



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 11, 2024 – 12:22 PM JST

PDB ID : 8Y96
Title : Crystal structure of a heterooligomeric aminotransferase from *Serratia* sp. ATCC 39006
Authors : Pramono, H.; Yoshida, A.; Nishiyama, M.
Deposited on : 2024-02-06
Resolution : 2.84 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

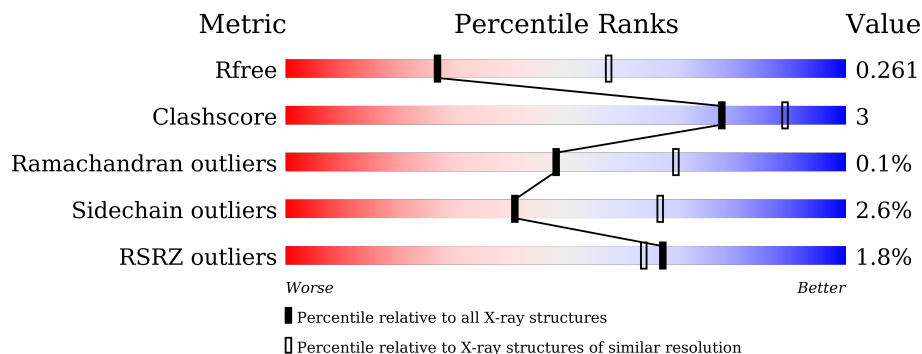
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.84 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}	164625	1367 (2.86-2.82)
Clashscore	180529	1455 (2.86-2.82)
Ramachandran outliers	177936	1422 (2.86-2.82)
Sidechain outliers	177891	1423 (2.86-2.82)
RSRZ outliers	164620	1368 (2.86-2.82)

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 $\leq 5\%$. 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	443	 86% 10% 4% 0% 0%
1	B	443	 88% 9% 3% 0% 0%
2	C	218	 83% 13% 3% 1% 0%
2	D	218	 79% 14% 4% 3% 0%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9730 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 DegT/DnrJ/EryC1/StrS family aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	428	3361	2157	572	615	17	0	0	0
1	B	427	3364	2157	575	615	17	0	1	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	438	HIS	-	expression tag	UNP A0A2I5TIB4
A	439	HIS	-	expression tag	UNP A0A2I5TIB4
A	440	HIS	-	expression tag	UNP A0A2I5TIB4
A	441	HIS	-	expression tag	UNP A0A2I5TIB4
A	442	HIS	-	expression tag	UNP A0A2I5TIB4
A	443	HIS	-	expression tag	UNP A0A2I5TIB4
B	438	HIS	-	expression tag	UNP A0A2I5TIB4
B	439	HIS	-	expression tag	UNP A0A2I5TIB4
B	440	HIS	-	expression tag	UNP A0A2I5TIB4
B	441	HIS	-	expression tag	UNP A0A2I5TIB4
B	442	HIS	-	expression tag	UNP A0A2I5TIB4
B	443	HIS	-	expression tag	UNP A0A2I5TIB4

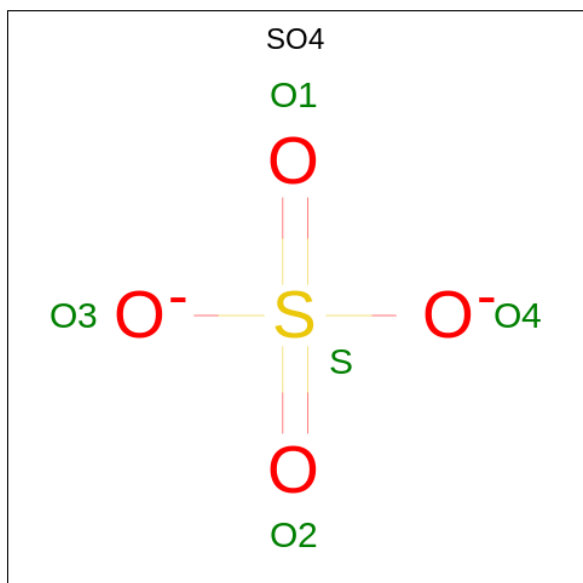
- Molecule 2 is a protein called DegT/DnrJ/EryC1/StrS aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	189	1470	950	244	271	5	0	0	0
2	D	188	1441	931	239	266	5	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP A0A2I5T5Y7
C	1	GLY	-	expression tag	UNP A0A2I5T5Y7
C	212	HIS	-	expression tag	UNP A0A2I5T5Y7
C	213	HIS	-	expression tag	UNP A0A2I5T5Y7
C	214	HIS	-	expression tag	UNP A0A2I5T5Y7
C	215	HIS	-	expression tag	UNP A0A2I5T5Y7
C	216	HIS	-	expression tag	UNP A0A2I5T5Y7
C	217	HIS	-	expression tag	UNP A0A2I5T5Y7
D	0	MET	-	initiating methionine	UNP A0A2I5T5Y7
D	1	GLY	-	expression tag	UNP A0A2I5T5Y7
D	212	HIS	-	expression tag	UNP A0A2I5T5Y7
D	213	HIS	-	expression tag	UNP A0A2I5T5Y7
D	214	HIS	-	expression tag	UNP A0A2I5T5Y7
D	215	HIS	-	expression tag	UNP A0A2I5T5Y7
D	216	HIS	-	expression tag	UNP A0A2I5T5Y7
D	217	HIS	-	expression tag	UNP A0A2I5T5Y7

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	23	Total 23	O 23	0	0
4	C	6	Total 6	O 6	0	0
4	B	43	Total 43	O 43	0	0
4	D	12	Total 12	O 12	0	0

4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	52.69Å 181.03Å 182.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.74 – 2.84 45.74 – 2.84	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.74-2.84) 99.9 (45.74-2.84)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.58 (at 2.86Å)	Xtrriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.204 , 0.259 0.207 , 0.261	Depositor DCC
R_{free} test set	2100 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	56.9	Xtrriage
Anisotropy	0.058	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 32.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.002 for -h,l,k	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9730	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.48% 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: SO4

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	0/3441	0.58	0/4662
1	B	0.30	0/3443	0.59	0/4662
2	C	0.30	0/1497	0.53	0/2023
2	D	0.30	0/1467	0.53	0/1985
All	All	0.30	0/9848	0.57	0/13332

There are no bond length outliers.

There are no bond angle outliers.

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	3361	0	3336	24	0
1	B	3364	0	3347	21	0
2	C	1470	0	1488	2	0
2	D	1441	0	1439	5	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	23	0	0	0	0
4	B	43	0	0	0	0
4	C	6	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	12	0	0	0	0
All	All	9730	0	9610	52	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 3.

The worst 5 of 52 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:300:MET:HE3	1:B:388:CYS:SG	2.39	0.63
1:B:21:GLN:HE21	1:B:22:LEU:H	1.49	0.60
1:B:397:TYR:CZ	1:B:401[A]:HIS:CE1	2.90	0.59
1:B:73:ARG:HD3	1:B:258:LEU:HD11	1.87	0.57
1:B:300:MET:CE	1:B:388:CYS:SG	2.92	0.56

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	426/443 (96%)	416 (98%)	10 (2%)	0	100	100
1	B	426/443 (96%)	413 (97%)	13 (3%)	0	100	100
2	C	185/218 (85%)	174 (94%)	11 (6%)	0	100	100
2	D	184/218 (84%)	177 (96%)	6 (3%)	1 (0%)	25	44
All	All	1221/1322 (92%)	1180 (97%)	40 (3%)	1 (0%)	48	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	165	GLY

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	351/369 (95%)	344 (98%)	7 (2%)	50 74
1	B	352/369 (95%)	344 (98%)	8 (2%)	45 70
2	C	160/192 (83%)	156 (98%)	4 (2%)	42 68
2	D	153/192 (80%)	146 (95%)	7 (5%)	23 45
All	All	1016/1122 (91%)	990 (97%)	26 (3%)	41 67

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	141	GLN
1	B	363	LEU
2	D	168	ASP
1	B	362	PHE
1	B	367	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	174	GLN
1	B	299	ASN
2	D	150	ASN
2	D	12	GLN
1	B	21	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

2 ligands are 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	$\# Z > 2$
3	SO4	B	501	-	4,4,4	0.36	0	6,6,6	0.09	0
3	SO4	A	501	-	4,4,4	0.40	0	6,6,6	0.07	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	428/443 (96%)	-0.15	6 (1%) 73 69	38, 56, 87, 112	0
1	B	427/443 (96%)	-0.41	1 (0%) 92 93	21, 50, 72, 96	1 (0%)
2	C	189/218 (86%)	-0.08	6 (3%) 50 45	41, 58, 88, 130	0
2	D	188/218 (86%)	0.24	9 (4%) 36 32	40, 65, 91, 103	0
All	All	1232/1322 (93%)	-0.17	22 (1%) 67 64	21, 55, 86, 130	1 (0%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	141	GLN	5.3
2	C	169	LEU	5.1
1	A	428	PRO	4.0
2	D	181	GLU	3.5
2	C	167	PHE	3.4

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	SO4	B	501	5/5	0.98	0.07	44,45,46,46	0
3	SO4	A	501	5/5	0.99	0.05	47,48,50,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.