wwPDB X-ray Structure Validation Summary Report (i)

Sep 12, 2023 – 09:05 PM EDT

PDB ID	:	4N0V
Title	:	Crystal structure of a glutathione S-transferase domain-containing protein (Marinobacter aquaeolei VT8), Target EFI-507332
Authors	:	Kim, J.; Toro, R.; Bhosle, R.; Al Obaidi, N.F.; Morisco, L.L.; Wasserman, S.R.; Sojitra, S.; Washington, E.; Glenn, A.S.; Chowdhury, S.; Evans, B.; Zhao, S.C.; Hillerich, B.; Love, J.; Seidel, R.D.; Imker, H.J.; Stead, M.; Jacobson, M.P.; Gerlt, J.A.; Almo, S.C.; Enzyme Function Initiative (EFI)
Deposited on	:	2013-10-02
Resolution	:	1.70 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 (i) 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 (i)) were used in the production of this report:

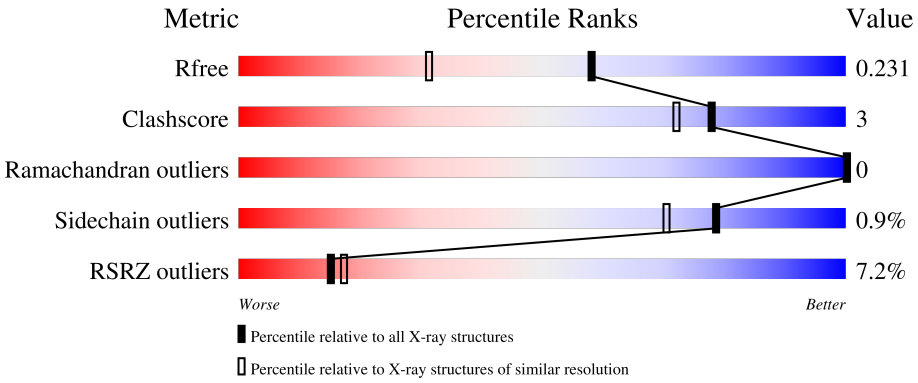
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 <=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	224	4% 89% 5% 5%
1	B	224	9% 87% · 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3408 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, N-terminal domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1629	1033	279	308	9			
1	B	205	Total	C	N	O	S	0	0	0
			1561	989	268	295	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	ALA	-	expression tag	UNP A1TY67
A	208	GLU	-	expression tag	UNP A1TY67
A	209	ASN	-	expression tag	UNP A1TY67
A	210	LEU	-	expression tag	UNP A1TY67
A	211	TYR	-	expression tag	UNP A1TY67
A	212	PHE	-	expression tag	UNP A1TY67
A	213	GLN	-	expression tag	UNP A1TY67
A	214	GLY	-	expression tag	UNP A1TY67
A	215	HIS	-	expression tag	UNP A1TY67
A	216	HIS	-	expression tag	UNP A1TY67
A	217	HIS	-	expression tag	UNP A1TY67
A	218	HIS	-	expression tag	UNP A1TY67
A	219	HIS	-	expression tag	UNP A1TY67
A	220	HIS	-	expression tag	UNP A1TY67
A	221	HIS	-	expression tag	UNP A1TY67
A	222	HIS	-	expression tag	UNP A1TY67
A	223	HIS	-	expression tag	UNP A1TY67
A	224	HIS	-	expression tag	UNP A1TY67
B	207	ALA	-	expression tag	UNP A1TY67
B	208	GLU	-	expression tag	UNP A1TY67
B	209	ASN	-	expression tag	UNP A1TY67
B	210	LEU	-	expression tag	UNP A1TY67
B	211	TYR	-	expression tag	UNP A1TY67
B	212	PHE	-	expression tag	UNP A1TY67
B	213	GLN	-	expression tag	UNP A1TY67

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	214	GLY	-	expression tag	UNP A1TY67
B	215	HIS	-	expression tag	UNP A1TY67
B	216	HIS	-	expression tag	UNP A1TY67
B	217	HIS	-	expression tag	UNP A1TY67
B	218	HIS	-	expression tag	UNP A1TY67
B	219	HIS	-	expression tag	UNP A1TY67
B	220	HIS	-	expression tag	UNP A1TY67
B	221	HIS	-	expression tag	UNP A1TY67
B	222	HIS	-	expression tag	UNP A1TY67
B	223	HIS	-	expression tag	UNP A1TY67
B	224	HIS	-	expression tag	UNP A1TY67

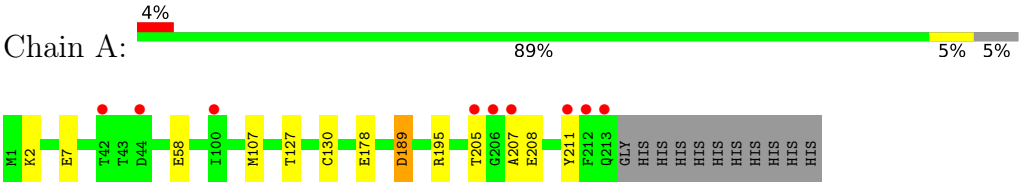
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	133	Total O 133 133	0	0
2	B	85	Total O 85 85	0	0

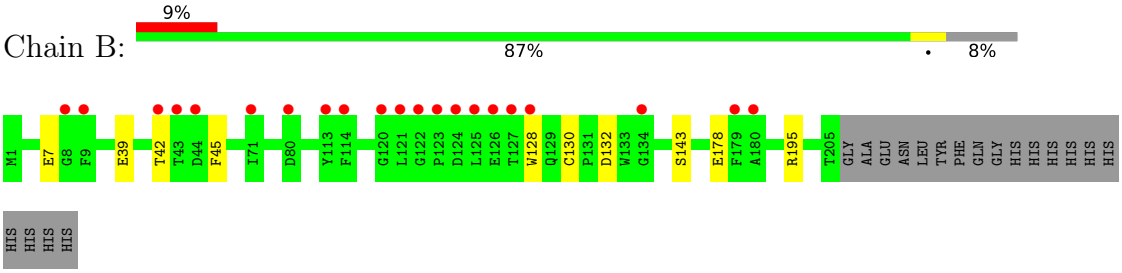
3 Residue-property plots (i)

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, N-terminal domain



• Molecule 1: Glutathione S-transferase, N-terminal domain



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.61Å 70.78Å 98.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.70 29.85 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-1.70) 99.5 (29.85-1.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.64 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.193 , 0.221 0.205 , 0.231	Depositor DCC
R_{free} test set	2658 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	22.3	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 46.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3408	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.67	0/1666	0.71	1/2267 (0.0%)
1	B	0.63	0/1596	0.69	0/2173
All	All	0.65	0/3262	0.70	1/4440 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	107	MET	CG-SD-CE	-5.12	92.01	100.20

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1629	0	1585	12	0
1	B	1561	0	1525	6	0
2	A	133	0	0	1	0
2	B	85	0	0	2	0
All	All	3408	0	3110	18	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 18 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:130:CYS:SG	2:B:369:HOH:O	2.48	0.71
1:B:39:GLU:O	1:B:42:THR:HG22	2.02	0.59
1:B:128:TRP:CH2	1:B:130:CYS:HB2	2.41	0.55
1:A:205:THR:CG2	1:A:207:ALA:HB3	2.39	0.53
1:B:42:THR:HG23	1:B:45:PHE:H	1.73	0.53

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	211/224 (94%)	207 (98%)	4 (2%)	0	100	100
1	B	203/224 (91%)	201 (99%)	2 (1%)	0	100	100
All	All	414/448 (92%)	408 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	166/178 (93%)	165 (99%)	1 (1%)	86	80
1	B	160/178 (90%)	158 (99%)	2 (1%)	69	56
All	All	326/356 (92%)	323 (99%)	3 (1%)	78	70

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	189	ASP
1	B	7	GLU
1	B	143	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	190	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)

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

6.1 Protein, DNA and RNA chains

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^{th} 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	213/224 (95%)	0.15	9 (4%) 36 40	13, 22, 43, 69	0
1	B	205/224 (91%)	0.55	21 (10%) 6 7	16, 30, 51, 75	0
All	All	418/448 (93%)	0.35	30 (7%) 15 17	13, 27, 49, 75	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	125	LEU	5.3
1	B	127	THR	5.2
1	A	206	GLY	4.7
1	B	44	ASP	4.4
1	B	123	PRO	4.3

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.