



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 1, 2022 – 02:54 pm GMT

PDB ID : 8BGQ
Title : Crystal structure of human Thiosulfate sulfurtransferase amino acid 1-291 in high resolution
Authors : Goo, H.v.; Groves, M.R.; AL-Dahmani, Z.M.; Wang, C.
Deposited on : 2022-10-28
Resolution : 1.28 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
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

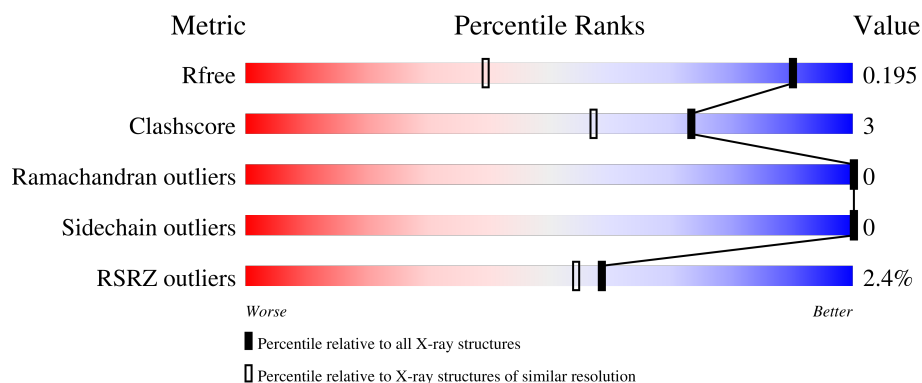
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

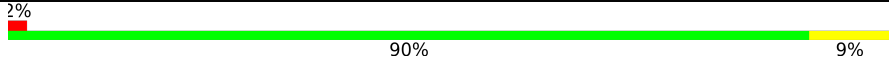
The reported resolution of this entry is 1.28 Å.

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	1850 (1.30-1.26)
Clashscore	141614	1926 (1.30-1.26)
Ramachandran outliers	138981	1860 (1.30-1.26)
Sidechain outliers	138945	1859 (1.30-1.26)
RSRZ outliers	127900	1807 (1.30-1.26)

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	291	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4888 atoms, of which 2246 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 Thiosulfate sulfurtransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	H	N	O	S	71	0	0
			4561	1480	2246	405	421	9			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	PHE	SER	conflict	UNP Q16762
A	281	HIS	ARG	conflict	UNP Q16762
A	288	TRP	ARG	conflict	UNP Q16762

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	327	Total	O	0	0
			327	327		

- Molecule 1: Thiosulfate sulfurtransferase



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	41.86Å 49.51Å 153.36Å 90.00° 97.16° 90.00°	Depositor
Resolution (Å)	38.89 – 1.28 38.86 – 1.28	Depositor EDS
% Data completeness (in resolution range)	74.9 (38.89-1.28) 73.6 (38.86-1.28)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.64 (at 1.28Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC 5.8.0267	Depositor
R, R_{free}	0.164 , 0.188 0.172 , 0.195	Depositor DCC
R_{free} test set	2990 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	9.2	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4888	wwPDB-VP
Average B, all atoms (Å ²)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.49% 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	0.90	4/2383 (0.2%)	0.96	9/3235 (0.3%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	102	GLU	CD-OE1	-14.21	1.10	1.25
1	A	166	GLU	CD-OE2	10.33	1.37	1.25
1	A	102	GLU	CB-CG	8.11	1.67	1.52
1	A	232	LEU	CG-CD1	5.26	1.71	1.51

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	123	THR	OG1-CB-CG2	-7.39	93.00	110.00
1	A	230	ARG	NE-CZ-NH2	-6.77	116.91	120.30
1	A	230	ARG	NE-CZ-NH1	6.45	123.53	120.30
1	A	183	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	A	132	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	A	30	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	A	117	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	A	111	ARG	NE-CZ-NH1	-5.28	117.66	120.30
1	A	249	ARG	NE-CZ-NH2	5.03	122.81	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	2315	2246	2229	15	1
2	A	327	0	0	2	0
All	All	2642	2246	2229	15	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 3.

All (15) 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:92:ASN:HD21	1:A:154:LYS:H	1.15	0.93
1:A:171:ASN:HD21	1:A:177:PHE:H	1.17	0.89
1:A:164:THR:H	1:A:167:GLN:HE21	1.32	0.76
1:A:87:ARG:NH2	2:A:301:HOH:O	2.28	0.66
1:A:102:GLU:HG3	1:A:105:GLY:H	1.60	0.66
1:A:277:PHE:O	1:A:281:HIS:HD2	1.79	0.64
1:A:107:PHE:O	1:A:256:HIS:HE1	1.81	0.63
1:A:281:HIS:HE1	2:A:356:HOH:O	1.83	0.61
1:A:48:TYR:O	1:A:52:HIS:HD2	1.85	0.58
1:A:102:GLU:HG3	1:A:104:LEU:H	1.76	0.51
1:A:102:GLU:HG3	1:A:104:LEU:N	2.27	0.50
1:A:171:ASN:ND2	1:A:177:PHE:H	1.96	0.49
1:A:102:GLU:CG	1:A:105:GLY:H	2.26	0.48
1:A:92:ASN:HD21	1:A:154:LYS:N	1.97	0.48
1:A:102:GLU:HG3	1:A:105:GLY:N	2.31	0.44

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:A:103:HIS:HD2	1:A:147:ARG:HH22[4_444]	1.22	0.38

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	289/291 (99%)	281 (97%)	8 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	242/246 (98%)	242 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	HIS
1	A	92	ASN
1	A	133	ASN
1	A	139	HIS
1	A	167	GLN
1	A	171	ASN
1	A	256	HIS
1	A	281	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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	291/291 (100%)	-0.19	7 (2%) 59 54	6, 10, 18, 26	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	87	ARG	3.6
1	A	93	HIS	3.0
1	A	235	THR	2.8
1	A	73	MET	2.7
1	A	176	ARG	2.3
1	A	232	LEU	2.2
1	A	67	THR	2.1

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.