



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 09:30 PM UTC

PDB ID : 3NIV / pdb_00003niv
Title : The crystal structure of Glutathione S-transferase from Legionella pneumophila
Authors : Zhang, Z.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2010-06-16
Resolution : 2.30 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

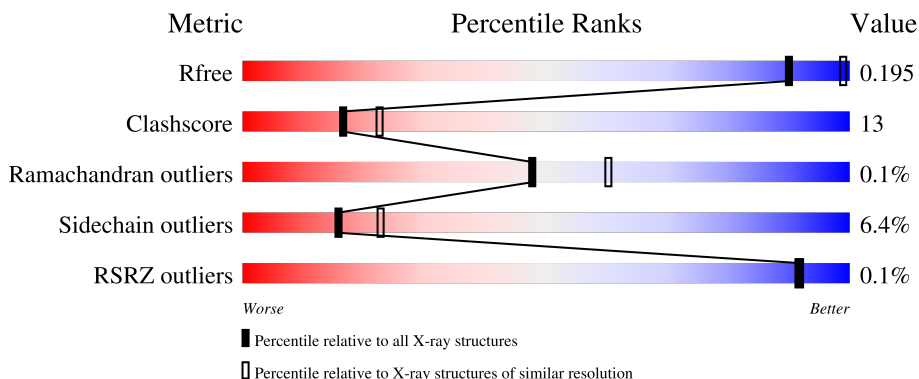
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

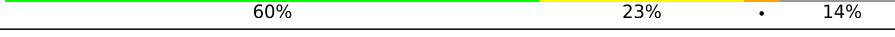
The reported resolution of this entry is 2.30 Å.

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}	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	222	 65% 19% • 14%
1	B	222	 57% 26% • 15%
1	C	222	 58% 23% 5% 14%
1	D	222	 60% 23% • 14%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6308 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 Glutathione S-transferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	191	Total	C	N	O	S	Se	0	1	0
			1552	1006	256	279	5	6			
1	B	188	Total	C	N	O	S	Se	0	0	0
			1525	989	250	275	5	6			
1	C	191	Total	C	N	O	S	Se	0	0	0
			1550	1004	256	279	5	6			
1	D	192	Total	C	N	O	S	Se	0	1	0
			1552	1006	256	279	5	6			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	expression tag	UNP Q5ZT82
A	0	SER	-	expression tag	UNP Q5ZT82
A	1	LEU	-	expression tag	UNP Q5ZT82
A	213	GLU	-	expression tag	UNP Q5ZT82
A	214	GLY	-	expression tag	UNP Q5ZT82
A	215	HIS	-	expression tag	UNP Q5ZT82
A	216	HIS	-	expression tag	UNP Q5ZT82
A	217	HIS	-	expression tag	UNP Q5ZT82
A	218	HIS	-	expression tag	UNP Q5ZT82
A	219	HIS	-	expression tag	UNP Q5ZT82
A	220	HIS	-	expression tag	UNP Q5ZT82
B	-1	MSE	-	expression tag	UNP Q5ZT82
B	0	SER	-	expression tag	UNP Q5ZT82
B	1	LEU	-	expression tag	UNP Q5ZT82
B	213	GLU	-	expression tag	UNP Q5ZT82
B	214	GLY	-	expression tag	UNP Q5ZT82
B	215	HIS	-	expression tag	UNP Q5ZT82
B	216	HIS	-	expression tag	UNP Q5ZT82
B	217	HIS	-	expression tag	UNP Q5ZT82
B	218	HIS	-	expression tag	UNP Q5ZT82
B	219	HIS	-	expression tag	UNP Q5ZT82

Continued on next page...

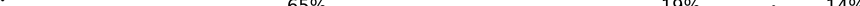
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	220	HIS	-	expression tag	UNP Q5ZT82
C	-1	MSE	-	expression tag	UNP Q5ZT82
C	0	SER	-	expression tag	UNP Q5ZT82
C	1	LEU	-	expression tag	UNP Q5ZT82
C	213	GLU	-	expression tag	UNP Q5ZT82
C	214	GLY	-	expression tag	UNP Q5ZT82
C	215	HIS	-	expression tag	UNP Q5ZT82
C	216	HIS	-	expression tag	UNP Q5ZT82
C	217	HIS	-	expression tag	UNP Q5ZT82
C	218	HIS	-	expression tag	UNP Q5ZT82
C	219	HIS	-	expression tag	UNP Q5ZT82
C	220	HIS	-	expression tag	UNP Q5ZT82
D	-1	MSE	-	expression tag	UNP Q5ZT82
D	0	SER	-	expression tag	UNP Q5ZT82
D	1	LEU	-	expression tag	UNP Q5ZT82
D	213	GLU	-	expression tag	UNP Q5ZT82
D	214	GLY	-	expression tag	UNP Q5ZT82
D	215	HIS	-	expression tag	UNP Q5ZT82
D	216	HIS	-	expression tag	UNP Q5ZT82
D	217	HIS	-	expression tag	UNP Q5ZT82
D	218	HIS	-	expression tag	UNP Q5ZT82
D	219	HIS	-	expression tag	UNP Q5ZT82
D	220	HIS	-	expression tag	UNP Q5ZT82

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	23	Total O 23 23	0	0
2	B	33	Total O 33 33	0	0
2	C	44	Total O 44 44	0	0
2	D	29	Total O 29 29	0	0

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

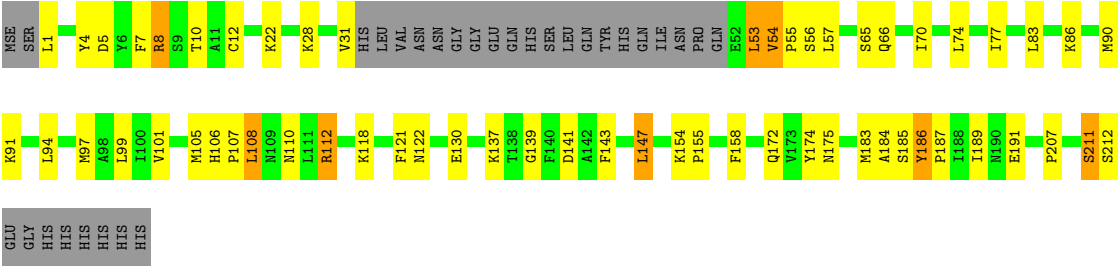
- Chain A:  65% 19% 14%
- | | |
|------|------|
| NSE | N110 |
| SER | L111 |
| LEU | R112 |
| I2 | V113 |
| Y6 | K118 |
| F7 | N122 |
| H8 | A123 |
| Y13 | N124 |
| I17 | E125 |
| L21 | E126 |
| K22 | Q127 |
| F23 | Y132 |
| L24 | L136 |
| L33 | G139 |
| VAL | F140 |
| ASN | F143 |
| ASN | E144 |
| GLY | E145 |
| GLY | K146 |
| GLU | L147 |
| GLN | K154 |
| HIS | P155 |
| SER | P156 |
| LEU | E161 |
| GLN | V162 |
| TYR | G163 |
| HIS | V167 |
| GLN | Q172 |
| ILE | I189 |
| ASN | Y195 |
| PRO | C196 |
| GLN | L197 |
| GLU | A206 |
| L53 | L210 |
| I63 | S211 |
| L64 | GLU |
| S65 | GLY |
| Y73 | HIS |
| I77 | HIS |
| H78 | HIS |
| P79 | HIS |
| K95 | HIS |
| S96 | HIS |
| I100 | HIS |
| V101 | HIS |
| M105 | HIS |
| L108 | HIS |
| H109 | HIS |

- Chain B:
-
- 57% 26% 15%
- MSE SER L1 I2 D5 Y6 F7 T10 A11 C12 Y13 R14 I17 K22 A25 V31 HIS LEU VAL ASN GLY GLU GLN HIS SER LEU GLN TTR HIS ILE ASN PRO GLN LEU V54 P55 S56 L57 D58 S65 Q66 I70 Y73 L74 E75 F76 I77
- H78 H79 E80 E81 P82 L83 M97 L108 M109 M110 L111 R112 K118 E119 A120 F121 M122 A123 N124 V128 L129 Y132 T138 A142 F143 E144 E145 K146 L147 D153 K154 P155 S160 E161 V162 D166 I170 P171 M175 S185 Y186 P187 I188 N189 N190 E191
- Y196 A206 F207 E208 A209 I210 SER SER GLU GLY HIS HIS HIS HIS HIS HIS

- [illegible]

- 
- WORLD WIDE
PDB
PROTEIN DATA BANK

Chain D: 60% 23% • 14%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	51.48Å 55.04Å 84.72Å 71.99° 84.33° 68.79°	Depositor
Resolution (Å)	80.56 – 2.30 80.56 – 2.30	Depositor EDS
% Data completeness (in resolution range)	91.2 (80.56-2.30) 91.2 (80.56-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.25 (at 2.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.197 , 0.249 0.197 , 0.195	Depositor DCC
R_{free} test set	1678 reflections (4.58%)	wwPDB-VP
Wilson B-factor (Å ²)	32.6	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 35.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.012 for -h,-k,-k+l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6308	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% 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	0.48	0/1589	0.85	4/2149 (0.2%)
1	B	0.52	1/1557 (0.1%)	0.87	6/2103 (0.3%)
1	C	0.48	0/1583	0.86	6/2139 (0.3%)
1	D	0.58	0/1588	0.83	2/2146 (0.1%)
All	All	0.52	1/6317 (0.0%)	0.85	18/8537 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	170	ILE	CA-CB	6.00	1.57	1.54

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	186	TYR	CA-C-N	6.72	125.96	118.97
1	D	186	TYR	C-N-CA	6.72	125.96	118.97
1	C	106	HIS	CA-C-N	-6.22	113.22	119.56
1	C	106	HIS	C-N-CA	-6.22	113.22	119.56
1	B	65	SER	CA-C-N	-6.20	114.17	122.72

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	1552	0	1532	35	0
1	B	1525	0	1516	50	0
1	C	1550	0	1534	41	0
1	D	1552	0	1540	39	0
2	A	23	0	0	0	0
2	B	33	0	0	1	0
2	C	44	0	0	0	0
2	D	29	0	0	0	0
All	All	6308	0	6122	159	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 13.

The worst 5 of 159 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:D:118:LYS:O	1:D:122:ASN:HA	1.68	0.93
1:C:2:ILE:HG23	1:C:29:ILE:HD13	1.51	0.92
1:A:110:ASN:HD22	1:A:112:ARG:H	1.18	0.87
1:C:97:MSE:HE1	1:C:147:LEU:HD13	1.63	0.80
1:A:132:TYR:O	1:A:136:LEU:HD23	1.83	0.78

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	188/222 (85%)	179 (95%)	9 (5%)	0	100	100
1	B	184/222 (83%)	180 (98%)	4 (2%)	0	100	100
1	C	187/222 (84%)	184 (98%)	2 (1%)	1 (0%)	24	31
1	D	189/222 (85%)	183 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	748/888 (84%)	726 (97%)	21 (3%)	1 (0%)	48 60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	66	GLN

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	169/190 (89%)	159 (94%)	10 (6%)	18 26
1	B	166/190 (87%)	161 (97%)	5 (3%)	36 53
1	C	169/190 (89%)	154 (91%)	15 (9%)	9 12
1	D	169/190 (89%)	156 (92%)	13 (8%)	12 16
All	All	673/760 (89%)	630 (94%)	43 (6%)	16 23

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	160	SER
1	D	54	VAL
1	C	170	ILE
1	D	28	LYS
1	D	91	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	172	GLN
1	D	134	HIS
1	C	172	GLN
1	D	110	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	110	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 oligosaccharides 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	185/222 (83%)	-0.15	0 100 100	24, 35, 49, 61	1 (0%)
1	B	182/222 (81%)	-0.19	1 (0%) 87 87	22, 33, 48, 63	0
1	C	185/222 (83%)	-0.18	0 100 100	24, 33, 47, 55	0
1	D	186/222 (83%)	-0.24	0 100 100	20, 33, 46, 57	1 (0%)
All	All	738/888 (83%)	-0.19	1 (0%) 92 91	20, 33, 48, 63	2 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	210	ILE	2.0

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 oligosaccharides 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.