



Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 01:06 am BST

PDB ID : 6OBT
Title : Structural insights into dehydratase substrate selection for the borrelidin and fluvirucin polyketide synthases
Authors : McAndrew, R.P.; Barajas, J.F.; Pereira, J.H.; Keasling, J.D.; Adams, P.D.
Deposited on : 2019-03-21
Resolution : 1.80 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 ⓘ](#)) 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

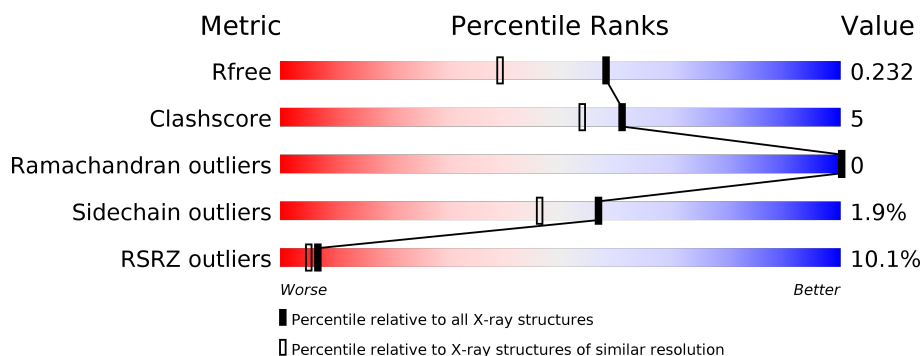
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	314	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2080 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 Borrelidin polyketide synthase, type I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	268	Total	C	N	O	S	0	0	0
			2032	1286	359	382	5			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q70I00
A	2	GLY	-	expression tag	UNP Q70I00
A	3	SER	-	expression tag	UNP Q70I00
A	4	SER	-	expression tag	UNP Q70I00
A	5	HIS	-	expression tag	UNP Q70I00
A	6	HIS	-	expression tag	UNP Q70I00
A	7	HIS	-	expression tag	UNP Q70I00
A	8	HIS	-	expression tag	UNP Q70I00
A	9	HIS	-	expression tag	UNP Q70I00
A	10	HIS	-	expression tag	UNP Q70I00
A	11	ASP	-	expression tag	UNP Q70I00
A	12	TYR	-	expression tag	UNP Q70I00
A	13	ASP	-	expression tag	UNP Q70I00
A	14	ILE	-	expression tag	UNP Q70I00
A	15	PRO	-	expression tag	UNP Q70I00
A	16	THR	-	expression tag	UNP Q70I00
A	17	THR	-	expression tag	UNP Q70I00
A	18	GLU	-	expression tag	UNP Q70I00
A	19	ASN	-	expression tag	UNP Q70I00
A	20	LEU	-	expression tag	UNP Q70I00
A	21	TYR	-	expression tag	UNP Q70I00
A	22	PHE	-	expression tag	UNP Q70I00
A	23	GLN	-	expression tag	UNP Q70I00
A	24	GLY	-	expression tag	UNP Q70I00
A	25	HIS	-	expression tag	UNP Q70I00
A	26	MET	-	expression tag	UNP Q70I00
A	349	VAL	-	expression tag	UNP Q70I00

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	350	LEU	-	expression tag	UNP Q70I00
A	351	ARG	-	expression tag	UNP Q70I00
A	352	SER	-	expression tag	UNP Q70I00
A	353	ALA	-	expression tag	UNP Q70I00
A	353A	ALA	-	expression tag	UNP Q70I00
A	353B	ALA	-	expression tag	UNP Q70I00
A	353C	ARG	-	expression tag	UNP Q70I00
A	353D	ARG	-	expression tag	UNP Q70I00
A	354	THR	-	expression tag	UNP Q70I00
A	355	GLY	-	expression tag	UNP Q70I00
A	356	ALA	-	expression tag	UNP Q70I00
A	357	ARG	-	expression tag	UNP Q70I00
A	358	ARG	-	expression tag	UNP Q70I00
A	359	GLN	-	expression tag	UNP Q70I00
A	360	ALA	-	expression tag	UNP Q70I00

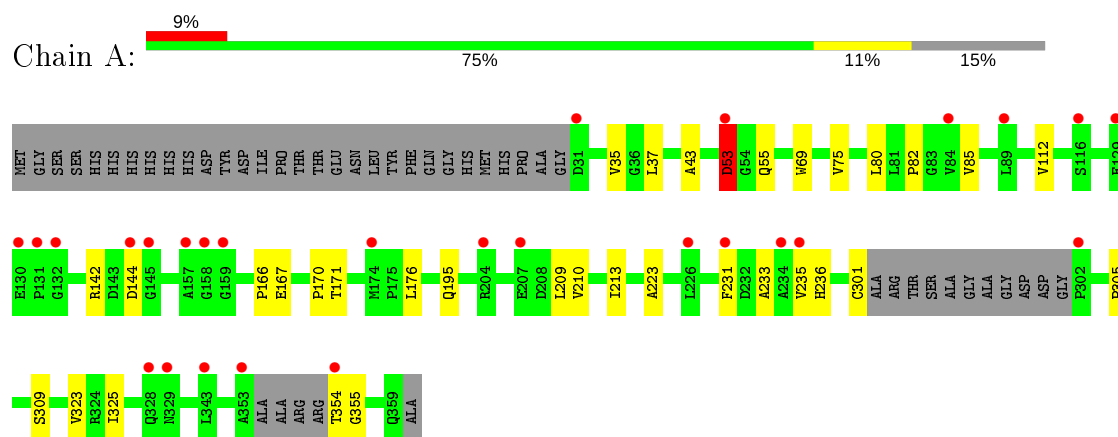
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	48	Total O 48 48	0	0

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: Borrelidin polyketide synthase, type I



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	108.21Å 63.35Å 40.67Å 90.00° 100.44° 90.00°	Depositor
Resolution (Å)	30.89 – 1.80 30.89 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.3 (30.89-1.80) 98.4 (30.89-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.60 (at 1.80Å)	Xtriage
Refinement program	PHENIX 1.15rc3_3435, PHENIX 1.15rc3_3435	Depositor
R, R_{free}	0.203 , 0.232 0.203 , 0.232	Depositor DCC
R_{free} test set	1245 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	33.2	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 53.7	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	0.97	EDS
Total number of atoms	2080	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.94% 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	0.36	0/2087	0.57	0/2858

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	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	171	THR	Peptide
1	A	53	ASP	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	2032	0	1963	18	0
2	A	48	0	0	0	0
All	All	2080	0	1963	18	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 5.

All (18) 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:A:354:THR:HG22	1:A:355:GLY:H	1.58	0.69
1:A:112:VAL:O	1:A:142:ARG:NH2	2.27	0.68
1:A:235:VAL:HG11	1:A:305:PRO:HG3	1.78	0.65
1:A:213:ILE:HD12	1:A:323:VAL:HG21	1.81	0.61
1:A:233:ALA:HA	1:A:236:HIS:CE1	2.38	0.59
1:A:209:LEU:HD23	1:A:325:ILE:HD12	1.90	0.54
1:A:75:VAL:HG12	1:A:80:LEU:HG	1.89	0.54
1:A:53:ASP:HB3	1:A:55:GLN:HB2	1.91	0.51
1:A:170:PRO:HG2	1:A:210:VAL:HG11	1.93	0.51
1:A:69:TRP:HB2	1:A:195:GLN:HG3	1.95	0.47
1:A:43:ALA:HB2	1:A:223:ALA:HB3	1.97	0.46
1:A:80:LEU:CD2	1:A:112:VAL:HG22	2.46	0.46
1:A:35:VAL:HG23	1:A:37:LEU:HG	1.97	0.46
1:A:176:LEU:HD23	1:A:301:CYS:HB2	1.98	0.46
1:A:82:PRO:HG2	1:A:85:VAL:HG23	1.97	0.46
1:A:80:LEU:HD22	1:A:112:VAL:HG22	1.98	0.45
1:A:166:PRO:HB2	1:A:167:GLU:OE1	2.17	0.44
1:A:53:ASP:HB3	1:A:55:GLN:H	1.84	0.43

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	261/314 (83%)	253 (97%)	8 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	213 / 246 (87%)	209 (98%)	4 (2%)	57 46

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	ASP
1	A	144	ASP
1	A	231	PHE
1	A	309	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 [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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	239:ALA	C	301:CYS	N	2.94

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, 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	268/314 (85%)	0.45	27 (10%) 7 5	31, 53, 90, 120	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	129	GLU	4.1
1	A	132	GLY	3.7
1	A	302	PRO	3.7
1	A	116	SER	3.5
1	A	131	PRO	3.5
1	A	231	PHE	3.3
1	A	328	GLN	3.3
1	A	174	MET	3.1
1	A	144	ASP	3.0
1	A	89	LEU	2.8
1	A	354	THR	2.8
1	A	159	GLY	2.8
1	A	53	ASP	2.6
1	A	234	ALA	2.6
1	A	84	VAL	2.6
1	A	31	ASP	2.5
1	A	157	ALA	2.5
1	A	343	LEU	2.4
1	A	226	LEU	2.2
1	A	145	GLY	2.2
1	A	353	ALA	2.2
1	A	207	GLU	2.2
1	A	204	ARG	2.2
1	A	158	GLY	2.1
1	A	130	GLU	2.1
1	A	235	VAL	2.0
1	A	329	ASN	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.