



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 05:33 pm BST

PDB ID : 6LR1  
Title : Hexachlorobenzene Monooxygenase (HcbA1) from *Nocardioides* sp. strain PD653  
Authors : Guo, Y.; Zheng, J.T.; Zhou, N.Y.  
Deposited on : 2020-01-15  
Resolution : 2.50 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](mailto: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 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	:	2.11
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)	:	2.11

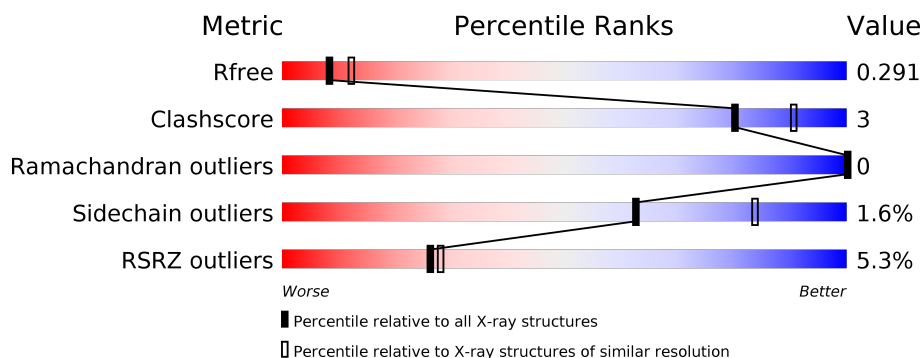
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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  $\geq 3$ , 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	451	<div> <div>5%</div> <div>95%</div> <div>5%</div> </div>
1	B	451	<div> <div>5%</div> <div>88%</div> <div>12%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7138 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 Hexachlorobenzene oxidative dehalogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3529	2219	621	671	18			
1	B	451	Total	C	N	O	S	0	0	0
			3529	2219	621	671	18			

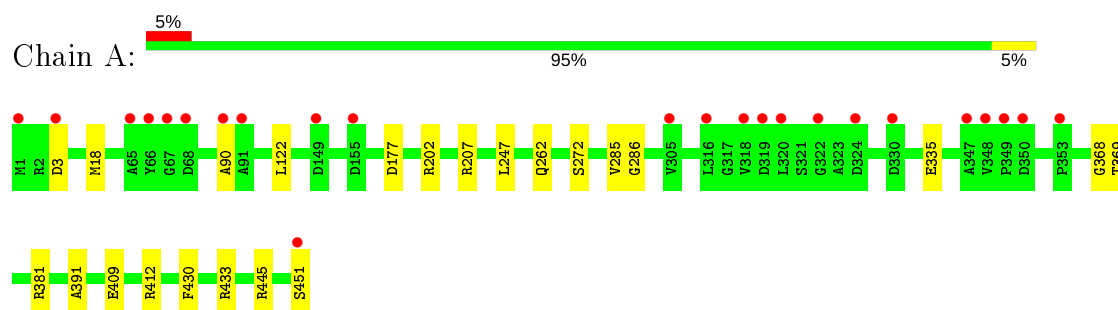
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	37	Total	O	0	0
			37	37		
2	B	43	Total	O	0	0
			43	43		

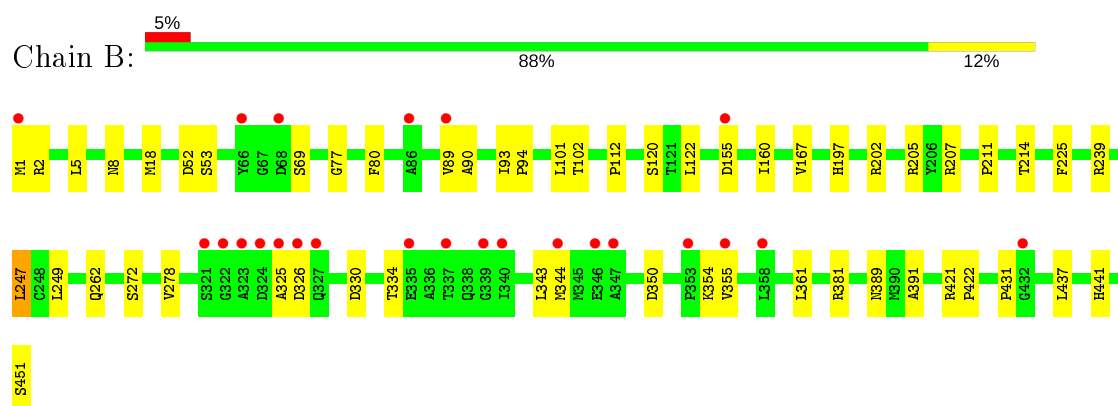
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Hexachlorobenzene oxidative dehalogenase



- Molecule 1: Hexachlorobenzene oxidative dehalogenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.09Å 139.31Å 169.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.33 – 2.50 47.34 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (87.33-2.50) 98.0 (47.34-2.50)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.85 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.228 , 0.294 0.234 , 0.291	Depositor DCC
$R_{free}$ test set	2236 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.1	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 31.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.56$ , $\langle L^2 \rangle = 0.41$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7138	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 56.11 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.8697e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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	0.86	0/3612	0.89	0/4898
1	B	0.94	0/3612	0.95	0/4898
All	All	0.90	0/7224	0.92	0/9796

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	3529	0	3407	11	0
1	B	3529	0	3407	27	0
2	A	37	0	0	0	0
2	B	43	0	0	0	0
All	All	7138	0	6814	38	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 3.

All (38) 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:249:LEU:CD1	1:B:344:MET:HA	2.07	0.85

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:247:LEU:HD22	1:B:361:LEU:O	1.90	0.72
1:B:262:GLN:OE1	1:B:272:SER:HB2	1.92	0.69
1:A:430:PHE:HB3	1:A:433:ARG:HD2	1.77	0.65
1:B:326:ASP:O	1:B:354:LYS:HD3	1.98	0.64
1:A:202:ARG:HD2	1:A:207:ARG:NH1	2.14	0.63
1:A:90:ALA:HB2	1:A:122:LEU:HD21	1.81	0.62
1:B:437:LEU:HD13	1:B:441:HIS:ND1	2.19	0.58
1:B:325:ALA:O	1:B:355:VAL:HG22	2.05	0.57
1:B:249:LEU:HD11	1:B:344:MET:HA	1.87	0.54
1:A:202:ARG:HD2	1:A:207:ARG:HH11	1.74	0.53
1:B:90:ALA:HB2	1:B:122:LEU:HD21	1.88	0.53
1:A:3:ASP:O	1:A:381:ARG:NH1	2.42	0.53
1:B:197:HIS:O	1:B:211:PRO:HB3	2.10	0.52
1:B:8:ASN:OD1	1:B:53:SER:OG	2.27	0.51
1:A:262:GLN:OE1	1:A:272:SER:HB2	2.11	0.50
1:B:239:ARG:HH11	1:B:239:ARG:HG2	1.79	0.48
1:B:202:ARG:HH21	1:B:207:ARG:HH12	1.61	0.47
1:A:177:ASP:OD2	1:A:451:SER:HB3	2.14	0.47
1:B:249:LEU:HD12	1:B:343:LEU:O	2.15	0.46
1:B:112:PRO:CG	1:B:167:VAL:HG11	2.46	0.45
1:B:160:ILE:HG23	1:B:205:ARG:HD3	1.99	0.45
1:B:120:SER:HB2	1:B:214:THR:HG21	1.98	0.44
1:B:278:VAL:HG12	1:B:389:ASN:HB2	2.00	0.43
1:A:285:VAL:HA	1:A:368:GLY:O	2.19	0.42
1:B:90:ALA:HB2	1:B:122:LEU:CD2	2.49	0.42
1:A:286:GLY:O	1:A:369:THR:HA	2.19	0.42
1:B:112:PRO:HG2	1:B:167:VAL:HG11	2.00	0.42
1:B:2:ARG:NH2	1:B:52:ASP:OD2	2.53	0.42
1:B:5:LEU:HD21	1:B:381:ARG:HG3	2.01	0.42
1:B:102:THR:HG21	1:B:225:PHE:CE1	2.55	0.42
1:B:77:GLY:HA2	1:B:80:PHE:O	2.21	0.41
1:A:18:MET:HG3	1:A:391:ALA:HB1	2.02	0.41
1:A:409:GLU:OE1	1:A:412:ARG:NH2	2.41	0.41
1:B:18:MET:HG3	1:B:391:ALA:HB1	2.03	0.40
1:B:89:VAL:HG13	1:B:101:LEU:HB3	2.03	0.40
1:B:421:ARG:HD2	1:B:422:PRO:HD2	2.04	0.40
1:B:93:ILE:HB	1:B:94:PRO:HD3	2.02	0.40

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	449/451 (100%)	429 (96%)	20 (4%)	0	100	100
1	B	449/451 (100%)	432 (96%)	17 (4%)	0	100	100
All	All	898/902 (100%)	861 (96%)	37 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	366/366 (100%)	363 (99%)	3 (1%)	81	93
1	B	366/366 (100%)	357 (98%)	9 (2%)	47	73
All	All	732/732 (100%)	720 (98%)	12 (2%)	62	84

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	247	LEU
1	A	335	GLU
1	A	445	ARG
1	B	1	MET
1	B	69	SER
1	B	155	ASP
1	B	247	LEU
1	B	330	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	334	THR
1	B	350	ASP
1	B	431	PRO
1	B	451	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	HIS
1	B	30	GLN
1	B	144	GLN
1	B	219	GLN

### 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	451/451 (100%)	0.09	24 (5%)	26 28	21, 36, 82, 117	0
1	B	451/451 (100%)	0.20	24 (5%)	26 28	22, 36, 78, 102	0
All	All	902/902 (100%)	0.14	48 (5%)	26 28	21, 36, 80, 117	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	66	TYR	8.2
1	B	323	ALA	6.7
1	B	68	ASP	6.3
1	A	1	MET	6.2
1	B	355	VAL	6.1
1	A	316	LEU	4.9
1	B	322	GLY	4.7
1	B	347	ALA	4.7
1	A	319	ASP	4.4
1	A	349	PRO	4.2
1	B	324	ASP	3.8
1	A	318	VAL	3.7
1	A	347	ALA	3.6
1	B	340	ILE	3.5
1	A	320	LEU	3.4
1	A	348	VAL	3.3
1	A	3	ASP	3.2
1	B	1	MET	3.1
1	B	66	TYR	3.0
1	A	155	ASP	3.0
1	A	65	ALA	2.9
1	B	89	VAL	2.9
1	A	451	SER	2.9
1	A	68	ASP	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	353	PRO	2.8
1	B	344	MET	2.8
1	B	346	GLU	2.8
1	B	335	GLU	2.7
1	A	353	PRO	2.7
1	A	322	GLY	2.7
1	B	86	ALA	2.6
1	B	337	THR	2.6
1	A	67	GLY	2.6
1	B	325	ALA	2.5
1	B	326	ASP	2.5
1	B	155	ASP	2.4
1	B	339	GLY	2.4
1	A	350	ASP	2.3
1	A	149	ASP	2.3
1	A	90	ALA	2.3
1	B	321	SER	2.3
1	A	330	ASP	2.2
1	B	327	GLN	2.2
1	A	324	ASP	2.2
1	A	91	ALA	2.1
1	B	358	LEU	2.1
1	B	432	GLY	2.1
1	A	305	VAL	2.0

## 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.