



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 6, 2022 – 10:22 am BST

PDB ID : 7QCW
Title : Apo-structure of serine hydroxymethyltransferase (PbzB) involved in benzobactin biosynthesis in *P. chlororaphis* subsp. *piscium* DSM 21509
Authors : Czech, L.; Bange, G.
Deposited on : 2021-11-25
Resolution : 2.81 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.31.2
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

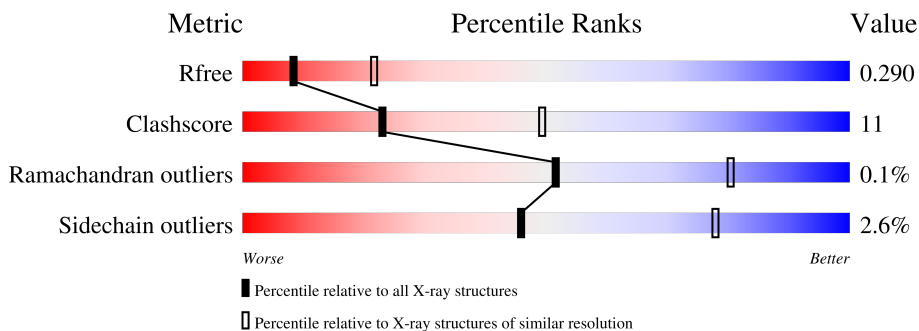
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)

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 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	460	 68% 18% 14%
1	B	460	 62% 25% 11%

2 Entry composition i

There is only 1 type of molecule in this entry. The entry contains 6027 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 Serine hydroxymethyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	396	2983	1852	542	572	17	0	0	0
1	B	408	3044	1886	555	585	18	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	SER	-	expression tag	UNP A0A3G7DQ80
A	-6	GLY	-	expression tag	UNP A0A3G7DQ80
A	-5	SER	-	expression tag	UNP A0A3G7DQ80
A	-4	HIS	-	expression tag	UNP A0A3G7DQ80
A	-3	MET	-	expression tag	UNP A0A3G7DQ80
A	-2	THR	-	expression tag	UNP A0A3G7DQ80
A	-1	MET	-	expression tag	UNP A0A3G7DQ80
A	0	GLY	-	expression tag	UNP A0A3G7DQ80
A	23	GLU	GLN	conflict	UNP A0A3G7DQ80
A	27	GLU	LYS	conflict	UNP A0A3G7DQ80
A	30	ARG	GLN	conflict	UNP A0A3G7DQ80
A	55	ARG	LEU	conflict	UNP A0A3G7DQ80
A	245	VAL	ALA	conflict	UNP A0A3G7DQ80
A	270	SER	ALA	conflict	UNP A0A3G7DQ80
A	289	MET	ILE	conflict	UNP A0A3G7DQ80
A	400	ARG	GLN	conflict	UNP A0A3G7DQ80
A	416	LEU	ILE	conflict	UNP A0A3G7DQ80
A	429	LYS	GLN	conflict	UNP A0A3G7DQ80
A	450	ILE	-	expression tag	UNP A0A3G7DQ80
A	451	LEU	-	expression tag	UNP A0A3G7DQ80
A	452	GLU	-	expression tag	UNP A0A3G7DQ80
B	-7	SER	-	expression tag	UNP A0A3G7DQ80
B	-6	GLY	-	expression tag	UNP A0A3G7DQ80
B	-5	SER	-	expression tag	UNP A0A3G7DQ80
B	-4	HIS	-	expression tag	UNP A0A3G7DQ80

Continued on next page...

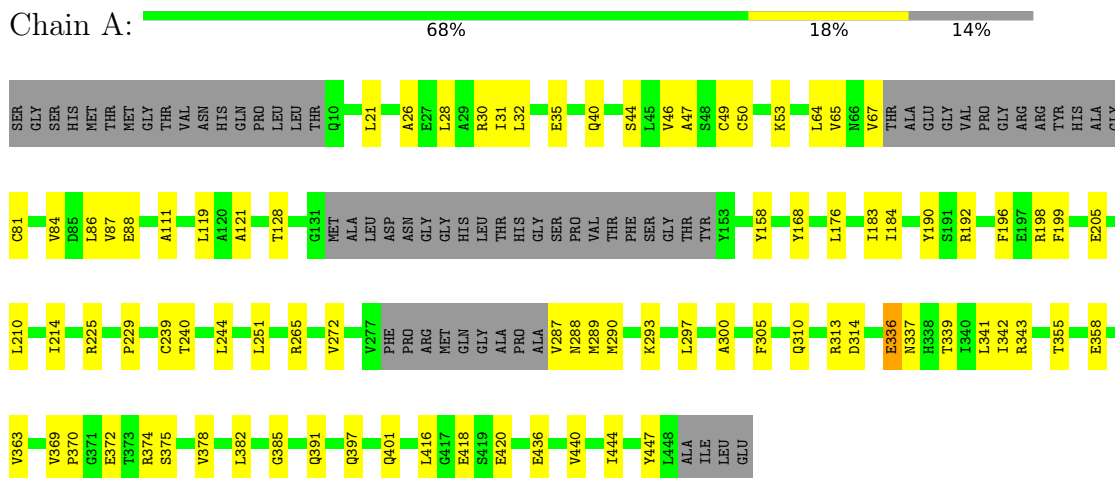
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	MET	-	expression tag	UNP A0A3G7DQ80
B	-2	THR	-	expression tag	UNP A0A3G7DQ80
B	-1	MET	-	expression tag	UNP A0A3G7DQ80
B	0	GLY	-	expression tag	UNP A0A3G7DQ80
B	23	GLU	GLN	conflict	UNP A0A3G7DQ80
B	27	GLU	LYS	conflict	UNP A0A3G7DQ80
B	30	ARG	GLN	conflict	UNP A0A3G7DQ80
B	55	ARG	LEU	conflict	UNP A0A3G7DQ80
B	245	VAL	ALA	conflict	UNP A0A3G7DQ80
B	270	SER	ALA	conflict	UNP A0A3G7DQ80
B	289	MET	ILE	conflict	UNP A0A3G7DQ80
B	400	ARG	GLN	conflict	UNP A0A3G7DQ80
B	416	LEU	ILE	conflict	UNP A0A3G7DQ80
B	429	LYS	GLN	conflict	UNP A0A3G7DQ80
B	450	ILE	-	expression tag	UNP A0A3G7DQ80
B	451	LEU	-	expression tag	UNP A0A3G7DQ80
B	452	GLU	-	expression tag	UNP A0A3G7DQ80

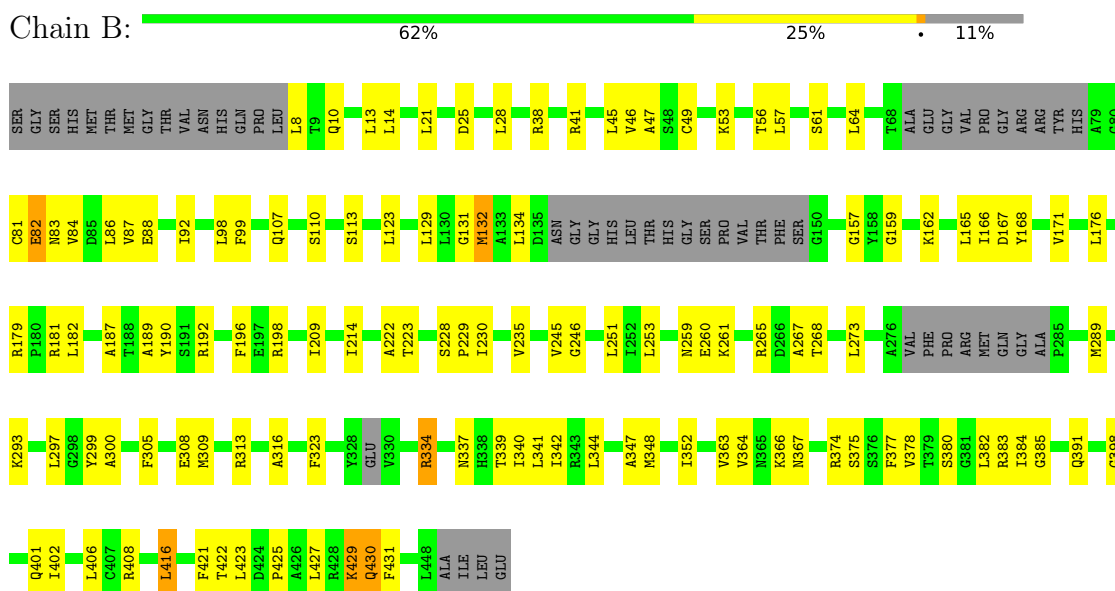
3 Residue-property plots

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. 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. 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: Serine hydroxymethyltransferase



- Molecule 1: Serine hydroxymethyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	132.40Å 59.64Å 112.12Å 90.00° 109.41° 90.00°	Depositor
Resolution (Å)	49.21 – 2.81 49.21 – 2.81	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.21-2.81) 99.7 (49.21-2.81)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.81Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.236 , 0.295 0.238 , 0.290	Depositor DCC
R_{free} test set	1016 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	68.5	Xtriage
Anisotropy	0.472	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	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	0.91	EDS
Total number of atoms	6027	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.10% 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 [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	0.30	0/3020	0.58	2/4088 (0.0%)
1	B	0.37	0/3079	0.64	1/4166 (0.0%)
All	All	0.34	0/6099	0.61	3/8254 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	86	LEU	CB-CG-CD2	-5.22	102.13	111.00
1	A	336	GLU	N-CA-C	-5.12	97.19	111.00
1	B	82	GLU	CA-CB-CG	5.09	124.61	113.40

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	2983	0	2988	58	0
1	B	3044	0	3051	83	1
All	All	6027	0	6039	137	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 11.

The worst 5 of 137 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:21:LEU:HD11	1:B:28:LEU:HD23	1.57	0.84
1:B:47:ALA:HA	1:B:385:GLY:HA3	1.64	0.79
1:A:378:VAL:HG21	1:A:418:GLU:HB3	1.64	0.79
1:B:214:ILE:HG21	1:B:229:PRO:HG2	1.64	0.78
1:A:40:GLN:HE22	1:A:444:ILE:H	1.32	0.77

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:260:GLU:OE2	1:B:261:LYS:CD[2_655]	1.75	0.45

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	388/460 (84%)	377 (97%)	11 (3%)	0	100	100
1	B	398/460 (86%)	383 (96%)	14 (4%)	1 (0%)	41	70
All	All	786/920 (85%)	760 (97%)	25 (3%)	1 (0%)	51	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	425	PRO

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	308/358 (86%)	306 (99%)	2 (1%)	86	95
1	B	312/358 (87%)	298 (96%)	14 (4%)	27	59
All	All	620/716 (87%)	604 (97%)	16 (3%)	46	78

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	429	LYS
1	B	427	LEU
1	B	337	ASN
1	B	423	LEU
1	B	334	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	83	ASN
1	B	337	ASN
1	B	401	GLN
1	B	397	GLN
1	B	10	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 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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.