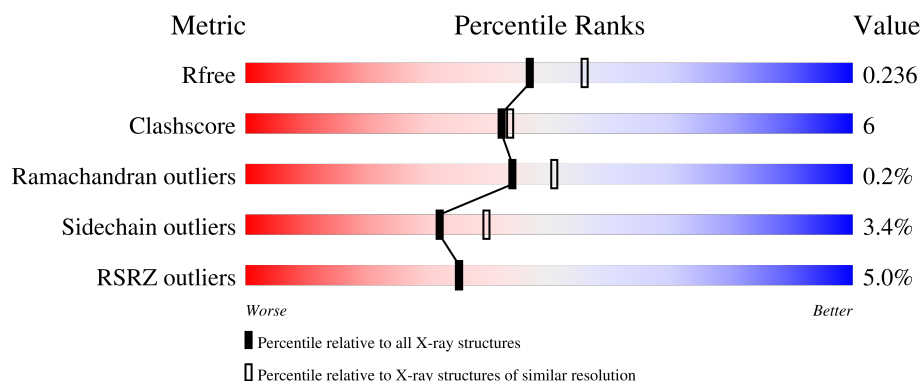
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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}	164625	1763 (2.26-2.26)
Clashscore	180529	1919 (2.26-2.26)
Ramachandran outliers	177936	1884 (2.26-2.26)
Sidechain outliers	177891	1885 (2.26-2.26)
RSRZ outliers	164620	1763 (2.26-2.26)

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	230	
1	B	230	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3405 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 Alpha/beta hydrolase fold-5 domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	0	0
			1643	1046	284	308	5			
1	B	218	Total	C	N	O	S	0	0	0
			1643	1046	284	308	5			

There are 20 discrepancies between the modelled and reference sequences:

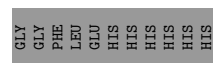
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	expression tag	UNP I0I495
A	2	MET	-	expression tag	UNP I0I495
A	223	LEU	-	expression tag	UNP I0I495
A	224	GLU	-	expression tag	UNP I0I495
A	225	HIS	-	expression tag	UNP I0I495
A	226	HIS	-	expression tag	UNP I0I495
A	227	HIS	-	expression tag	UNP I0I495
A	228	HIS	-	expression tag	UNP I0I495
A	229	HIS	-	expression tag	UNP I0I495
A	230	HIS	-	expression tag	UNP I0I495
B	1	HIS	-	expression tag	UNP I0I495
B	2	MET	-	expression tag	UNP I0I495
B	223	LEU	-	expression tag	UNP I0I495
B	224	GLU	-	expression tag	UNP I0I495
B	225	HIS	-	expression tag	UNP I0I495
B	226	HIS	-	expression tag	UNP I0I495
B	227	HIS	-	expression tag	UNP I0I495
B	228	HIS	-	expression tag	UNP I0I495
B	229	HIS	-	expression tag	UNP I0I495
B	230	HIS	-	expression tag	UNP I0I495

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	52	Total 52	O 52	0	0
2	B	67	Total 67	O 67	0	0

i

- Molecule 1: Alpha/beta hydrolase fold-5 domain-containing protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.19Å 46.82Å 76.47Å 90.00° 101.97° 90.00°	Depositor
Resolution (Å)	40.50 – 2.25 40.50 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.7 (40.50-2.25) 99.6 (40.50-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.67 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, R_{free}	0.186 , 0.240 0.184 , 0.236	Depositor DCC
R_{free} test set	904 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	19.5	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 27.0	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	0.94	EDS
Total number of atoms	3405	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 8.32% 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

5.1 Standard geometry

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	0.40	0/1686	0.79	1/2305 (0.0%)
1	B	0.39	0/1685	0.73	1/2302 (0.0%)
All	All	0.40	0/3371	0.76	2/4607 (0.0%)

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
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	187	TYR	CB-CA-C	5.61	121.62	110.40
1	B	70	THR	CB-CA-C	5.16	125.52	111.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	188	GLY	Peptide
1	B	188	GLY	Peptide
1	B	36	ARG	Sidechain

5.2 Too-close contacts ⓘ

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	1643	0	1582	11	0
1	B	1643	0	1581	30	0
2	A	52	0	0	0	0
2	B	67	0	0	5	0
All	All	3405	0	3163	40	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 6.

All (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)
1:B:80:ASP:HA	1:B:82:ASP:HB2	1.37	1.07
1:B:80:ASP:CA	1:B:82:ASP:HB2	2.08	0.83
1:B:79:PHE:O	1:B:82:ASP:N	2.26	0.69
1:B:33:ALA:O	1:B:35:PRO:HD3	1.97	0.64
1:A:186:TRP:HE1	1:A:204:GLN:HE21	1.46	0.64
1:B:41:ILE:HG23	1:B:70:THR:HG21	1.79	0.64
1:A:47:ARG:HE	1:B:79:PHE:HE1	1.43	0.62
1:B:16:ASP:OD1	1:B:58:ARG:HD2	2.00	0.61
1:A:104:LEU:O	1:A:108:MET:HG3	2.01	0.60
1:B:80:ASP:N	2:B:301:HOH:O	2.31	0.59
1:A:186:TRP:HE1	1:A:204:GLN:NE2	2.01	0.58
1:B:79:PHE:O	1:B:82:ASP:CB	2.51	0.58
1:B:186:TRP:HE1	1:B:204:GLN:HE21	1.54	0.55
1:B:114:HIS:HD2	1:B:115:ASN:OD1	1.89	0.55
1:B:80:ASP:HB3	2:B:301:HOH:O	2.07	0.54
1:B:145:TYR:CE1	1:B:172:ARG:HG3	2.43	0.54
1:A:99:ILE:O	1:A:123:VAL:HA	2.08	0.53
1:B:41:ILE:HG23	1:B:70:THR:CG2	2.40	0.50
1:B:73:PRO:HB3	1:B:78:VAL:HB	1.93	0.50
1:B:73:PRO:HD3	1:B:83:ARG:HD2	1.95	0.49
1:B:80:ASP:CB	1:B:82:ASP:HB2	2.42	0.48
1:B:205:GLN:HG3	2:B:357:HOH:O	2.13	0.48
1:B:186:TRP:HE1	1:B:204:GLN:NE2	2.13	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:30:GLN:NE2	1:B:31:PRO:HD2	2.30	0.46
1:A:70:THR:HG21	1:A:87:VAL:HG21	1.97	0.46
1:A:83:ARG:HA	1:A:83:ARG:HD3	1.68	0.45
1:A:56:GLN:HG3	1:A:204:GLN:HG3	1.99	0.45
1:B:68:VAL:O	1:B:70:THR:HG22	2.18	0.44
1:B:79:PHE:O	1:B:82:ASP:CG	2.56	0.44
1:B:46:GLY:O	1:B:47:ARG:HB2	2.17	0.43
1:B:80:ASP:HA	1:B:82:ASP:CB	2.27	0.43
1:B:79:PHE:O	1:B:82:ASP:HB2	2.17	0.43
1:B:205:GLN:CG	2:B:357:HOH:O	2.67	0.43
1:A:61:ALA:HA	1:A:65:PHE:O	2.19	0.42
1:A:59:ALA:HB3	1:A:208:VAL:HG21	2.02	0.42
1:B:82:ASP:C	1:B:84:ALA:H	2.23	0.42
1:A:127:ALA:HB2	1:A:184:PHE:CE2	2.56	0.41
1:B:33:ALA:HB3	2:B:304:HOH:O	2.20	0.41
1:B:82:ASP:C	1:B:84:ALA:N	2.74	0.41
1:B:111:ASN:HD22	1:B:134:ASP:CG	2.24	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	216/230 (94%)	203 (94%)	12 (6%)	1 (0%)	25	25
1	B	214/230 (93%)	203 (95%)	11 (5%)	0	100	100
All	All	430/460 (94%)	406 (94%)	23 (5%)	1 (0%)	44	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	33	ALA

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	162/173 (94%)	156 (96%)	6 (4%)	29	35
1	B	162/173 (94%)	157 (97%)	5 (3%)	35	43
All	All	324/346 (94%)	313 (97%)	11 (3%)	32	40

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	ASP
1	A	72	MET
1	A	76	LEU
1	A	78	VAL
1	A	79	PHE
1	A	83	ARG
1	B	17	ASP
1	B	70	THR
1	B	72	MET
1	B	83	ARG
1	B	172	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	114	HIS
1	A	204	GLN
1	B	30	GLN
1	B	75	ASN
1	B	114	HIS
1	B	204	GLN

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	80:ASP	C	81:ALA	N	6.51

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	218/230 (94%)	-0.24	9 (4%) 42 42	7, 17, 49, 97	0
1	B	218/230 (94%)	-0.18	13 (5%) 29 28	7, 15, 48, 129	0
All	All	436/460 (94%)	-0.21	22 (5%) 35 35	7, 16, 49, 129	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	79	PHE	7.1
1	B	80	ASP	5.4
1	B	78	VAL	5.3
1	A	79	PHE	5.1
1	B	82	ASP	4.6
1	B	15	GLY	4.3
1	A	76	LEU	3.6
1	A	78	VAL	3.5
1	A	75	ASN	3.2
1	B	81	ALA	3.1
1	B	16	ASP	3.0
1	A	74	LEU	2.9
1	B	77	ALA	2.8
1	B	188	GLY	2.8
1	B	74	LEU	2.8
1	B	108	MET	2.7
1	A	188	GLY	2.6
1	A	77	ALA	2.5
1	B	73	PRO	2.5
1	B	33	ALA	2.4
1	A	14	GLN	2.4
1	A	33	ALA	2.4

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.