Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 07:35 am BST

PDB ID	:	5XAV
Title	:	Structure of PhaC from Chromobacterium sp. USM2
Authors	:	Chek, M.F.; Kim, S.Y.; Mori, T.; Arsad, H.; Samian, M.R.; Sudesh, K.; Hakoshima, T.
Deposited on	:	2017-03-15
Resolution	:	1.48 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 following versions of software and data (see references ①) were used in the production of this report:

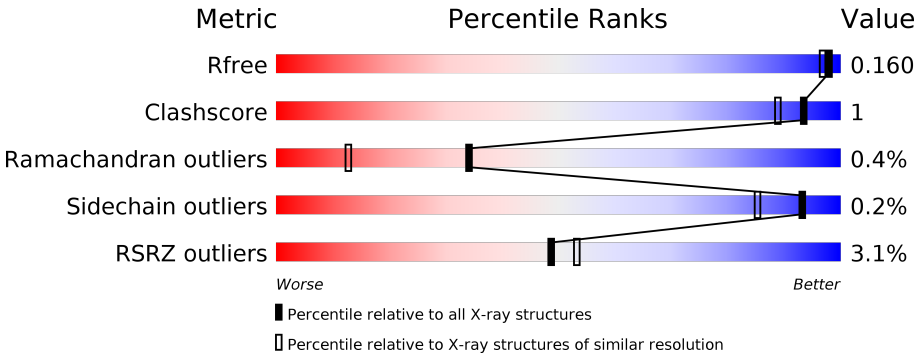
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.48 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	4690 (1.50-1.46)
Clashscore	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

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 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 <=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	395	3%, 90%, 5%
1	B	395	3%, 90%, 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6968 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 Intracellular polyhydroxyalkanoate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	376	Total	C	N	O	S	0	0	0
			2952	1923	486	528	15			
1	B	378	Total	C	N	O	S	0	0	0
			2967	1932	489	531	15			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	173	GLY	-	expression tag	UNP E1APK1
A	174	PRO	-	expression tag	UNP E1APK1
B	173	GLY	-	expression tag	UNP E1APK1
B	174	PRO	-	expression tag	UNP E1APK1

- Molecule 2 is water.

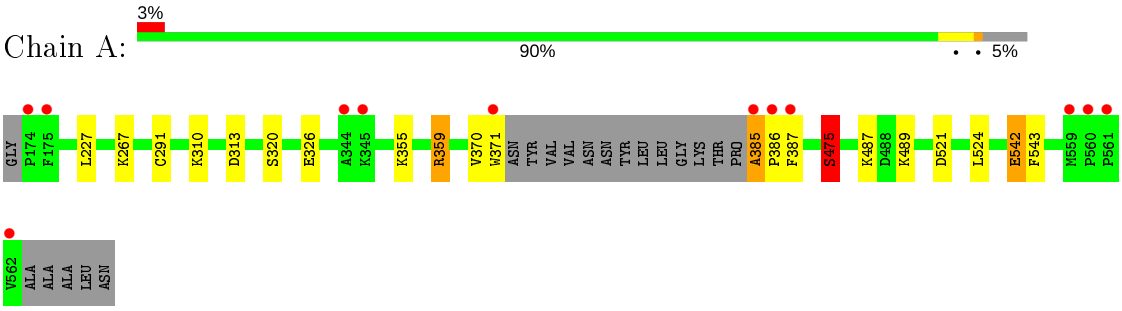
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	548	Total	O	0	0
			548	548		
2	B	501	Total	O	0	0
			501	501		

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.

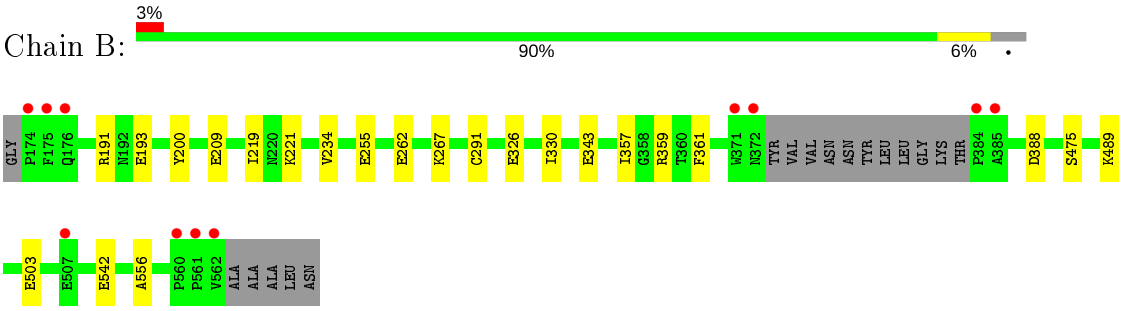
- Molecule 1: Intracellular polyhydroxyalkanoate synthase

Chain A:



- Molecule 1: Intracellular polyhydroxyalkanoate synthase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	117.33Å 117.33Å 105.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.10 – 1.48 39.31 – 1.48	Depositor EDS
% Data completeness (in resolution range)	99.8 (36.10-1.48) 99.8 (39.31-1.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.67 (at 1.48Å)	Xtriage
Refinement program	REFMAC, Coot, PHENIX (1.10.1_2155)	Depositor
R, R_{free}	0.121 , 0.157 0.123 , 0.160	Depositor DCC
R_{free} test set	6910 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	15.5	Xtriage
Anisotropy	0.115	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 50.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.028 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	6968	wwPDB-VP
Average B, all atoms (Å^2)	22.0	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 30.39 % 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 1.3176e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.99	6/3041 (0.2%)	0.85	3/4146 (0.1%)
1	B	1.00	9/3057 (0.3%)	0.84	4/4168 (0.1%)
All	All	0.99	15/6098 (0.2%)	0.85	7/8314 (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

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	388	ASP	CB-CG	7.41	1.67	1.51
1	B	255	GLU	CD-OE2	7.11	1.33	1.25
1	B	326	GLU	CD-OE1	6.72	1.33	1.25
1	A	542	GLU	CD-OE1	6.31	1.32	1.25
1	B	209	GLU	CD-OE2	6.03	1.32	1.25
1	A	475	SER	CB-OG	-6.03	1.34	1.42
1	B	343	GLU	CD-OE2	-5.71	1.19	1.25
1	A	524	LEU	CB-CG	-5.54	1.36	1.52
1	B	262	GLU	CD-OE1	-5.45	1.19	1.25
1	A	326	GLU	CD-OE1	5.35	1.31	1.25
1	A	386	PRO	C-O	-5.20	1.12	1.23
1	B	542	GLU	CG-CD	5.19	1.59	1.51
1	B	503	GLU	CG-CD	5.07	1.59	1.51
1	B	542	GLU	CD-OE1	5.06	1.31	1.25
1	A	320	SER	CB-OG	-5.04	1.35	1.42

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	524	LEU	CA-CB-CG	6.78	130.90	115.30
1	A	359	ARG	NE-CZ-NH1	6.49	123.55	120.30
1	B	359	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	B	359	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	521	ASP	CB-CG-OD1	5.52	123.27	118.30
1	B	191	ARG	NE-CZ-NH1	5.21	122.91	120.30
1	B	388	ASP	CB-CG-OD1	5.05	122.84	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	385	ALA	Peptide

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	2952	0	2937	8	0
1	B	2967	0	2951	9	0
2	A	548	0	0	2	1
2	B	501	0	0	2	2
All	All	6968	0	5888	17	3

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 1.

All (17) 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:291:CYS:SG	2:B:944:HOH:O	2.29	0.89
1:B:489:LYS:NZ	2:B:601:HOH:O	2.04	0.82
1:A:291:CYS:SG	2:A:998:HOH:O	2.19	0.61
1:B:193:GLU:OE1	1:B:267:LYS:NZ	2.30	0.61
1:B:200:TYR:CE2	1:B:234:VAL:HG11	2.36	0.60
1:B:193:GLU:H	1:B:193:GLU:CD	2.05	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:200:TYR:CZ	1:B:234:VAL:HG11	2.42	0.55
1:A:355:LYS:HE2	1:A:359:ARG:NH2	2.24	0.53
1:B:221:LYS:NZ	1:B:556:ALA:O	2.36	0.51
1:A:487:LYS:HB3	1:A:489:LYS:HE3	1.93	0.51
1:A:267:LYS:NZ	2:A:602:HOH:O	2.40	0.49
1:A:310:LYS:HE3	1:A:313:ASP:OD1	2.12	0.49
1:A:370:VAL:HA	1:A:371:TRP:C	2.34	0.48
1:B:330:ILE:HD12	1:B:361:PHE:CE1	2.50	0.46
1:B:219:ILE:HG23	1:B:357:ILE:CG2	2.47	0.45
1:A:385:ALA:HA	1:A:387:PHE:H	1.81	0.45
1:A:542:GLU:HG2	1:A:543:PHE:CE2	2.52	0.45

All (3) 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:809:HOH:O	2:A:1013:HOH:O[5_554]	1.66	0.54
2:B:1067:HOH:O	2:B:1089:HOH:O[6_445]	2.05	0.15
2:B:696:HOH:O	2:B:966:HOH:O[6_445]	2.05	0.15

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	372/395 (94%)	359 (96%)	11 (3%)	2 (0%)	29 9
1	B	374/395 (95%)	364 (97%)	9 (2%)	1 (0%)	41 18
All	All	746/790 (94%)	723 (97%)	20 (3%)	3 (0%)	34 13

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	475	SER
1	A	227	LEU
1	A	475	SER

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	316/330 (96%)	315 (100%)	1 (0%)	92 84
1	B	318/330 (96%)	318 (100%)	0	100 100
All	All	634/660 (96%)	633 (100%)	1 (0%)	93 85

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	475	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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	376/395 (95%)	-0.08	12 (3%) 47 52	10, 17, 38, 59	0
1	B	378/395 (95%)	-0.02	11 (2%) 51 56	12, 18, 39, 54	0
All	All	754/790 (95%)	-0.05	23 (3%) 49 53	10, 17, 39, 59	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	371	TRP	10.6
1	B	562	VAL	10.4
1	B	174	PRO	9.9
1	A	562	VAL	9.2
1	B	175	PHE	7.0
1	B	561	PRO	6.7
1	A	174	PRO	6.5
1	A	561	PRO	5.9
1	B	372	ASN	5.6
1	B	384	PRO	5.5
1	A	560	PRO	4.8
1	B	371	TRP	4.6
1	A	175	PHE	4.6
1	A	344	ALA	4.2
1	B	385	ALA	3.6
1	A	385	ALA	3.5
1	A	387	PHE	3.3
1	B	560	PRO	3.2
1	A	386	PRO	2.6
1	A	559	MET	2.4
1	B	176	GLN	2.3
1	B	507	GLU	2.2
1	A	345	LYS	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.