



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:00 am BST

PDB ID : 6F93
Title : Helicobacter pylori serine hydroxymethyl transferase in apo form
Authors : Sodolescu, A.; Dian, C.; Terradot, L.; Bouzahir-Sima, L.; Lestini, R.; Myllykallio, H.; Skouloubris, S.; Liebl, U.
Deposited on : 2017-12-13
Resolution : 2.80 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 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

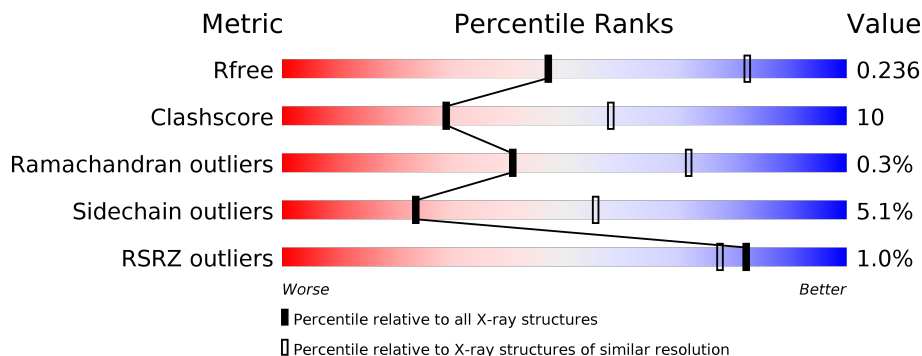
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	424	 % 67% 22% • 9%
1	B	424	 % 70% 19% • 9%

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 5909 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	387	Total	C	N	O	S	0	1	0
			2948	1876	503	556	13			
1	B	385	Total	C	N	O	S	0	3	0
			2945	1876	499	558	12			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	ARG	-	expression tag	UNP P56089
A	418	SER	-	expression tag	UNP P56089
A	419	HIS	-	expression tag	UNP P56089
A	420	HIS	-	expression tag	UNP P56089
A	421	HIS	-	expression tag	UNP P56089
A	422	HIS	-	expression tag	UNP P56089
A	423	HIS	-	expression tag	UNP P56089
A	424	HIS	-	expression tag	UNP P56089
B	417	ARG	-	expression tag	UNP P56089
B	418	SER	-	expression tag	UNP P56089
B	419	HIS	-	expression tag	UNP P56089
B	420	HIS	-	expression tag	UNP P56089
B	421	HIS	-	expression tag	UNP P56089
B	422	HIS	-	expression tag	UNP P56089
B	423	HIS	-	expression tag	UNP P56089
B	424	HIS	-	expression tag	UNP P56089

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	5	Total	O	0	0
			5	5		
2	B	11	Total	O	0	0
			11	11		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	57.35Å 87.55Å 162.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.85 – 2.80 46.85 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.6 (46.85-2.80) 98.6 (46.85-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.62 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.194 , 0.236 0.194 , 0.236	Depositor DCC
R_{free} test set	1287 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	39.9	Xtrriage
Anisotropy	0.733	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 39.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5909	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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	0.31	0/3003	0.50	1/4058 (0.0%)
1	B	0.31	0/3002	0.51	1/4062 (0.0%)
All	All	0.31	0/6005	0.50	2/8120 (0.0%)

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	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	315	LEU	CA-CB-CG	6.69	130.69	115.30
1	A	314	LEU	CA-CB-CG	5.14	127.13	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	310	SER	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	2948	0	2907	63	0
1	B	2945	0	2853	59	0
2	A	5	0	0	0	0
2	B	11	0	0	1	0
All	All	5909	0	5760	112	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 10.

The worst 5 of 112 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:145:ASN:HB3	1:A:147:ASP:H	1.46	0.78
1:B:312:HIS:CE1	1:B:313:LEU:HD23	2.19	0.78
1:A:272:LYS:NZ	1:B:8:THR:O	2.17	0.76
1:B:207:THR:HG21	1:B:283:ALA:HB3	1.66	0.75
1:A:174:TYR:O	1:A:311:ASN:ND2	2.23	0.71

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	382/424 (90%)	364 (95%)	16 (4%)	2 (0%)	29	61
1	B	382/424 (90%)	365 (96%)	17 (4%)	0	100	100
All	All	764/848 (90%)	729 (95%)	33 (4%)	2 (0%)	41	72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	253	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	308/350 (88%)	290 (94%)	18 (6%)	20	50
1	B	303/350 (87%)	289 (95%)	14 (5%)	27	60
All	All	611/700 (87%)	579 (95%)	32 (5%)	24	55

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	350	SER
1	A	391	VAL
1	B	315	LEU
1	A	373	GLU
1	B	10	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	225	HIS
1	B	379	ASN
1	B	340	ASN
1	B	210	HIS
1	B	311	ASN

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 [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	387/424 (91%)	-0.48	4 (1%) 82 77	24, 39, 65, 102	0
1	B	385/424 (90%)	-0.44	4 (1%) 82 77	24, 42, 68, 108	0
All	All	772/848 (91%)	-0.46	8 (1%) 82 77	24, 40, 68, 108	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	131	SER	4.6
1	B	255	THR	3.8
1	B	254	GLY	3.2
1	B	253	PRO	3.0
1	A	253	PRO	2.5

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.