



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 19, 2022 – 06:55 AM JST

PDB ID : 7ESG
Title : Crystal structure of Haloarcula marismortui CheB with Glutathione S-transferase expression tag
Authors : Chen, K.L.; Yang, C.S.
Deposited on : 2021-05-10
Resolution : 2.53 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
Xtriage (Phenix) : 1.13
EDS : 2.29
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) : 2.29

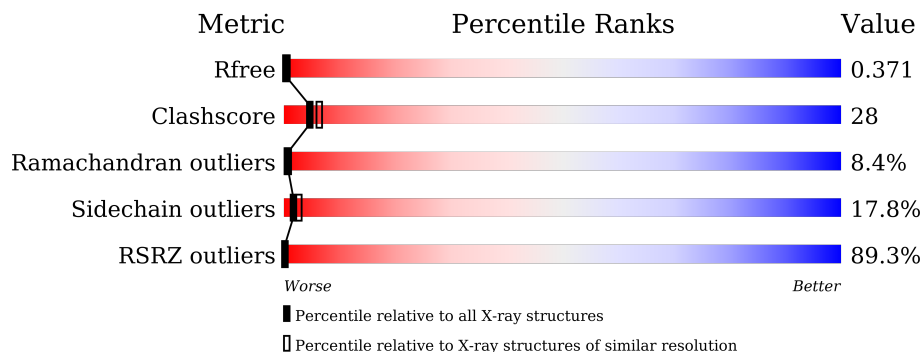
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.53 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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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	598	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4553 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,Protein-glutamate methylesterase/protein-glutamine glutaminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	589	4514	2822	767	895	30	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	219	SER	-	linker	UNP P08515
A	220	ASP	-	linker	UNP P08515
A	221	LEU	-	linker	UNP P08515
A	222	VAL	-	linker	UNP P08515
A	223	PRO	-	linker	UNP P08515
A	224	ARG	-	linker	UNP P08515
A	225	GLY	-	linker	UNP P08515
A	226	SER	-	linker	UNP P08515
A	227	PRO	-	linker	UNP P08515
A	228	GLU	-	linker	UNP P08515
A	229	PHE	-	linker	UNP P08515

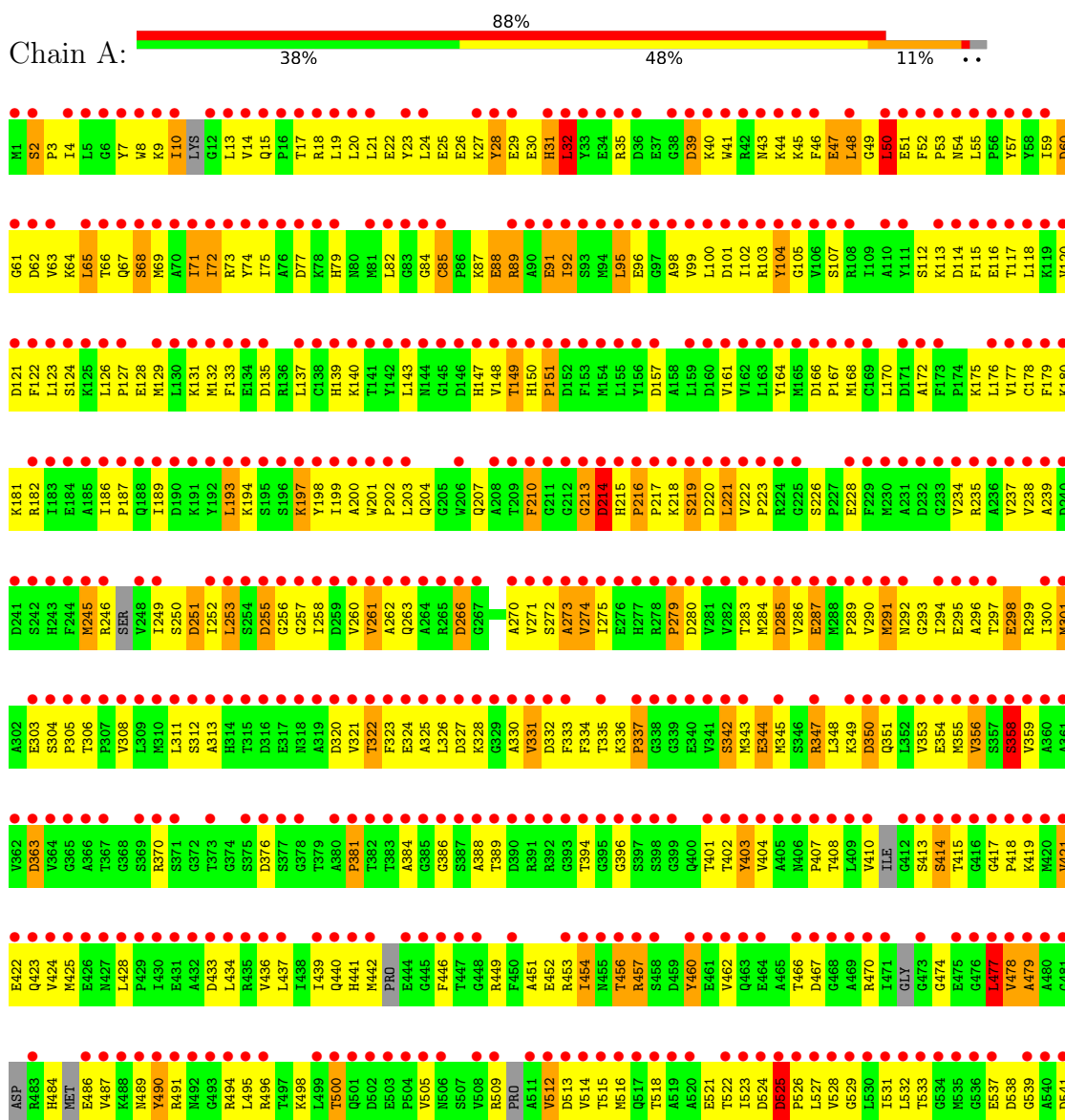
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	39	Total	O	0	0
			39	39		

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 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.

- Molecule 1: Glutathione S-transferase,Protein-glutamate methylesterase/protein-glutamine glutaminase



G542	I543	R544	R545	I546	K547	Q548	R549	G550	G551	K552	T553	T554	ALA	Q556	D557	E558	A559	T560	S561	A562	V563	Y564	G565	M566	P567	R568	R569	A570	A571	E572	T573	G574	C575	V576	D577	T578	V579	L580	P581	I582	D583	D584	I585	A586	D587	G588	V589	I590	R591	T592	I593	T594	T595	E596	V597	T598
------	------	------	------	------	------	------	------	------	------	------	------	------	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.06Å 93.91Å 93.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.86 – 2.53 23.86 – 2.53	Depositor EDS
% Data completeness (in resolution range)	99.5 (23.86-2.53) 98.9 (23.86-2.53)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	120.76 (at 2.53Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.349 , 0.372 0.346 , 0.371	Depositor DCC
R_{free} test set	1753 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å ²)	35.4	Xtrriage
Anisotropy	0.457	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.14 , 19.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.318 for -h,l,k	Xtrriage
F_o, F_c correlation	0.75	EDS
Total number of atoms	4553	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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	0.26	0/4585	0.54	0/6191

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	4514	0	4445	253	0
2	A	39	0	0	11	0
All	All	4553	0	4445	253	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 28.

The worst 5 of 253 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:295:GLU:OE1	2:A:601:HOH:O	1.85	0.94
1:A:140:LYS:HA	1:A:143:LEU:HB2	1.51	0.92
1:A:161:VAL:O	2:A:602:HOH:O	1.91	0.89
1:A:137:LEU:HB3	1:A:151:PRO:HB3	1.58	0.86
1:A:246:ARG:H	1:A:249:ILE:HB	1.41	0.85

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	569/598 (95%)	399 (70%)	122 (21%)	48 (8%)	1 0

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	LYS
1	A	149	THR
1	A	216	PRO
1	A	358	SER
1	A	381	PRO

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	488/495 (99%)	401 (82%)	87 (18%)	2 3

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	345	MET
1	A	454	ILE
1	A	350	ASP
1	A	401	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	477	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

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	589/598 (98%)	4.32	526 (89%) 0 0	23, 27, 31, 35	0

The worst 5 of 526 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	341	VAL	15.0
1	A	286	VAL	14.5
1	A	523	ILE	14.3
1	A	258	ILE	13.6
1	A	275	ILE	13.6

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.