



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:06 pm BST

PDB ID : 2CY9
Title : Crystal structure of thioesterase superfamily member2 from *Mus musculus*
Authors : Hosaka, T.; Murayama, K.; Kishishita, S.; Shirouzu, M.; Yokoyama, S.;
RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2005-07-06
Resolution : 2.72 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 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 : 2.11
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) : 2.11

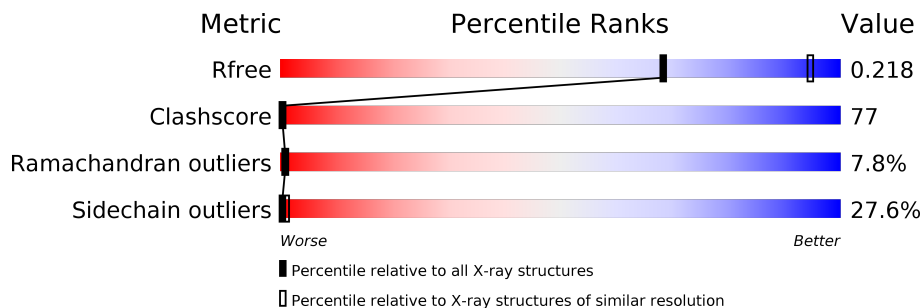
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.72 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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)

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 on 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	140	
1	B	140	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 1959 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 Thioesterase superfamily member 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	122	912	580	153	171	2	6	0	0	0
1	B	133	1004	631	171	191	2	9	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	4	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	12	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	15	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	42	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	70	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	73	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	86	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
A	91	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	4	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	12	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	15	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	42	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	70	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	73	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	86	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4
B	91	MSE	MET	MODIFIED RESIDUE	UNP Q9CQR4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	24	Total	O	0	0
			24	24		

Continued on next page...

Continued from previous page...

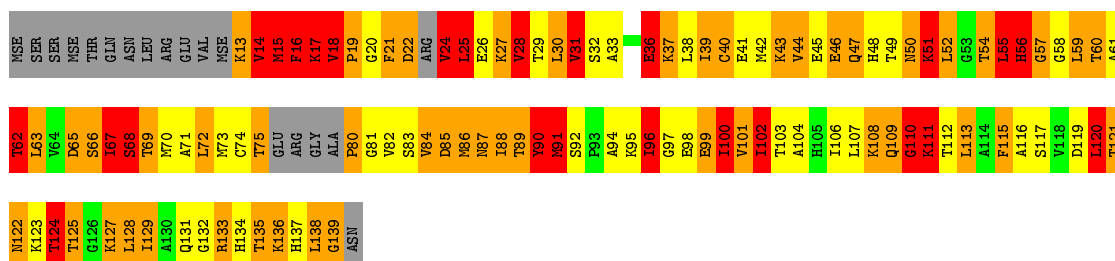
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	19	Total	O	0	0
			19	19		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

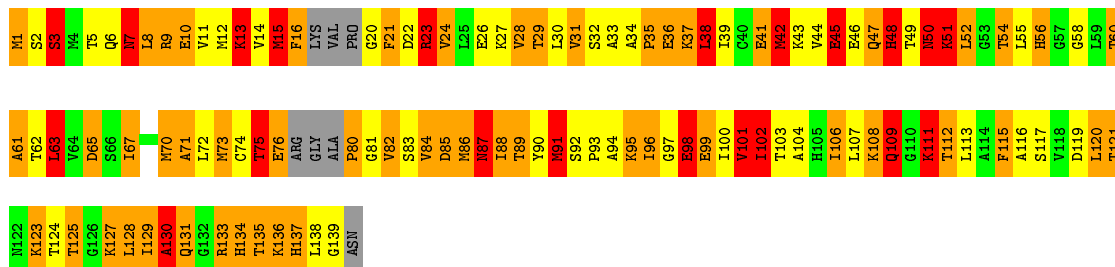
- Molecule 1: Thioesterase superfamily member 2

Chain A: 



- Molecule 1: Thioesterase superfamily member 2

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	66.52Å 90.19Å 61.21Å 90.00° 118.32° 90.00°	Depositor
Resolution (Å)	49.10 – 2.72 49.12 – 2.72	Depositor EDS
% Data completeness (in resolution range)	95.4 (49.10-2.72) 95.4 (49.12-2.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.56 (at 2.73Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.191 , 0.262 0.216 , 0.218	Depositor DCC
R_{free} test set	768 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	50.8	Xtrriage
Anisotropy	0.311	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 109.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	1959	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.82% 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	3.43	128/916 (14.0%)	2.36	53/1222 (4.3%)
1	B	3.42	136/1005 (13.5%)	2.34	56/1335 (4.2%)
All	All	3.42	264/1921 (13.7%)	2.35	109/2557 (4.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	B	0	2
All	All	0	7

The worst 5 of 264 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	36	GLU	CD-OE1	14.22	1.41	1.25
1	B	117	SER	CB-OG	13.96	1.60	1.42
1	A	135	THR	CA-CB	13.47	1.88	1.53
1	B	41	GLU	CD-OE2	13.44	1.40	1.25
1	A	90	TYR	CE1-CZ	13.34	1.55	1.38

The worst 5 of 109 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	133	ARG	NE-CZ-NH1	-15.97	112.32	120.30
1	B	65	ASP	CB-CG-OD2	11.32	128.49	118.30
1	A	121	THR	CA-CB-CG2	-10.27	98.03	112.40
1	A	66	SER	CA-CB-OG	-9.71	84.99	111.20
1	B	119	ASP	CB-CG-OD1	9.71	127.03	118.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	125	THR	Peptide
1	A	14	VAL	Peptide
1	A	16	PHE	Peptide
1	A	57	GLY	Peptide
1	A	91	MSE	Mainchain

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	912	0	963	181	0
1	B	1004	0	1053	137	0
2	A	24	0	0	2	0
2	B	19	0	0	0	0
All	All	1959	0	2016	303	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 77.

The worst 5 of 303 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:111:LYS:CG	1:B:111:LYS:CD	1.75	1.65
1:A:24:VAL:CG1	1:A:24:VAL:CB	1.74	1.65
1:B:88:ILE:CB	1:B:88:ILE:CG2	1.77	1.62
1:B:75:THR:CG2	1:B:75:THR:CB	1.80	1.59
1:A:30:LEU:CG	1:A:30:LEU:CD1	1.75	1.59

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	116/140 (83%)	82 (71%)	20 (17%)	14 (12%)	0	0
1	B	127/140 (91%)	111 (87%)	11 (9%)	5 (4%)	3	6
All	All	243/280 (87%)	193 (79%)	31 (13%)	19 (8%)	1	1

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	15	MSE
1	A	51	LYS
1	A	110	GLY
1	B	109	GLN
1	B	131	GLN

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	103/110 (94%)	73 (71%)	30 (29%)	0	1
1	B	114/110 (104%)	84 (74%)	30 (26%)	0	1
All	All	217/220 (99%)	157 (72%)	60 (28%)	0	1

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	127	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	8	LEU
1	B	107	LEU
1	B	1	MSE
1	B	15	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	122	ASN
1	A	134	HIS
1	B	56	HIS
1	A	109	GLN
1	B	50	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 carbohydrates 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	52:LEU	C	53:GLY	N	1.20

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.