



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 20, 2023 – 02:36 PM JST

PDB ID : 7CQY  
Title : Tetrathionate hydrolase from Acidithiobacillus ferrooxidans mutant - D325N  
Authors : Tamada, T.; Hirano, Y.  
Deposited on : 2020-08-12  
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](mailto: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 : 2.36  
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) : 2.36

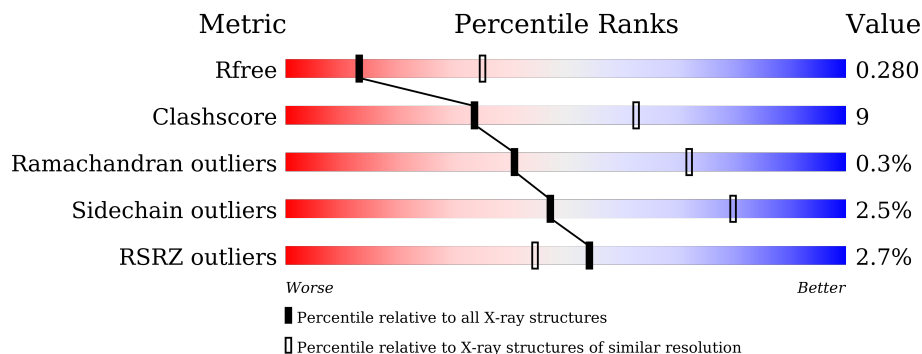
# 1 Overall quality at a glance

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	470	 75% 15% • 10%
1	B	470	 74% 18% • 7%
1	C	470	 65% 21% • 13%
1	D	470	 75% 22% ••
1	E	470	 77% 18% ••
1	F	470	 76% 17% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
2	SO4	E	501	-	-	X	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 19937 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 Tetrathionate hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	425	Total 3202	C 2058	N 541	O 591	S 12	0	0	0
1	B	436	Total 3302	C 2120	N 562	O 609	S 11	0	0	0
1	C	409	Total 3105	C 1997	N 524	O 574	S 10	0	0	0
1	D	464	Total 3502	C 2248	N 600	O 642	S 12	0	0	0
1	E	451	Total 3406	C 2185	N 582	O 627	S 12	0	0	0
1	F	445	Total 3365	C 2161	N 575	O 618	S 11	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

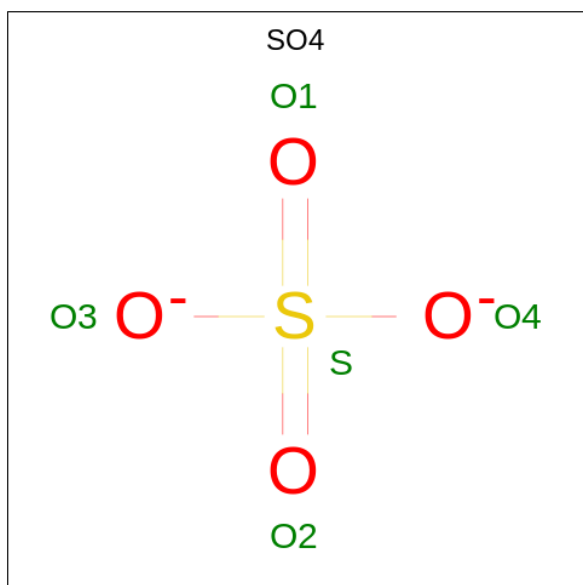
Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	expression tag	UNP B7J3C9
A	31	SER	-	expression tag	UNP B7J3C9
A	32	ALA	-	expression tag	UNP B7J3C9
A	325	ASN	ASP	engineered mutation	UNP B7J3C9
B	30	MET	-	expression tag	UNP B7J3C9
B	31	SER	-	expression tag	UNP B7J3C9
B	32	ALA	-	expression tag	UNP B7J3C9
B	325	ASN	ASP	engineered mutation	UNP B7J3C9
C	30	MET	-	expression tag	UNP B7J3C9
C	31	SER	-	expression tag	UNP B7J3C9
C	32	ALA	-	expression tag	UNP B7J3C9
C	325	ASN	ASP	engineered mutation	UNP B7J3C9
D	30	MET	-	expression tag	UNP B7J3C9
D	31	SER	-	expression tag	UNP B7J3C9
D	32	ALA	-	expression tag	UNP B7J3C9
D	325	ASN	ASP	engineered mutation	UNP B7J3C9
E	30	MET	-	expression tag	UNP B7J3C9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	31	SER	-	expression tag	UNP B7J3C9
E	32	ALA	-	expression tag	UNP B7J3C9
E	325	ASN	ASP	engineered mutation	UNP B7J3C9
F	30	MET	-	expression tag	UNP B7J3C9
F	31	SER	-	expression tag	UNP B7J3C9
F	32	ALA	-	expression tag	UNP B7J3C9
F	325	ASN	ASP	engineered mutation	UNP B7J3C9

- Molecule 2 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	O	0	0
			2	2		
3	B	2	Total	O	0	0
			2	2		
3	D	4	Total	O	0	0
			4	4		
3	E	1	Total	O	0	0
			1	1		
3	F	1	Total	O	0	0
			1	1		



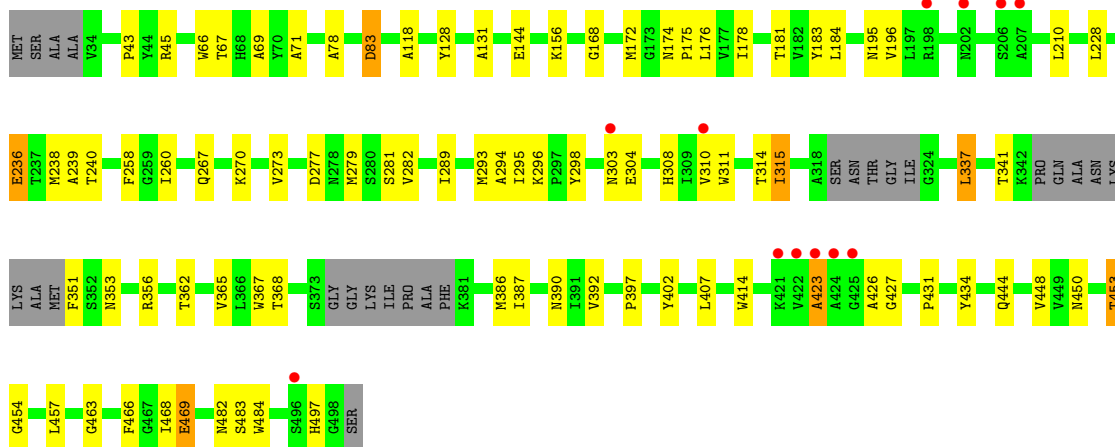




S496  
H497  
GLY  
SER

• Molecule 1: Tetrathionate hydrolase

Chain F: 3% 76% 17% 5%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.31Å 94.31Å 235.57Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.24 – 2.80 46.24 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.6 (46.24-2.80) 99.6 (46.24-2.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.03 (at 2.81Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.225 , 0.280 0.226 , 0.280	Depositor DCC
$R_{free}$ test set	2770 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	62.2	Xtrriage
Anisotropy	0.580	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l 0.034 for h,-h-k,-l 0.025 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	19937	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.28	0/3298	0.53	0/4514
1	B	0.27	0/3403	0.52	0/4664
1	C	0.27	0/3198	0.52	0/4381
1	D	0.27	0/3611	0.50	0/4945
1	E	0.27	0/3510	0.52	0/4806
1	F	0.28	0/3468	0.52	1/4748 (0.0%)
All	All	0.27	0/20488	0.52	1/28058 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	337	LEU	CA-CB-CG	5.66	128.31	115.30

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	3202	0	3092	44	0
1	B	3302	0	3180	57	0
1	C	3105	0	2993	64	0
1	D	3502	0	3397	71	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	3406	0	3290	55	0
1	F	3365	0	3248	57	0
2	A	5	0	0	0	0
2	B	10	0	0	1	0
2	D	5	0	0	0	0
2	E	15	0	0	2	0
2	F	10	0	0	1	0
3	A	2	0	0	1	0
3	B	2	0	0	2	0
3	D	4	0	0	4	0
3	E	1	0	0	0	0
3	F	1	0	0	2	0
All	All	19937	0	19200	334	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 9.

The worst 5 of 334 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:450:ASN:ND2	1:B:453:THR:OG1	2.16	0.78
1:D:172:MET:SD	3:D:604:HOH:O	2.42	0.76
1:D:132:LEU:HD22	1:D:173:GLY:HA2	1.70	0.73
1:F:78:ALA:HB1	1:F:434:TYR:HD1	1.54	0.72
1:B:172:MET:SD	3:B:602:HOH:O	2.47	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	415/470 (88%)	402 (97%)	12 (3%)	1 (0%)	47 78
1	B	428/470 (91%)	413 (96%)	14 (3%)	1 (0%)	47 78
1	C	397/470 (84%)	380 (96%)	16 (4%)	1 (0%)	41 72
1	D	462/470 (98%)	445 (96%)	16 (4%)	1 (0%)	47 78
1	E	445/470 (95%)	429 (96%)	15 (3%)	1 (0%)	47 78
1	F	437/470 (93%)	417 (95%)	18 (4%)	2 (0%)	29 61
All	All	2584/2820 (92%)	2486 (96%)	91 (4%)	7 (0%)	41 72

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	295	ILE
1	B	295	ILE
1	C	295	ILE
1	D	295	ILE
1	E	295	ILE

### 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	328/361 (91%)	319 (97%)	9 (3%)	44 78
1	B	339/361 (94%)	331 (98%)	8 (2%)	49 81
1	C	320/361 (89%)	312 (98%)	8 (2%)	47 80
1	D	358/361 (99%)	350 (98%)	8 (2%)	52 83
1	E	349/361 (97%)	340 (97%)	9 (3%)	46 79
1	F	344/361 (95%)	335 (97%)	9 (3%)	46 79
All	All	2038/2166 (94%)	1987 (98%)	51 (2%)	47 80

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	327	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	198	ARG
1	F	469	GLU
1	D	377	ILE
1	D	484	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	211	ASN
1	F	353	ASN
1	F	390	ASN
1	F	388	HIS
1	C	388	HIS

### 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](#)

9 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
2	SO4	E	501	-	4,4,4	0.14	0	6,6,6	0.12	0
2	SO4	F	501	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	E	502	-	4,4,4	0.15	0	6,6,6	0.18	0
2	SO4	E	503	-	4,4,4	0.14	0	6,6,6	0.09	0
2	SO4	A	501	-	4,4,4	0.15	0	6,6,6	0.09	0
2	SO4	B	502	-	4,4,4	0.14	0	6,6,6	0.10	0
2	SO4	D	501	-	4,4,4	0.15	0	6,6,6	0.10	0
2	SO4	F	502	-	4,4,4	0.13	0	6,6,6	0.10	0
2	SO4	B	501	-	4,4,4	0.14	0	6,6,6	0.09	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.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	501	SO4	2	0
2	F	501	SO4	1	0
2	B	502	SO4	1	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	425/470 (90%)	0.14	6 (1%) 75 70	43, 60, 80, 100	0
1	B	436/470 (92%)	0.16	8 (1%) 68 61	43, 63, 82, 92	0
1	C	409/470 (87%)	0.52	33 (8%) 12 6	57, 75, 93, 100	0
1	D	464/470 (98%)	0.09	8 (1%) 70 63	47, 61, 74, 86	0
1	E	451/470 (95%)	-0.04	5 (1%) 80 75	38, 53, 77, 89	0
1	F	445/470 (94%)	0.11	12 (2%) 54 44	41, 59, 79, 89	0
All	All	2630/2820 (93%)	0.16	72 (2%) 54 44	38, 62, 85, 100	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	419	PRO	4.8
1	D	55	SER	4.7
1	C	114	THR	4.4
1	D	380	PHE	4.3
1	C	238	MET	3.9

### 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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
2	SO4	B	502	5/5	0.91	0.14	81,83,86,102	0
2	SO4	E	503	5/5	0.92	0.20	71,75,83,89	0
2	SO4	E	502	5/5	0.93	0.19	50,55,63,70	0
2	SO4	E	501	5/5	0.93	0.16	57,65,73,90	0
2	SO4	A	501	5/5	0.94	0.16	70,74,87,92	0
2	SO4	D	501	5/5	0.95	0.14	67,69,78,86	0
2	SO4	F	501	5/5	0.95	0.10	54,56,65,70	0
2	SO4	B	501	5/5	0.96	0.11	54,60,64,72	0
2	SO4	F	502	5/5	0.97	0.18	62,65,72,78	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.