wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2023 – 08:44 AM EDT

PDB ID	:	4LA5
Title	:	Crystal structure of 2-methylisoborneol synthase from Streptomyces coelicolor A3(2)
Authors	:	Koksal, M.; Christianson, D.W.
Deposited on	:	2013-06-19
Resolution	:	1.85 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

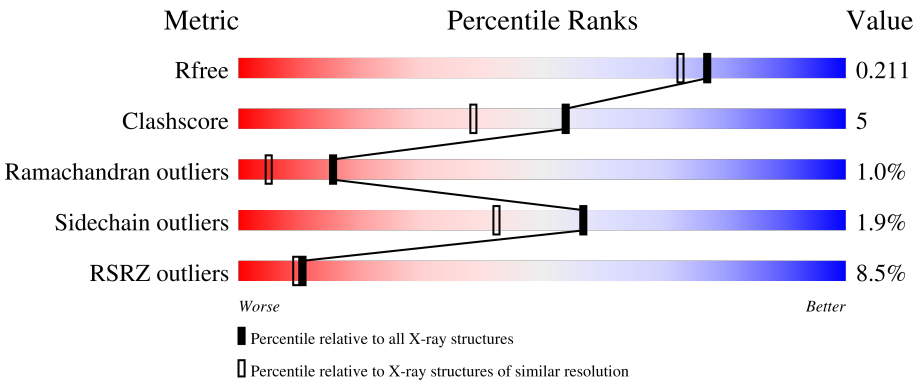
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 <=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	461	6% 53% 14% • 31%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2679 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 2-methylisoborneol synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	319	2534	1600	458	463	13	0	2	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP Q9F1Y6
A	-19	GLY	-	expression tag	UNP Q9F1Y6
A	-18	SER	-	expression tag	UNP Q9F1Y6
A	-17	SER	-	expression tag	UNP Q9F1Y6
A	-16	HIS	-	expression tag	UNP Q9F1Y6
A	-15	HIS	-	expression tag	UNP Q9F1Y6
A	-14	HIS	-	expression tag	UNP Q9F1Y6
A	-13	HIS	-	expression tag	UNP Q9F1Y6
A	-12	HIS	-	expression tag	UNP Q9F1Y6
A	-11	HIS	-	expression tag	UNP Q9F1Y6
A	-10	SER	-	expression tag	UNP Q9F1Y6
A	-9	SER	-	expression tag	UNP Q9F1Y6
A	-8	GLY	-	expression tag	UNP Q9F1Y6
A	-7	LEU	-	expression tag	UNP Q9F1Y6
A	-6	VAL	-	expression tag	UNP Q9F1Y6
A	-5	PRO	-	expression tag	UNP Q9F1Y6
A	-4	ARG	-	expression tag	UNP Q9F1Y6
A	-3	GLY	-	expression tag	UNP Q9F1Y6
A	-2	SER	-	expression tag	UNP Q9F1Y6
A	-1	HIS	-	expression tag	UNP Q9F1Y6
A	0	MET	-	expression tag	UNP Q9F1Y6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	145	145	145	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: 2-methylisoborneol synthase

Chain A: 6% 53% 14% • 31%

MET GLY SER SER HIS HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET MET PRO ASP SER GLY THR LEU GLY THR PRO PRO PRO GLU GLN GLY PRO THR PRO PRO THR THR LEU PRO ASP VAL PRO ALA PRO VAL ILE PRO SER ALA SER VAL THR SER ALA ALA

SER ASP PHE LEU ALA ALA LEU HIS PRO PRO VAL THR VAL PRO ASP PRO ALA PRO PRO PRO PRO PRO ALA PRO ALA ALA GLY ASN PRO PRO ASP THR VAL THR GLY ASP SER VAL LEU GLN ARG ILE LEU ARG GLY PRO THR GLY PRO GLY THR THR SER LEU ALA PRO ALA VAL ARG TYR

GLY ARG GLN PRO GLY PRO GLU ALA PRO ALA SER A111 P112 P113 A114 A115 G116 R117 L122 Y123 P126 E129 P130 D131 V135 E136 E137 R140 R144 Y153 P154 GLU GLU TRP GLU GLY GLN F161 F164 Y169 T187 R188 L189 M190 E193 D197 D198 C199

Y200 C201 GLU ASP HIS GLY GLY S207 P208 V209 R214 L215 L216 F225 H226 S227 T232 P233 T234 W235 S238 A243 P244 R245 R246 S250 V256 R257 A258 A259 Q263 S264 Y267 R268 E281 W284 P291 E292 V293 W294 E295 Y296 L297 Q301 Y318

D329 R332 N339 A340 L354 N355 S356 P357 G358 R359 E371 Q372 L373 C374 E375 R376 D377 L380 E384 H391 R414 G415 W426 H427 R428 T429 N430 T431 Y432 R433 Y434 S435 L436 P437 D438 F439 W440

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	99.45Å 99.45Å 104.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.73 – 1.85 49.73 – 1.85	Depositor EDS
% Data completeness (in resolution range)	93.8 (49.73-1.85) 93.9 (49.73-1.85)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 1.86Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.186 , 0.214 0.182 , 0.211	Depositor DCC
R_{free} test set	1983 reflections (4.36%)	wwPDB-VP
Wilson B-factor (Å^2)	29.2	Xtriage
Anisotropy	0.267	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 39.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2679	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.22% of the height of the origin peak. No significant pseudotranslation is detected.*

[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 (i)

5.1 Standard geometry (i)

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	1.65	33/2623 (1.3%)	1.31	21/3584 (0.6%)

The worst 5 of 33 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	193	GLU	CG-CD	9.27	1.65	1.51
1	A	193	GLU	CD-OE2	9.02	1.35	1.25
1	A	295	GLU	CG-CD	8.94	1.65	1.51
1	A	268	ARG	CB-CG	7.96	1.74	1.52
1	A	332	ARG	CZ-NH1	7.33	1.42	1.33

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	188	ARG	NE-CZ-NH1	12.15	126.38	120.30
1	A	188	ARG	NE-CZ-NH2	-10.99	114.80	120.30
1	A	257[A]	ARG	NE-CZ-NH1	10.61	125.60	120.30
1	A	257[B]	ARG	NE-CZ-NH1	10.61	125.60	120.30
1	A	257[A]	ARG	NE-CZ-NH2	-8.54	116.03	120.30

There are no chirality outliers.

There are no planarity outliers.

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	2534	0	2404	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	145	0	0	3	0
All	All	2679	0	2404	26	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.

The worst 5 of 26 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:257[A]:ARG:NH2	2:A:598:HOH:O	1.57	1.29
1:A:374:CYS:SG	1:A:377:ASP:OD2	2.22	0.97
1:A:257[A]:ARG:CZ	2:A:598:HOH:O	2.03	0.88
1:A:112:PRO:HG3	1:A:438:ASP:HB2	1.79	0.63
1:A:153:TYR:CE2	1:A:161:PHE:HB2	2.37	0.59

There are no symmetry-related clashes.

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	315/461 (68%)	304 (96%)	8 (2%)	3 (1%)	15 5

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	372	GLN
1	A	357	PRO
1	A	356	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	262/372 (70%)	257 (98%)	5 (2%)	57 42

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	153	TYR
1	A	169	TYR
1	A	198	ASP
1	A	234	THR
1	A	430	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	391	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 monosaccharides 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	319/461 (69%)	0.26	27 (8%) 10 9	17, 32, 70, 108	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	357	PRO	7.4
1	A	112	PRO	7.3
1	A	111	ALA	7.0
1	A	432	TYR	6.1
1	A	114	ALA	5.9

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.