



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 01:10 pm BST

PDB ID : 3LKD
Title : Crystal Structure of the type I restriction-modification system methyltransferase subunit from *Streptococcus thermophilus*, Northeast Structural Genomics Consortium Target SuR80
Authors : Vorobiev, S.; Su, M.; Seetharaman, J.; Mao, M.; Xiao, R.; Foote, E.L.; Ciccosanti, C.; Wang, D.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2010-01-27
Resolution : 2.25 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

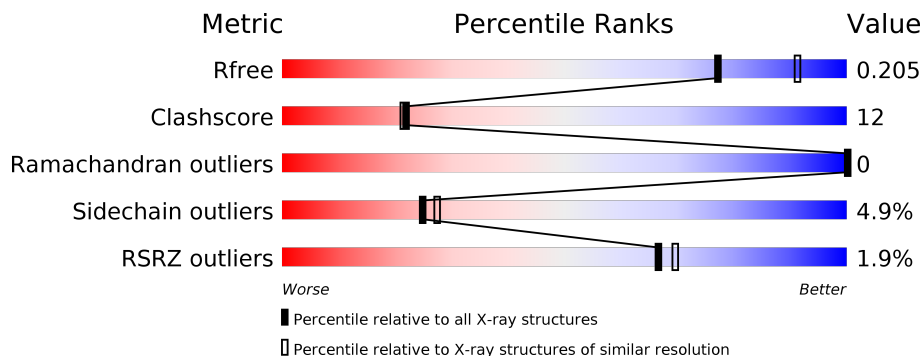
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 $\leq 5\%$. 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	542	 2% 66% 18% • 14%
1	B	542	 % 67% 19% • 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7930 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 Type I restriction-modification system methyltransferase subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	468	3733	2381	606	731	15	0	0	0
1	B	472	3766	2400	612	739	15	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	535	LEU	-	EXPRESSION TAG	UNP Q5M500
A	536	GLU	-	EXPRESSION TAG	UNP Q5M500
A	537	HIS	-	EXPRESSION TAG	UNP Q5M500
A	538	HIS	-	EXPRESSION TAG	UNP Q5M500
A	539	HIS	-	EXPRESSION TAG	UNP Q5M500
A	540	HIS	-	EXPRESSION TAG	UNP Q5M500
A	541	HIS	-	EXPRESSION TAG	UNP Q5M500
A	542	HIS	-	EXPRESSION TAG	UNP Q5M500
B	535	LEU	-	EXPRESSION TAG	UNP Q5M500
B	536	GLU	-	EXPRESSION TAG	UNP Q5M500
B	537	HIS	-	EXPRESSION TAG	UNP Q5M500
B	538	HIS	-	EXPRESSION TAG	UNP Q5M500
B	539	HIS	-	EXPRESSION TAG	UNP Q5M500
B	540	HIS	-	EXPRESSION TAG	UNP Q5M500
B	541	HIS	-	EXPRESSION TAG	UNP Q5M500
B	542	HIS	-	EXPRESSION TAG	UNP Q5M500

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	205	Total	O	0	0
			205	205		
2	B	226	Total	O	0	0
			226	226		

THR
THR
PRO
GLU
ALA
ASP
GLU
GLU
LEU
LYS
ALA
PHE
VAL
LYS
LYS
PHE
LYS
GLY
LEU
GLU
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	60.87Å 83.13Å 96.68Å 90.00° 98.13° 90.00°	Depositor
Resolution (Å)	48.79 – 2.25 48.79 – 2.25	Depositor EDS
% Data completeness (in resolution range)	94.9 (48.79-2.25) 99.5 (48.79-2.25)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.49 (at 2.24Å)	Xtrriage
Refinement program	CNS 1.2, REFMAC	Depositor
R, R_{free}	0.186 , 0.210 0.187 , 0.205	Depositor DCC
R_{free} test set	2279 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	19.8	Xtrriage
Anisotropy	0.181	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 42.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7930	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% 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	0.33	0/3800	0.58	0/5124
1	B	0.34	0/3834	0.59	0/5169
All	All	0.34	0/7634	0.59	0/10293

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	3733	0	3593	83	0
1	B	3766	0	3635	92	0
2	A	205	0	0	5	0
2	B	226	0	0	9	0
All	All	7930	0	7228	174	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 12.

The worst 5 of 174 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:230:MSE:HE3	1:B:236:LEU:HD12	1.27	1.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:158:ALA:HA	1:B:161:MSE:HE3	1.36	1.07
1:B:268:MSE:HE2	1:B:280:LEU:HD21	1.39	1.03
1:A:230:MSE:HE3	1:A:236:LEU:HD12	1.44	0.99
1:B:157:VAL:HG12	1:B:161:MSE:HE2	1.44	0.97

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	464/542 (86%)	444 (96%)	20 (4%)	0	100	100
1	B	470/542 (87%)	454 (97%)	16 (3%)	0	100	100
All	All	934/1084 (86%)	898 (96%)	36 (4%)	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	399/449 (89%)	381 (96%)	18 (4%)	27	31
1	B	404/449 (90%)	383 (95%)	21 (5%)	23	24
All	All	803/898 (89%)	764 (95%)	39 (5%)	25	27

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	470	ASN
1	B	61	GLU
1	B	441	LEU
1	B	7	THR
1	B	12	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	470	ASN
1	B	127	GLN
1	B	408	ASN
1	B	96	HIS
1	B	203	GLN

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

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	453/542 (83%)	-0.08	10 (2%) 62 65	9, 23, 45, 68	0
1	B	457/542 (84%)	-0.21	7 (1%) 73 75	9, 20, 36, 49	0
All	All	910/1084 (83%)	-0.14	17 (1%) 66 69	9, 21, 41, 68	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	196	ALA	4.2
1	A	172	HIS	4.0
1	A	59	GLU	4.0
1	A	197	GLY	3.5
1	B	478	PHE	3.3

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 carbohydrates 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.