



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:03 PM GMT

PDB ID : 4GYR
Title : Granulibacter bethesdensis allophanate hydrolase apo
Authors : Lin, Y.; St Maurice, M.
Deposited on : 2012-09-05
Resolution : 2.80 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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

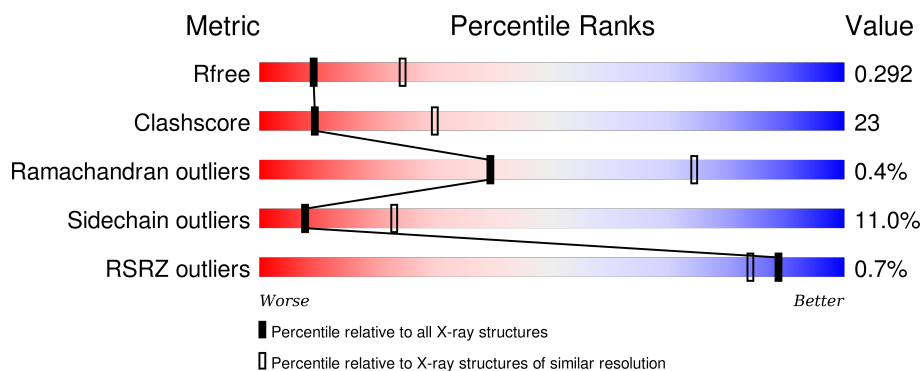
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

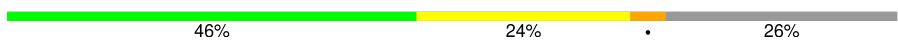

The reported resolution of this entry is 2.80 Å.

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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	A	621	 46% 24% • 26%
1	B	621	 46% 24% • 26%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6930 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 Allophanate hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	461	Total	C	N	O	S	11	1	0
			3404	2159	583	651	11			
1	B	461	Total	C	N	O	S	9	1	0
			3410	2165	586	648	11			

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-28	MET	-	EXPRESSION TAG	UNP Q0BRB0
A	-27	GLY	-	EXPRESSION TAG	UNP Q0BRB0
A	-26	SER	-	EXPRESSION TAG	UNP Q0BRB0
A	-25	SER	-	EXPRESSION TAG	UNP Q0BRB0
A	-24	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-23	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-22	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-21	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-20	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-19	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-18	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-17	HIS	-	EXPRESSION TAG	UNP Q0BRB0
A	-16	ASP	-	EXPRESSION TAG	UNP Q0BRB0
A	-15	TYR	-	EXPRESSION TAG	UNP Q0BRB0
A	-14	ASP	-	EXPRESSION TAG	UNP Q0BRB0
A	-13	ILE	-	EXPRESSION TAG	UNP Q0BRB0
A	-12	PRO	-	EXPRESSION TAG	UNP Q0BRB0
A	-11	THR	-	EXPRESSION TAG	UNP Q0BRB0
A	-10	SER	-	EXPRESSION TAG	UNP Q0BRB0
A	-9	GLU	-	EXPRESSION TAG	UNP Q0BRB0
A	-8	ASN	-	EXPRESSION TAG	UNP Q0BRB0
A	-7	LEU	-	EXPRESSION TAG	UNP Q0BRB0
A	-6	TYR	-	EXPRESSION TAG	UNP Q0BRB0
A	-5	PHE	-	EXPRESSION TAG	UNP Q0BRB0
A	-4	GLN	-	EXPRESSION TAG	UNP Q0BRB0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP Q0BRB0
A	-2	LEU	-	EXPRESSION TAG	UNP Q0BRB0
A	-1	LEU	-	EXPRESSION TAG	UNP Q0BRB0
A	0	GLN	-	EXPRESSION TAG	UNP Q0BRB0
B	-28	MET	-	EXPRESSION TAG	UNP Q0BRB0
B	-27	GLY	-	EXPRESSION TAG	UNP Q0BRB0
B	-26	SER	-	EXPRESSION TAG	UNP Q0BRB0
B	-25	SER	-	EXPRESSION TAG	UNP Q0BRB0
B	-24	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-23	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-22	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-21	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-20	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-19	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-18	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-17	HIS	-	EXPRESSION TAG	UNP Q0BRB0
B	-16	ASP	-	EXPRESSION TAG	UNP Q0BRB0
B	-15	TYR	-	EXPRESSION TAG	UNP Q0BRB0
B	-14	ASP	-	EXPRESSION TAG	UNP Q0BRB0
B	-13	ILE	-	EXPRESSION TAG	UNP Q0BRB0
B	-12	PRO	-	EXPRESSION TAG	UNP Q0BRB0
B	-11	THR	-	EXPRESSION TAG	UNP Q0BRB0
B	-10	SER	-	EXPRESSION TAG	UNP Q0BRB0
B	-9	GLU	-	EXPRESSION TAG	UNP Q0BRB0
B	-8	ASN	-	EXPRESSION TAG	UNP Q0BRB0
B	-7	LEU	-	EXPRESSION TAG	UNP Q0BRB0
B	-6	TYR	-	EXPRESSION TAG	UNP Q0BRB0
B	-5	PHE	-	EXPRESSION TAG	UNP Q0BRB0
B	-4	GLN	-	EXPRESSION TAG	UNP Q0BRB0
B	-3	GLY	-	EXPRESSION TAG	UNP Q0BRB0
B	-2	LEU	-	EXPRESSION TAG	UNP Q0BRB0
B	-1	LEU	-	EXPRESSION TAG	UNP Q0BRB0
B	0	GLN	-	EXPRESSION TAG	UNP Q0BRB0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	65	Total O 65 65	0	0
2	B	51	Total O 51 51	0	0

PRO	VAL	F163	Q345	S453	PRO
GLY	GLY				GLY
LEU	LEU	T167	L351	D456	LEU
VAL	VAL	D168	P457	P457	VAL
ALA	ALA	T169	Q354	E460	ALA
ASP	ASP	A170	R361	E461	ASP
PRO	PRO	G171	K362	T462	PRO
ASP	ASP	S172		ASN	ASP
PHE	PHE	G173	L365	PRO	PHE
GLY	GLY	R174	A264	GLY	GLY
GLY	GLY	V175	V379	GLN	GLY
TRP	TRP	P176		ILE	TRP
ARG	ARG	A177	D382	ALA	ARG
ALA	ALA	A178	P383	ALA	ALA
ILE	ILE	F179	V384	LEU	ILE
GLU	GLU	N180	R385	ALA	GLU
ILE	ILE	N181	L386	VAL	ILE
LEU	LEU	L182	D274	VAL	LEU
LEU	LEU		N387	GLY	LEU
TRP	TRP	K186		GLY	TRP
SER	SER	P187	L390	ALA	SER
MET	MET		G391	HIS	MET
ASP	ASP	L191	H392	LEU	ASP
ALA	ALA	L192	D280	GLY	ALA
GLU	GLU	S193	Y393	GLY	GLU
ALA	ALA	T194	T394	GLN	ALA
PHE	PHE	S195	N395	PRO	PHE
GLY	GLY		A404	LEU	GLY
THR	THR	C201	I405	HIS	THR
THR	THR	R202	P408	TRP	THR
THR	THR	S203	A409	GLN	THR
ALA	ALA	L204	G410	LEU	ALA
LEU	LEU	D205	F411	THR	LEU
LEU	LEU	T208	I412	ALA	LEU
ALA	ALA	V214		THR	ALA
ALA	ALA		L416	LEU	ALA
ILE	ILE	T218	P417	LEU	ILE
GLY	GLY	L219		VAL	GLY
THR	THR	I220	V420	ALA	THR
LEU	LEU	R221		ARG	LEU
ARG	ARG	R222	V423	THR	ARG
LEU	LEU	E225	S428	ARG	LEU
ALA	ALA		D429	THR	ALA
ASP	ASP	D231	D430	ALA	ASP
GLY	GLY		S431	PRO	GLY
THR	THR	S234	M432	TYR	THR
VAL	VAL	S237	A433	ARG	SER
LYS	LYS	Q238	L434	LEU	VAL
GLY	GLY	R239		TYR	GLY
PHE	PHE	R240	R438	ALA	PHE
VAL	VAL			LEU	VAL
CYS	CYS	V245	L443	ALA	CYS
GLU	GLU	H244	E444	GLU	GLU
PRO	PRO	V245	P445	THR	PRO
ALA	ALA	V251	G446	ILE	ALA
LEU	LEU	P252	R451	PRO	LEU
			A452	LYS	

4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α , β , γ	78.09 Å 78.09 Å 397.54 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.05 – 2.80 40.05 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.05-2.80) 99.9 (40.05-2.80)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	14.99 (at 2.81 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.256 , 0.307 0.243 , 0.292	Depositor DCC
R_{free} test set	2025 reflections (6.05%)	DCC
Wilson B-factor (Å ²)	40.3	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
Estimated twinning fraction	0.477 for h,-h-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 33510 reflections	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	6930	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.32	1/3478 (0.0%)	0.52	3/4752 (0.1%)
1	B	0.31	1/3484 (0.0%)	0.52	2/4757 (0.0%)
All	All	0.32	2/6962 (0.0%)	0.52	5/9509 (0.1%)

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	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	460	GLU	CD-OE1	-8.74	1.16	1.25
1	A	460	GLU	CD-OE1	-8.44	1.16	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	222	ARG	NE-CZ-NH2	-6.63	116.99	120.30
1	A	222	ARG	NE-CZ-NH2	-6.50	117.05	120.30
1	A	460	GLU	CG-CD-OE1	-5.73	106.84	118.30
1	B	460	GLU	CG-CD-OE1	-5.68	106.94	118.30
1	A	460	GLU	CG-CD-OE2	5.02	128.33	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	262	ASN	Peptide
1	B	262	ASN	Peptide

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	3404	0	3333	157	0
1	B	3410	0	3347	149	1
2	A	65	0	0	23	1
2	B	51	0	0	10	0
All	All	6930	0	6680	305	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 23.

The worst 5 of 305 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:411:PHE:CE1	1:B:417:PRO:HB3	1.64	1.32
1:B:411:PHE:CE2	1:B:451:ARG:HD2	1.73	1.24
1:A:411:PHE:CE2	1:A:451:ARG:HD2	1.74	1.21
1:B:411:PHE:HE2	1:B:451:ARG:CD	1.54	1.20
1:A:218:THR:CG2	1:A:222:ARG:HH12	1.57	1.17

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)
1:B:279:LEU:O	2:A:661:HOH:O[1_565]	2.03	0.17

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	460/621 (74%)	429 (93%)	29 (6%)	2 (0%)	39	74
1	B	460/621 (74%)	428 (93%)	30 (6%)	2 (0%)	39	74
All	All	920/1242 (74%)	857 (93%)	59 (6%)	4 (0%)	39	74

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	263	THR
1	B	263	THR
1	A	140	ASP
1	B	140	ASP

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	340/482 (70%)	303 (89%)	37 (11%)	8	23
1	B	340/482 (70%)	302 (89%)	38 (11%)	7	22
All	All	680/964 (70%)	605 (89%)	75 (11%)	8	23

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	429	ASP
1	B	96	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	405	ILE
1	A	431	SER
1	B	50	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	244	HIS
1	A	354	GLN
1	B	244	HIS
1	B	354	GLN
1	B	440	HIS

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](#)

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	461/621 (74%)	-0.31	3 (0%) 89 84	25, 40, 59, 73	9 (1%)
1	B	461/621 (74%)	-0.32	3 (0%) 89 84	25, 41, 59, 74	8 (1%)
All	All	922/1242 (74%)	-0.32	6 (0%) 89 84	25, 40, 59, 74	17 (1%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	274	ASP	2.9
1	B	244	HIS	2.9
1	B	281	ALA	2.3
1	A	274	ASP	2.1
1	A	414	THR	2.1

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.