Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2024 – 06:09 PM EDT

PDB ID	:	9BPE
Title	:	Joint X-ray/neutron structure of Thermus thermophilus serine hydroxymethyltransferase (TthSHMT) in internal aldimine state and folinic acid bound
Authors	:	Drago, V.N.; Kovalevsky, A.
Deposited on	:	2024-05-07
Resolution	:	2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

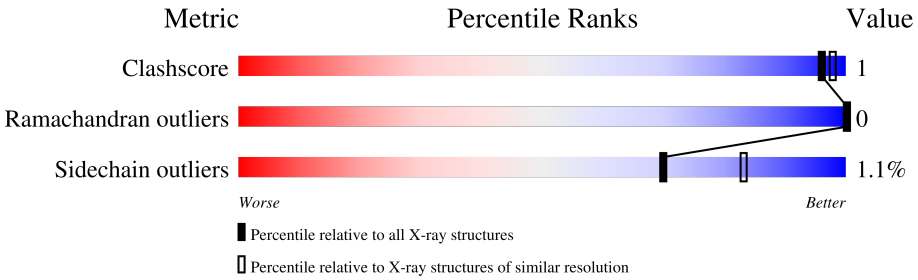
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (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	402	98%
1	B	402	98%

2 Entry composition

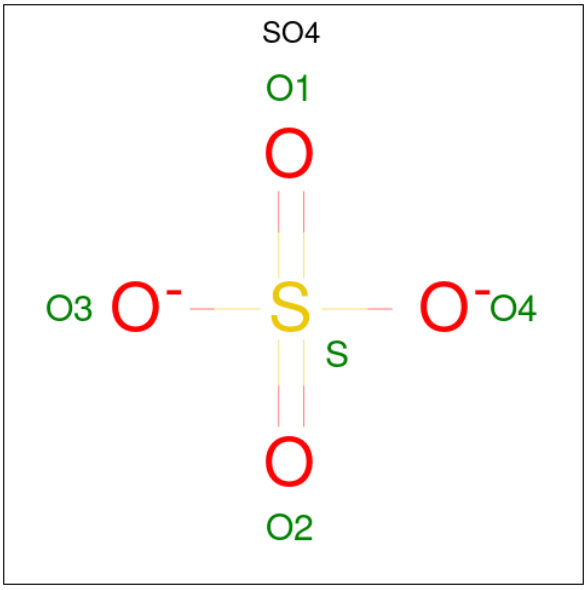
There are 5 unique types of molecules in this entry. The entry contains 15066 atoms, of which 6182 are hydrogens and 2161 are deuteriums.

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.

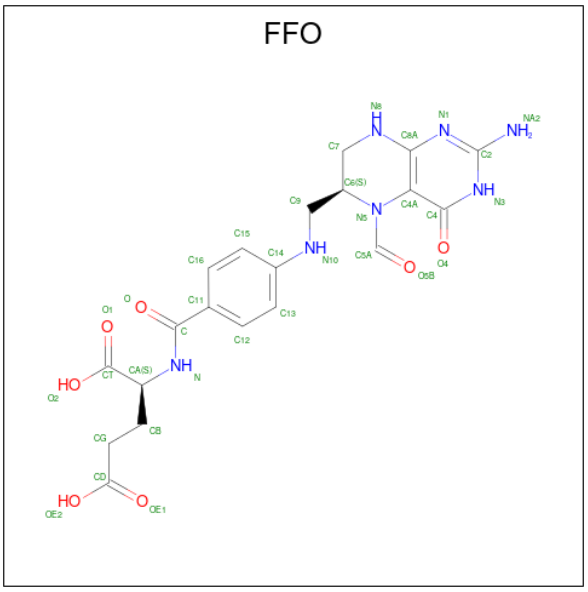
Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
1	A	401	Total	C	D	H	N	O	P	S	0	326	0
			6833	1979	641	3074	555	574	1	9			
1	B	402	Total	C	D	H	N	O	P	S	0	334	0
			6855	1984	640	3090	556	575	1	9			

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



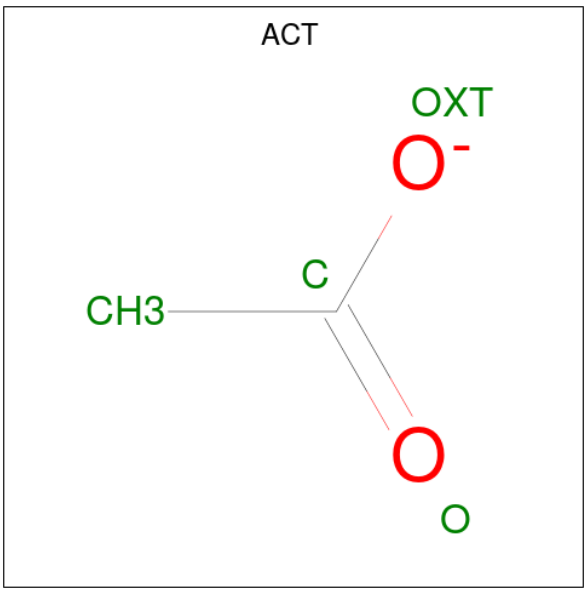
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is N-[4-({[(6S)-2-amino-5-formyl-4-oxo-3,4,5,6,7,8-hexahydropteridin-6-yl]methyl }amino)benzoyl]-L-glutamic acid (three-letter code: FFO) (formula: $C_{20}H_{23}N_7O_7$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	B	1	Total	C	D	H	N	O	0	0
			55	20	6	15	7	7		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	218	Total 654	D 436	O 218	0	0
5	B	219	Total 657	D 438	O 219	0	0

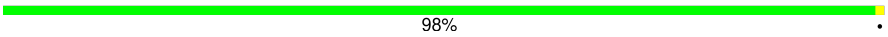
3 Residue-property plots (i)

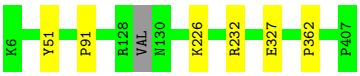
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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

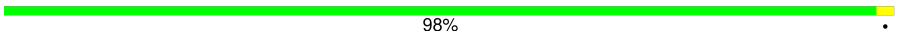
Chain A:





- 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	58.76Å 83.48Å 95.56Å 90.00° 91.52° 90.00°	Depositor
Resolution (Å)	14.51 – 2.30	Depositor
% Data completeness (in resolution range)	92.5 (14.51-2.30)	Depositor
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.00 (at 2.00Å)	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	0.212 , 0.239	Depositor
Wilson B-factor (Å^2)	5.0	Xtriage
Anisotropy	0.702	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.044 for h,-k,-l	Xtriage
Total number of atoms	15066	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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

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

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	0.35	0/5765	0.59	0/7792
1	B	0.35	0/5832	0.59	0/7889
All	All	0.35	0/11597	0.59	0/15681

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

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	3759	3074	553	3	0
1	B	3765	3090	494	3	0
2	A	5	0	0	1	0
3	B	40	15	21	0	0
4	B	4	3	3	0	0
5	A	654	0	0	0	0
5	B	657	0	0	1	0
All	All	8884	6182	1071	8	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 1.

All (8) 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	719/402 (179%)	701 (98%)	18 (2%)	0	100	100
1	B	732/402 (182%)	712 (97%)	20 (3%)	0	100	100
All	All	1451/804 (180%)	1413 (97%)	38 (3%)	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	576/319 (181%)	572 (99%)	4 (1%)	81	90
1	B	585/319 (183%)	575 (98%)	10 (2%)	56	72
All	All	1161/638 (182%)	1147 (99%)	14 (1%)	70	81

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51[A]	TYR
1	A	51[B]	TYR
1	A	232[A]	ARG
1	A	232[B]	ARG
1	B	1051[A]	TYR
1	B	1051[B]	TYR
1	B	1281[A]	GLU
1	B	1281[B]	GLU
1	B	1342[A]	ASN
1	B	1342[B]	ASN
1	B	1399[A]	ARG
1	B	1399[B]	ARG
1	B	1406[A]	MET
1	B	1406[B]	MET

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	1.88	5 (21%)	25,32,34	2.52	10 (40%)
1	LLP	B	1226	1	23,24,25	1.71	6 (26%)	25,32,34	2.51	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	-	6/16/17/19	0/1/1/1
1	LLP	B	1226	1	-	6/16/17/19	0/1/1/1

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	226	LLP	C3-C2	5.02	1.46	1.41
1	A	226	LLP	C4-C5	3.82	1.47	1.42
1	B	1226	LLP	C3-C2	3.70	1.44	1.41
1	A	226	LLP	C4-C3	3.43	1.46	1.41
1	B	1226	LLP	C4-C5	3.34	1.46	1.42
1	B	1226	LLP	CB-CA	3.22	1.58	1.53
1	B	1226	LLP	C4-C3	3.00	1.46	1.41
1	A	226	LLP	CB-CA	2.64	1.57	1.53
1	B	1226	LLP	OP4-C5'	-2.25	1.36	1.44
1	B	1226	LLP	CD-CE	2.22	1.59	1.51
1	A	226	LLP	P-OP4	2.06	1.66	1.60

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	226	LLP	OP4-C5'-C5	7.92	124.20	109.36
1	B	1226	LLP	OP4-C5'-C5	7.89	124.14	109.36
1	A	226	LLP	C6-N1-C2	5.07	128.38	119.20
1	B	1226	LLP	C6-N1-C2	4.88	128.05	119.20
1	B	1226	LLP	C3-C2-N1	-3.26	116.85	120.96
1	A	226	LLP	C2'-C2-C3	3.18	124.51	120.80
1	A	226	LLP	C3-C2-N1	-3.14	117.00	120.96
1	B	1226	LLP	C5-C4-C4'	3.02	126.12	121.47
1	B	1226	LLP	C2'-C2-C3	2.99	124.30	120.80
1	B	1226	LLP	C5-C6-N1	-2.97	119.00	123.83
1	A	226	LLP	C5-C4-C4'	2.94	126.00	121.47
1	A	226	LLP	C5-C6-N1	-2.90	119.12	123.83
1	A	226	LLP	C3-C4-C5	-2.40	116.35	118.28
1	B	1226	LLP	CE-NZ-C4'	2.38	126.35	118.72
1	A	226	LLP	C4-C3-C2	-2.38	118.80	120.14
1	A	226	LLP	CE-NZ-C4'	2.19	125.72	118.72
1	B	1226	LLP	C4-C3-C2	-2.18	118.92	120.14
1	B	1226	LLP	C3-C4-C5	-2.08	116.61	118.28
1	A	226	LLP	O3-C3-C4	2.00	124.83	119.44

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	226	LLP	C4-C5-C5'-OP4
1	A	226	LLP	C6-C5-C5'-OP4
1	A	226	LLP	O-C-CA-CB
1	A	226	LLP	CG-CD-CE-NZ
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	B	1226	LLP	CG-CD-CE-NZ
1	B	1226	LLP	CD-CE-NZ-C4'
1	A	226	LLP	CD-CE-NZ-C4'
1	B	1226	LLP	C3-C4-C4'-NZ
1	A	226	LLP	C3-C4-C4'-NZ

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	A	501	-	4,4,4	0.31	0	6,6,6	0.15	0
4	ACT	B	2002	-	3,3,3	1.29	0	3,3,3	1.61	1 (33%)
3	FFO	B	2001	-	33,36,36	1.65	8 (24%)	35,50,50	1.42	4 (11%)

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
3	FFO	B	2001	-	-	6/24/37/37	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2001	FFO	C4A-N5	4.16	1.44	1.38
3	B	2001	FFO	C15-C14	2.97	1.44	1.39
3	B	2001	FFO	C13-C14	2.82	1.44	1.39
3	B	2001	FFO	C8A-N1	2.79	1.40	1.36
3	B	2001	FFO	C16-C11	2.78	1.43	1.39
3	B	2001	FFO	C12-C11	2.66	1.43	1.39
3	B	2001	FFO	C7-C6	2.09	1.54	1.52
3	B	2001	FFO	CG-CD	2.06	1.55	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	2001	FFO	C2-N1-C8A	4.13	120.65	113.36
3	B	2001	FFO	C4A-C4-N3	3.74	117.53	110.94
3	B	2001	FFO	C2-N3-C4	-2.32	120.90	125.11
4	B	2002	ACT	OXT-C-O	2.08	129.76	122.03
3	B	2001	FFO	CG-CB-CA	2.08	117.00	113.16

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	2001	FFO	O5B-C5A-N5-C6
3	B	2001	FFO	C6-C9-N10-C14
3	B	2001	FFO	OE1-CD-CG-CB
3	B	2001	FFO	OE2-CD-CG-CB
3	B	2001	FFO	C13-C14-N10-C9
3	B	2001	FFO	CT-CA-CB-CG

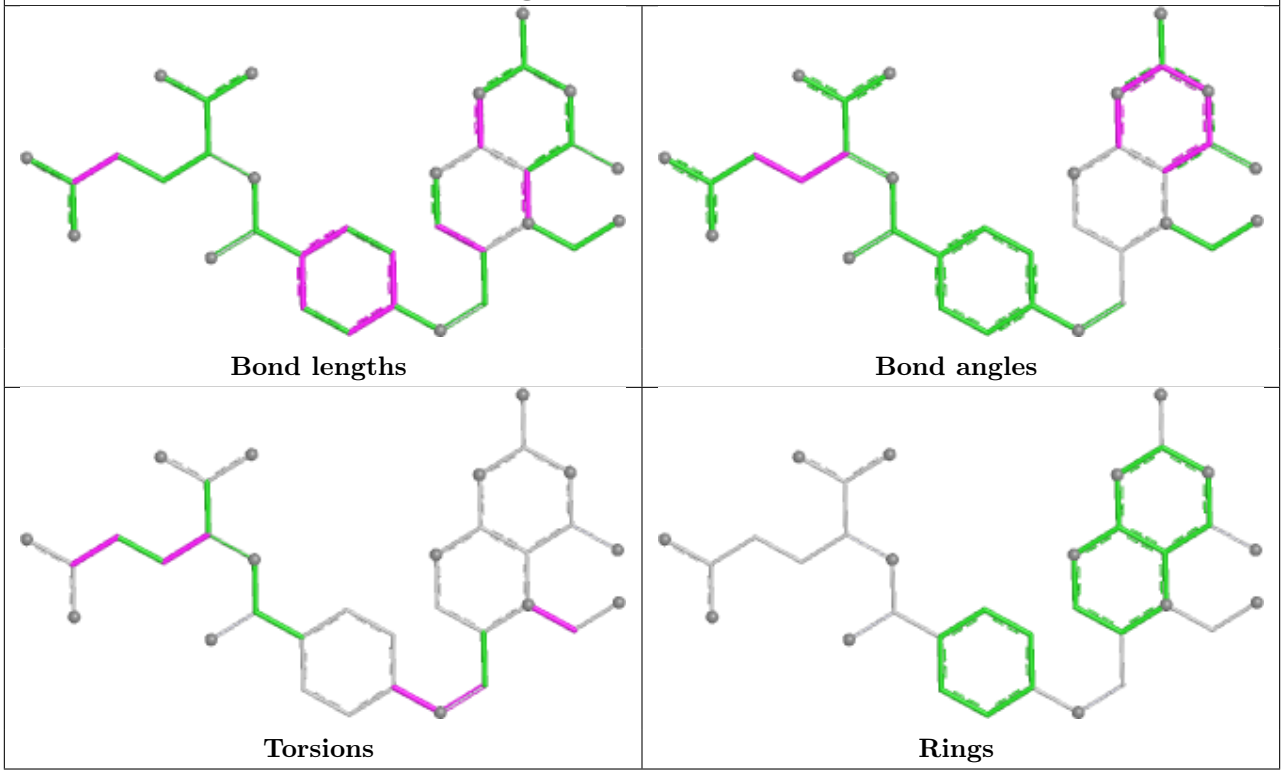
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	SO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand FFO B 2001



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

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

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

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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