



Full wwPDB X-ray Structure Validation Report ⓘ

Jul 24, 2023 – 08:07 PM JST

PDB ID : 7YOD
Title : Crystal structure of A68P single mutant of O-acetylserine sulphydrylase from Haemophilus influenzae in complex with high-affinity inhibitory peptide from serine acetyltransferase of Salmonella typhimurium at 2.1 Å
Authors : Kumar, N.; Rahisuddin, R.; Saini, N.; Singh, A.K.
Deposited on : 2022-08-01
Resolution : 2.10 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 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.34
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) : 2.34

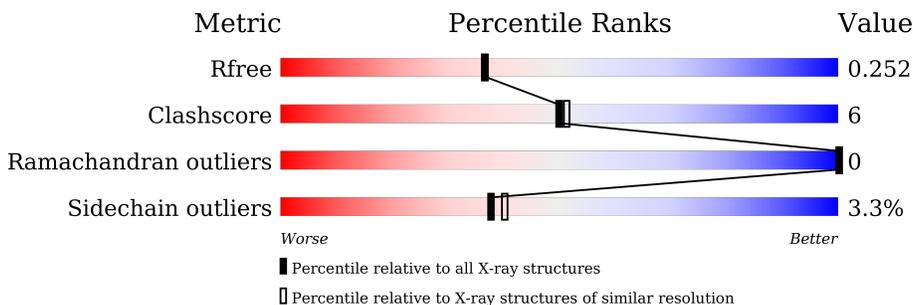
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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	316	 85% 10%
2	B	8	 12% 75% 12%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2344 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 Cysteine synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	303	2258	1420	391	437	1	9	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	68	PRO	ALA	engineered mutation	UNP P45040

- Molecule 2 is a protein called peptide from serine acetyltransferase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	7	49	31	7	11	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	37	Total	O	0	0
			37	37		

3 Residue-property plots [i](#)

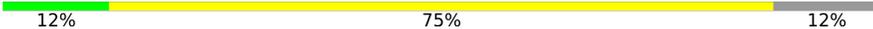
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: Cysteine synthase

Chain A:  85% 10%



- Molecule 2: peptide from serine acetyltransferase

Chain B:  12% 75% 12%



4 Data and refinement statistics i

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	112.60Å 112.60Å 45.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.81 – 2.10 39.81 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.81-2.10) 99.9 (39.81-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	9.63 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.222 , 0.255 0.221 , 0.252	Depositor DCC
R_{free} test set	824 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	32.5	Xtrriage
Anisotropy	0.035	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 36.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.278 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2344	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.41% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: LLP

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	0.40	0/2266	0.73	2/3069 (0.1%)
2	B	0.52	0/49	1.13	0/64
All	All	0.40	0/2315	0.74	2/3133 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	35	ARG	CG-CD-NE	-7.61	95.83	111.80
1	A	39	TYR	CB-CA-C	-5.11	100.18	110.40

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	2258	0	2320	24	0
2	B	49	0	37	9	0
3	A	37	0	0	5	0
All	All	2344	0	2357	27	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 6.

All (27) 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:58:THR:HB	3:A:405:HOH:O	1.63	0.97
1:A:120:MET:HE1	3:A:433:HOH:O	1.66	0.94
1:A:178:THR:CG2	2:B:401:ILE:HD11	2.28	0.64
1:A:92:MET:CE	1:A:96:MET:HG3	2.32	0.60
1:A:59:LEU:HD13	1:A:139:VAL:HG11	1.85	0.59
1:A:96:MET:CE	1:A:100:ARG:HG2	2.34	0.57
2:B:401:ILE:HD12	2:B:401:ILE:N	2.19	0.57
1:A:178:THR:HG23	2:B:401:ILE:HD11	1.87	0.56
1:A:58:THR:CB	3:A:405:HOH:O	2.38	0.52
1:A:92:MET:HE2	1:A:96:MET:HG3	1.93	0.51
1:A:98:LEU:O	1:A:102:ARG:HG3	2.10	0.51
1:A:143:GLN:OE1	2:B:401:ILE:HG12	2.10	0.50
1:A:61:LYS:HB2	3:A:425:HOH:O	2.12	0.50
1:A:92:MET:HE3	1:A:96:MET:HG3	1.94	0.48
1:A:269:ALA:HB1	1:A:273:SER:CB	2.43	0.48
1:A:178:THR:HG22	2:B:401:ILE:HD11	1.96	0.47
2:B:403:ASP:O	2:B:404:GLY:C	2.52	0.47
1:A:144:PHE:CZ	2:B:401:ILE:HD13	2.50	0.46
1:A:42:LLP:H2'1	1:A:72:ASN:ND2	2.32	0.45
1:A:38:SER:HB2	1:A:75:ILE:HG22	2.00	0.44
1:A:139:VAL:CG2	3:A:405:HOH:O	2.66	0.43
1:A:71:GLY:HA3	2:B:401:ILE:HG22	2.00	0.43
2:B:402:GLY:O	2:B:407:PHE:CD2	2.72	0.42
1:A:96:MET:HE3	1:A:100:ARG:HG2	2.01	0.42
1:A:102:ARG:HH11	1:A:102:ARG:HG2	1.86	0.41
1:A:172:VAL:HG11	1:A:185:ILE:HG21	2.03	0.40
1:A:153:HIS:HB2	1:A:184:GLY:HA3	2.04	0.40

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	300/316 (95%)	291 (97%)	9 (3%)	0	100	100
2	B	5/8 (62%)	1 (20%)	4 (80%)	0	100	100
All	All	305/324 (94%)	292 (96%)	13 (4%)	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	238/248 (96%)	231 (97%)	7 (3%)	42	46
2	B	4/6 (67%)	3 (75%)	1 (25%)	0	0
All	All	242/254 (95%)	234 (97%)	8 (3%)	38	40

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	67	ASP
1	A	69	THR
1	A	79	TYR
1	A	121	LYS
1	A	140	MET
1	A	227	GLN
2	B	406	GLU

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	227	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](#)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	LLP	A	42	1	23,24,25	0.57	0	25,32,34	0.57	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	A	42	1	-	1/16/17/19	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	42	LLP	CG-CD-CE-NZ

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	42	LLP	1	0

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](#)

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

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

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

6.4 Ligands [i](#)

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

6.5 Other polymers [i](#)

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