



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 26, 2026 – 12:26 PM EST

PDB ID : 9O5G / pdb_00009o5g
Title : Room-temperature joint X-ray/Neutron structure of *Thermus thermophilus* SHMT in complex with PLP-Gly external aldimine and 5-methyl-tetrahydrofolate (5MTHF)
Authors : Kovalevsky, A.; Drago, V.N.; Phillips, R.S.
Deposited on : 2025-04-10
Resolution : 2.10 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
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) : 2.47

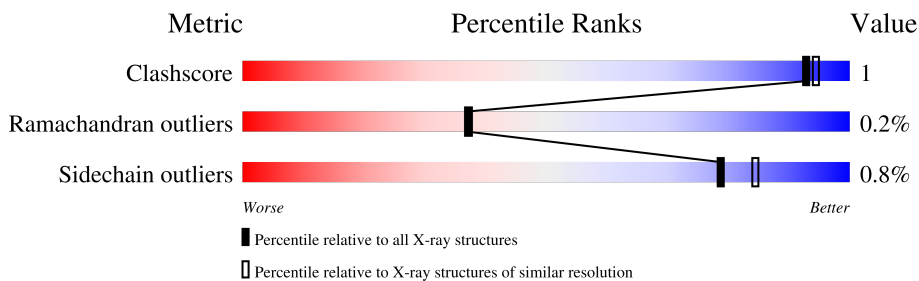
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION, NEUTRON DIFFRACTION

The reported resolution of this entry is 2.10 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	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	405	96% . .
1	B	405	95% . .

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 14376 atoms, of which 5665 are hydrogens and 2016 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	402	Total	C	D	H	N	O	S	0	207	0
			6538	1976	609	2819	555	570	9			
1	B	402	Total	C	D	H	N	O	S	0	203	0
			6535	1976	605	2820	555	570	9			

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



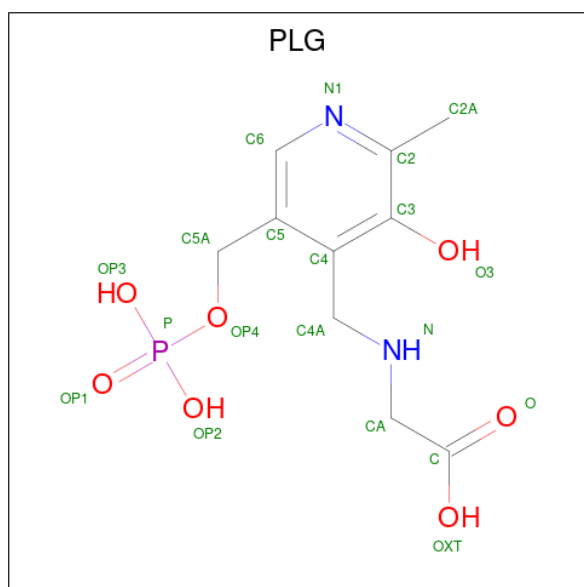
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

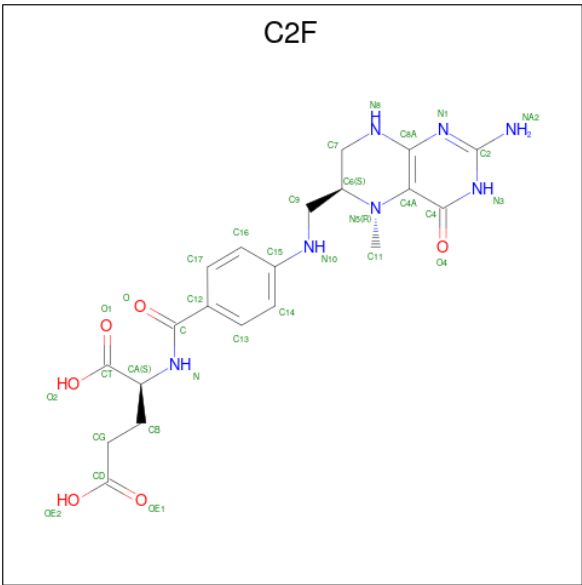
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is N-GLYCINE-[3-HYDROXY-2-METHYL-5-PHOSPHONOXYMETHYL-PYRIDIN-4-YL-METHANE] (CCD ID: PLG) (formula: $C_{10}H_{15}N_2O_7P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
3	B	1	Total	C	D	H	N	O	P	0	0
			31	10	2	9	2	7	1		

- Molecule 4 is 5-METHYL-5,6,7,8-TETRAHYDROFOLIC ACID (CCD ID: C2F) (formula: $C_{20}H_{25}N_7O_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	B	1	Total	C	D	H	N	O	0	0
			56	20	6	17	7	6		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	197	Total	D	O	0	0
			591	394	197		
5	B	200	Total	D	O	0	0
			600	400	200		

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:  96%



- Molecule 1: Serine hydroxymethyltransferase

Chain B:  95%



4 Data and refinement statistics

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	58.92Å 83.68Å 95.52Å 90.00° 91.64° 90.00°	Depositor
Resolution (Å)	36.58 – 2.10	Depositor
% Data completeness (in resolution range)	67.5 (36.58-2.10)	Depositor
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.00 (at 1.70Å)	Xtriage
Refinement program	nCNS 1.0.8	Depositor
R, R_{free}	0.292 , 0.327	Depositor
Wilson B-factor (Å ²)	31.5	Xtriage
Anisotropy	0.342	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.053 for h,-k,-l	Xtriage
Total number of atoms	14376	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.74% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: C2F, SO4, DOD, PLG

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	0.49	0/4847	0.88	8/6563 (0.1%)
1	B	0.53	0/4820	0.84	11/6531 (0.2%)
All	All	0.51	0/9667	0.86	19/13094 (0.1%)

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1225	HIS	N-CA-C	7.32	121.92	112.92
1	A	138	VAL	N-CA-C	6.60	118.10	108.53
1	A	223	THR	N-CA-C	-6.57	99.95	110.14
1	A	30	ALA	N-CA-C	6.33	119.85	111.75
1	B	1341	LYS	N-CA-C	-6.08	102.08	110.35
1	B	1030	ALA	N-CA-C	5.79	118.37	111.71
1	A	200	HIS	N-CA-C	5.75	117.22	111.07
1	B	1138	VAL	N-CA-C	5.66	116.74	108.53
1	B	1223	THR	N-CA-C	-5.44	102.15	110.14
1	A	90[A]	GLN	CA-C-N	-5.31	114.10	119.83
1	A	90[A]	GLN	C-N-CA	-5.31	114.10	119.83
1	A	90[B]	GLN	CA-C-N	-5.31	114.10	119.83
1	A	90[B]	GLN	C-N-CA	-5.31	114.10	119.83
1	B	1310	ASP	N-CA-C	-5.13	107.05	113.72
1	B	1255	ILE	N-CA-C	5.11	119.00	113.43
1	B	1361[A]	THR	CA-C-N	-5.06	113.44	119.05
1	B	1361[A]	THR	C-N-CA	-5.06	113.44	119.05
1	B	1361[B]	THR	CA-C-N	-5.06	113.44	119.05
1	B	1361[B]	THR	C-N-CA	-5.06	113.44	119.05

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	3719	2819	1477	7	0
1	B	3715	2820	1515	6	0
2	A	20	0	0	0	0
2	B	5	0	0	0	0
3	B	22	9	11	1	0
4	B	39	17	23	0	0
5	A	591	0	0	2	1
5	B	600	0	0	2	0
All	All	8711	5665	3026	15	1

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 (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)
1:B:1341:LYS:NZ	5:B:2102:DOD:O	2.41	0.54
5:A:4101:DOD:O	3:B:2001:PLG:H4A1	2.02	0.53
1:B:1372:GLU:CD	1:B:1372:GLU:H	2.16	0.48
1:A:333:ASP:HB2	5:A:4271:DOD:O	2.11	0.45
1:A:223:THR:HB	1:A:225:HIS:CE1	2.48	0.42
1:B:1313:LEU:C	1:B:1313:LEU:HD12	2.40	0.42

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:4200:DOD:O	5:A:4262:DOD:D1[2_546]	1.41	0.79

5.3 Torsion angles

5.3.1 Protein backbone

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	607/405 (150%)	593 (98%)	13 (2%)	1 (0%)	44	45
1	B	603/405 (149%)	582 (96%)	20 (3%)	1 (0%)	44	45
All	All	1210/810 (149%)	1175 (97%)	33 (3%)	2 (0%)	37	45

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	LYS
1	B	1254	GLY

5.3.2 Protein sidechains

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	489/323 (151%)	486 (99%)	3 (1%)	84	89
1	B	485/323 (150%)	479 (99%)	6 (1%)	67	74
All	All	974/646 (151%)	965 (99%)	9 (1%)	79	82

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	TYR
1	A	340[A]	ASN
1	A	340[B]	ASN
1	B	1051[A]	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1051[B]	TYR
1	B	1117[A]	LEU
1	B	1117[B]	LEU
1	B	1342[A]	ASN
1	B	1342[B]	ASN

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](#)

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

7 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	2003	-	4,4,4	0.36	0	6,6,6	0.08	0
2	SO4	A	4004	-	4,4,4	0.33	0	6,6,6	0.23	0
2	SO4	A	4003	-	4,4,4	0.34	0	6,6,6	0.08	0
4	C2F	B	2002	-	33,35,35	1.42	3 (9%)	35,49,49	2.02	6 (17%)
2	SO4	A	4002	-	4,4,4	0.36	0	6,6,6	0.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	PLG	B	2001	-	20,20,20	2.59	7 (35%)	26,28,28	3.01	12 (46%)
2	SO4	A	4001	-	4,4,4	0.37	0	6,6,6	0.12	0

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	PLG	B	2001	-	-	2/12/12/12	0/1/1/1
4	C2F	B	2002	-	-	6/22/35/35	0/3/3/3

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2001	PLG	C4A-C4	-6.70	1.42	1.52
3	B	2001	PLG	O3-C3	-4.56	1.26	1.36
4	B	2002	C2F	C11-N5	-4.01	1.36	1.45
3	B	2001	PLG	C4A-N	-3.96	1.28	1.45
4	B	2002	C2F	C4-N3	3.50	1.39	1.33
3	B	2001	PLG	P-OP4	3.43	1.71	1.60
3	B	2001	PLG	C3-C4	3.20	1.44	1.40
3	B	2001	PLG	C3-C2	3.00	1.44	1.41
3	B	2001	PLG	OXT-C	-2.98	1.20	1.30
4	B	2002	C2F	C2-N3	-2.92	1.30	1.35

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2001	PLG	C4A-N-CA	11.13	125.68	112.72
4	B	2002	C2F	C11-N5-C4A	6.99	122.80	113.30
4	B	2002	C2F	C8A-C4A-C4	6.88	119.77	114.44
3	B	2001	PLG	C6-N1-C2	4.13	126.69	119.20
3	B	2001	PLG	C4-C4A-N	4.07	119.02	111.50
3	B	2001	PLG	C3-C4-C5	3.57	121.96	118.73
4	B	2002	C2F	C2-N3-C4	3.29	120.53	115.96
3	B	2001	PLG	O3-C3-C4	2.78	126.23	118.18
3	B	2001	PLG	C4-C3-C2	-2.67	115.88	119.91
3	B	2001	PLG	C5-C6-N1	-2.65	119.51	123.83
3	B	2001	PLG	C2A-C2-N1	2.56	122.47	117.64
3	B	2001	PLG	C4A-C4-C3	-2.44	116.73	119.98

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	2002	C2F	C2-N1-C8A	2.42	120.35	114.59
3	B	2001	PLG	OP4-C5A-C5	2.27	113.61	109.36
3	B	2001	PLG	OXT-C-O	2.09	128.71	123.33
4	B	2002	C2F	O1-CT-CA	-2.06	115.60	122.26
3	B	2001	PLG	C3-C2-N1	-2.04	118.38	120.96
4	B	2002	C2F	NA2-C2-N3	2.02	120.25	117.22

There are no chirality outliers.

All (8) torsion outliers are listed below:

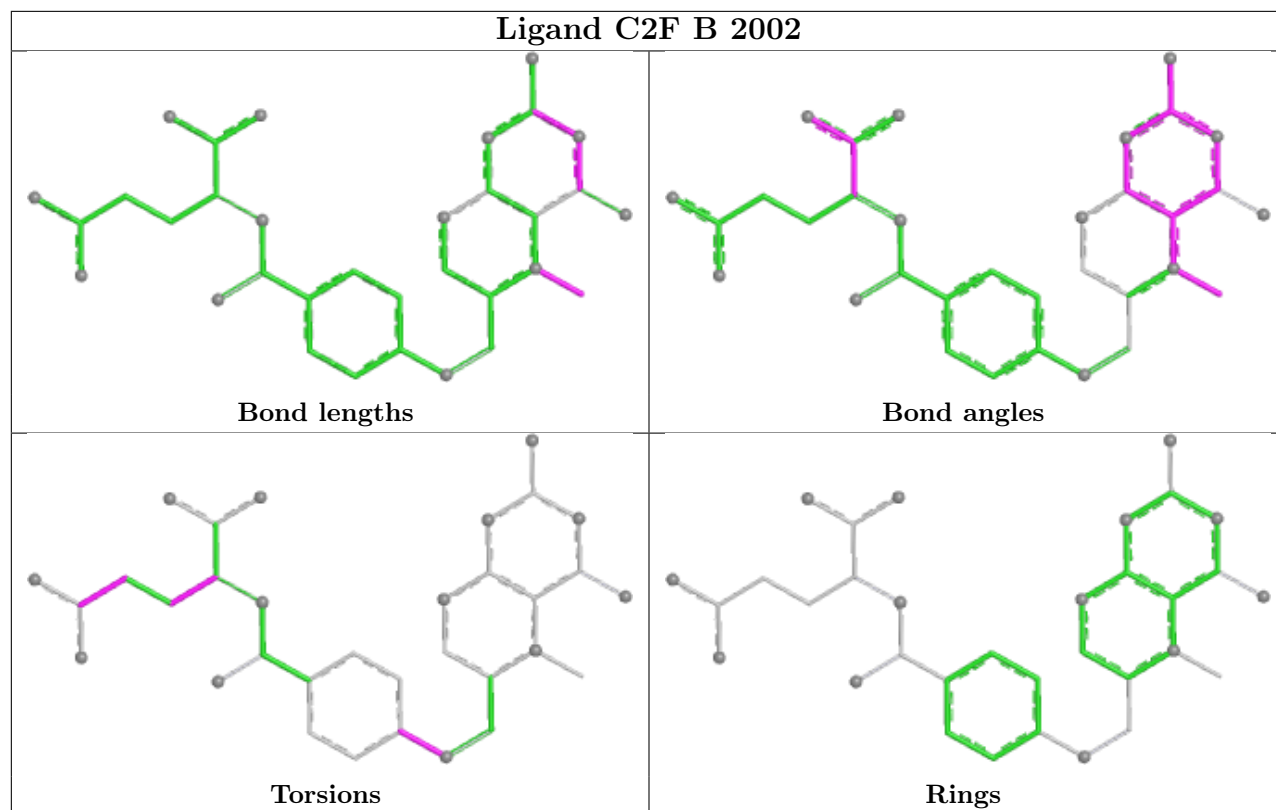
Mol	Chain	Res	Type	Atoms
3	B	2001	PLG	C5-C4-C4A-N
4	B	2002	C2F	C14-C15-N10-C9
4	B	2002	C2F	C16-C15-N10-C9
3	B	2001	PLG	C3-C4-C4A-N
4	B	2002	C2F	CT-CA-CB-CG
4	B	2002	C2F	OE2-CD-CG-CB
4	B	2002	C2F	N-CA-CB-CG
4	B	2002	C2F	OE1-CD-CG-CB

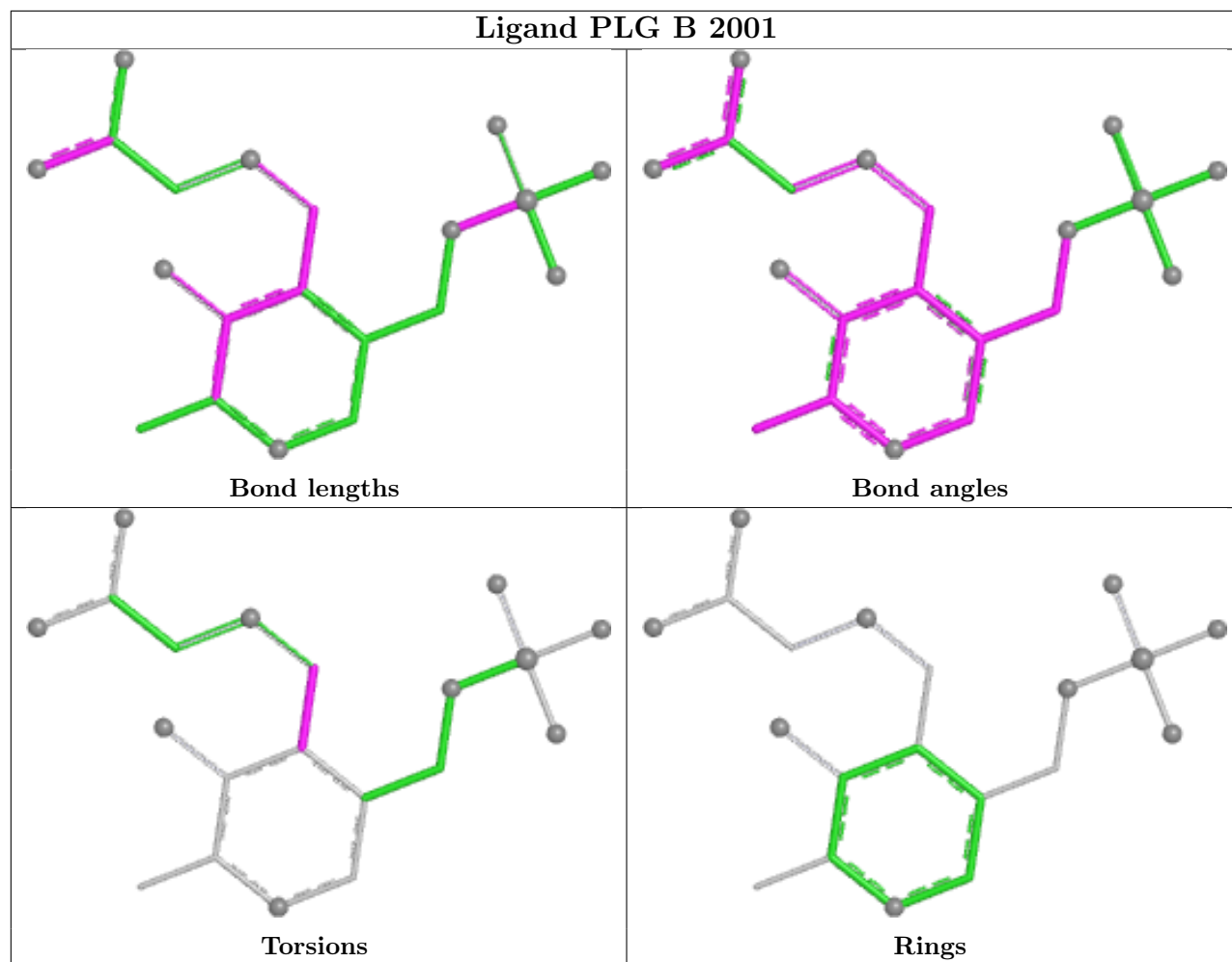
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	2001	PLG	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 than 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.





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.