wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 03:32 PM EST

PDB ID	:	8SUJ
Title	:	Joint X-ray/neutron structure of Thermus thermophilus serine hydroxymeth yltransferase (TthSHMT) in internal aldimine state
Authors	:	Drago, V.N.; Kovalevsky, A.
Deposited on	:	2023-05-12
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 (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:

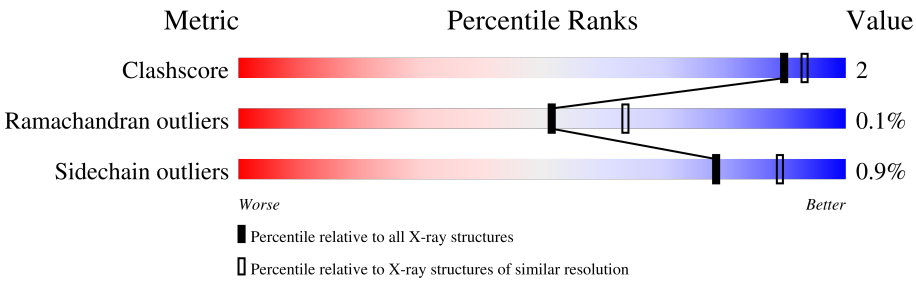
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION, NEUTRON 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))
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (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 <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	405	97%
1	B	405	97%

2 Entry composition (i)

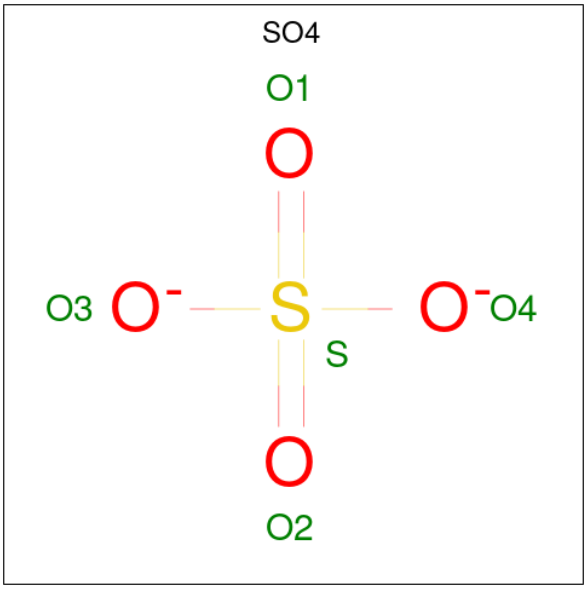
There are 3 unique types of molecules in this entry. The entry contains 14922 atoms, of which 6085 are hydrogens and 2148 are deuteriums.

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 Serine hydroxymethyltransferase.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
			Total	C	D	H	N	O	P	S			
1	A	402	6810	1984	647	3038	556	575	1	9	0	305	0
1	B	402	6815	1984	643	3047	556	575	1	9	0	322	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	D	O		
3	A	221	663	442	221	0	0
			Total	D	O		
3	B	208	624	416	208	0	0

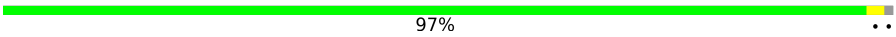
3 Residue-property plots (i)

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. 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.

Note EDS failed to run properly.

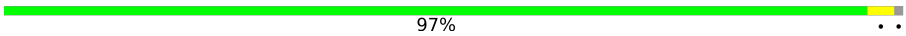
- Molecule 1: Serine hydroxymethyltransferase

Chain A:



- Molecule 1: Serine hydroxymethyltransferase

Chain B:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.81Å 83.33Å 95.57Å 90.00° 91.68° 90.00°	Depositor
Resolution (Å)	40.00 – 2.30	Depositor
% Data completeness (in resolution range)	75.5 (40.00-2.30)	Depositor
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.00 (at 2.00Å)	Xtriage
Refinement program	nCNS 1.0.8	Depositor
R, R_{free}	0.210 , 0.239	Depositor
Wilson B-factor (Å^2)	11.0	Xtriage
Anisotropy	0.053	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtriage
Total number of atoms	14922	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.13% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: LLP, DOD, SO4

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	0.37	0/5640	0.62	0/7633
1	B	0.38	0/5753	0.62	0/7778
All	All	0.38	0/11393	0.62	0/15411

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	B	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1040[A]	ARG	Sidechain
1	B	1040[B]	ARG	Sidechain
1	B	1156[A]	ARG	Sidechain
1	B	1156[B]	ARG	Sidechain

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	3772	3038	649	4	0
1	B	3768	3047	596	6	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
3	A	663	0	0	2	0
3	B	624	0	0	2	0
All	All	8837	6085	1245	15	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 2.

The worst 5 of 15 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)

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	703/405 (174%)	687 (98%)	16 (2%)	0	100 100
1	B	720/405 (178%)	702 (98%)	16 (2%)	2 (0%)	41 50
All	All	1423/810 (176%)	1389 (98%)	32 (2%)	2 (0%)	51 64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1361[A]	THR
1	B	1361[B]	THR

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	570/322 (177%)	562 (99%)	8 (1%)	67	81
1	B	575/322 (179%)	571 (99%)	4 (1%)	84	92
All	All	1145/644 (178%)	1133 (99%)	12 (1%)	78	87

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	391[B]	GLU
1	B	1051[A]	TYR
1	B	1157[B]	ARG
1	B	1051[B]	TYR
1	A	157[B]	ARG

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)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	LLP	A	226	1	23,24,25	2.01	6 (26%)	25,32,34	2.57	10 (40%)
1	LLP	B	1226	1	23,24,25	1.94	5 (21%)	25,32,34	2.58	9 (36%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	A	226	1	-	4/16/17/19	0/1/1/1
1	LLP	B	1226	1	-	5/16/17/19	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1226	LLP	C3-C2	5.49	1.46	1.40
1	A	226	LLP	C4-C3	4.74	1.47	1.40
1	A	226	LLP	C3-C2	4.23	1.45	1.40
1	A	226	LLP	C4-C5	3.92	1.46	1.42
1	B	1226	LLP	C4-C3	3.90	1.46	1.40

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1226	LLP	OP4-C5'-C5	8.09	124.77	109.35
1	A	226	LLP	OP4-C5'-C5	8.00	124.60	109.35
1	B	1226	LLP	C6-N1-C2	5.14	128.70	119.17
1	A	226	LLP	C6-N1-C2	5.05	128.52	119.17
1	B	1226	LLP	C3-C2-N1	-3.60	116.12	120.77

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	226	LLP	O-C-CA-CB
1	B	1226	LLP	C4-C5-C5'-OP4
1	B	1226	LLP	C6-C5-C5'-OP4
1	B	1226	LLP	O-C-CA-CB
1	A	226	LLP	CD-CE-NZ-C4'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	B	1501	-	4,4,4	0.14	0	6,6,6	0.18	0
2	SO4	A	2001	-	4,4,4	0.24	0	6,6,6	0.16	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	SO4	1	0

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)

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates (i)

EDS failed to run properly - this section is therefore empty.

6.4 Ligands (i)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers (i)

EDS failed to run properly - this section is therefore empty.