wwPDB X-ray Structure Validation Summary Report (i)

Oct 4, 2023 – 07:03 PM EDT

PDB ID	:	6PF7
Title	:	Crystal structure of TS-DHFR from Cryptosporidium hominis in complex with NADPH, FdUMP and 2-(4-((2-amino-4-oxo-4,7-dihydro-3H-pyrrolo[2,3-d]pyrimidin-5-yl)methyl)benzamido)benzoic acid
Authors	:	Czyzyk, D.J.; Valhondo, M.; Jorgensen, W.L.; Anderson, K.S.
Deposited on	:	2019-06-21
Resolution	:	2.79 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:

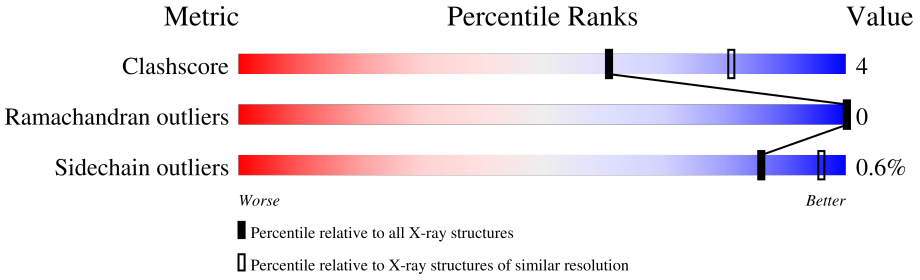
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

2 Entry composition (i)

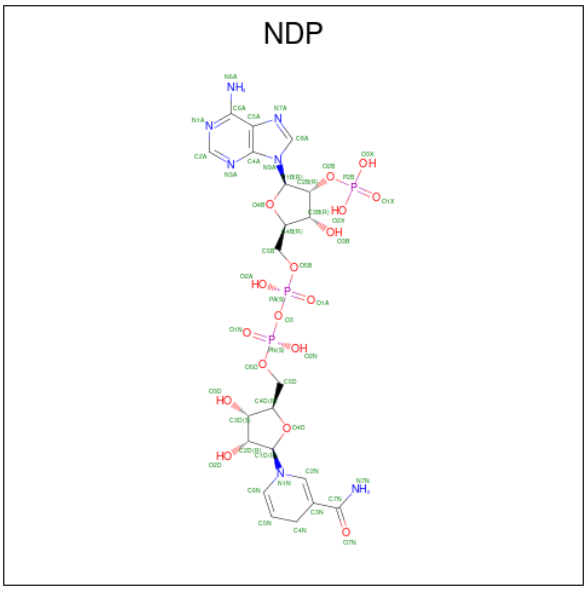
There are 7 unique types of molecules in this entry. The entry contains 21040 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bifunctional dihydrofolate reductase-thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	507	4086	2612	686	766	22	0	0	0
1	B	507	4070	2601	684	764	21	0	0	0
1	C	508	4044	2587	683	753	21	0	0	0
1	D	507	4050	2588	681	760	21	0	0	0
1	E	508	3971	2533	666	751	21	0	0	0

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$).



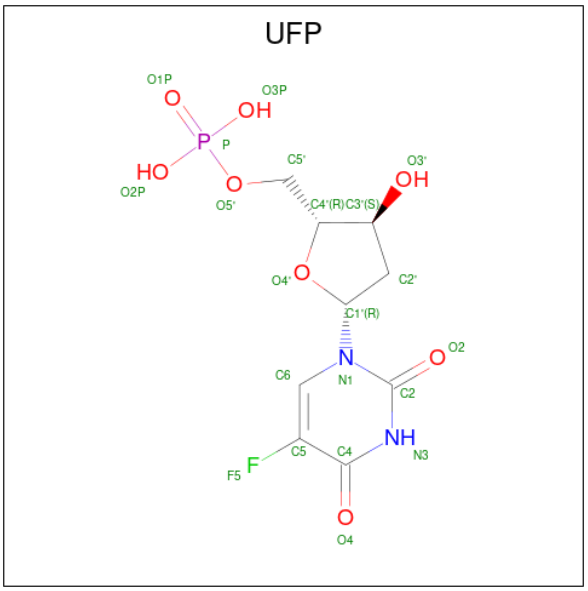
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	48	21	7	17	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	B	1	48	21	7	17	3	0	0
			Total	C	N	O	P		
2	C	1	48	21	7	17	3	0	0
			Total	C	N	O	P		
2	D	1	48	21	7	17	3	0	0
			Total	C	N	O	P		
2	E	1	48	21	7	17	3	0	0

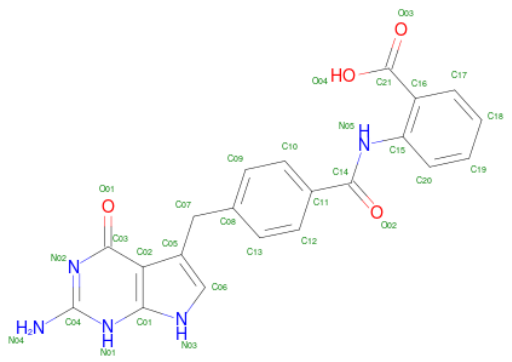
- Molecule 3 is 5-FLUORO-2'-DEOXYURIDINE-5'-MONOPHOSPHATE (three-letter code: UFP) (formula: $C_9H_{12}FN_2O_8P$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	F	N	O	P		
3	A	1	21	9	1	2	8	1	0	0
			Total	C	F	N	O	P		
3	B	1	21	9	1	2	8	1	0	0
			Total	C	F	N	O	P		
3	C	1	21	9	1	2	8	1	0	0
			Total	C	F	N	O	P		
3	D	1	21	9	1	2	8	1	0	0
			Total	C	F	N	O	P		
3	E	1	21	9	1	2	8	1	0	0

- Molecule 4 is 2-({4-[(2-amino-4-oxo-4,7-dihydro-1H-pyrrolo[2,3-d]pyrimidin-5-yl)methyl]benzene-1-carbonyl}amino)benzoic acid (three-letter code: OG7) (formula: $C_{21}H_{17}N_5O_4$) (labeled as "Ligand of Interest" by depositor).

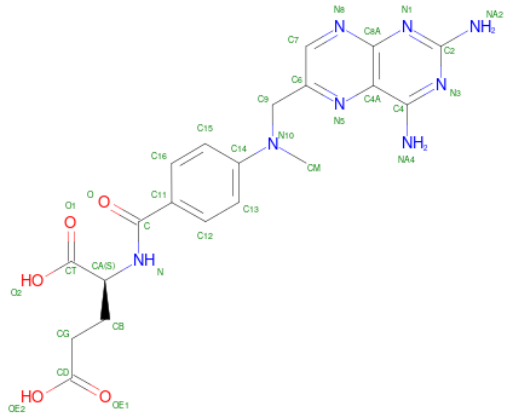
OG7



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	30	21	5	4	0	0
			Total	C	N	O		
4	B	1	30	21	5	4	0	0
			Total	C	N	O		
4	C	1	30	21	5	4	0	0
			Total	C	N	O		
4	D	1	30	21	5	4	0	0
			Total	C	N	O		
4	E	1	30	21	5	4	0	0

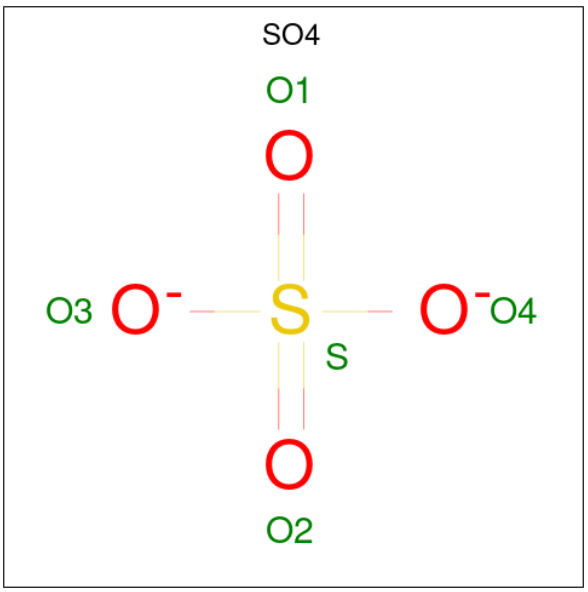
- Molecule 5 is METHOTREXATE (three-letter code: MTX) (formula: $C_{20}H_{22}N_8O_5$).

MTX



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	33	20	8	5	0	0
5	B	1	33	20	8	5	0	0
5	C	1	33	20	8	5	0	0
5	D	1	33	20	8	5	0	0
5	E	1	33	20	8	5	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
6	A	1	5	4	1	0	0
6	B	1	5	4	1	0	0
6	C	1	5	4	1	0	0
6	D	1	5	4	1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	55	55	55	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	27	Total O 27 27	0	0
7	C	28	Total O 28 28	0	0
7	D	25	Total O 25 25	0	0
7	E	4	Total O 4 4	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	214.77Å 115.79Å 220.14Å 90.00° 94.68° 90.00°	Depositor
Resolution (Å)	49.09 – 2.79	Depositor
% Data completeness (in resolution range)	99.5 (49.09-2.79)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.34 (at 2.81Å)	Xtriage
Refinement program	PHENIX (1.15.2_3472)	Depositor
R, R_{free}	0.222 , 0.253	Depositor
Wilson B-factor (Å^2)	67.6	Xtriage
Anisotropy	0.340	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	21040	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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.

4 Model quality (i)

4.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: OG7, MTX, SO4, UFP, NDP

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	0.25	0/4181	0.43	0/5660
1	B	0.25	0/4165	0.43	0/5643
1	C	0.25	0/4139	0.42	0/5613
1	D	0.25	0/4145	0.42	0/5621
1	E	0.24	0/4066	0.41	0/5533
All	All	0.25	0/20696	0.42	0/28070

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	4086	0	3973	44	0
1	B	4070	0	3938	46	0
1	C	4044	0	3896	42	0
1	D	4050	0	3897	39	0
1	E	3971	0	3707	27	0
2	A	48	0	26	1	0
2	B	48	0	26	1	0
2	C	48	0	26	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	48	0	26	0	0
2	E	48	0	26	1	0
3	A	21	0	10	2	0
3	B	21	0	10	2	0
3	C	21	0	10	1	0
3	D	21	0	10	1	0
3	E	21	0	10	1	0
4	A	30	0	0	0	0
4	B	30	0	0	1	0
4	C	30	0	0	0	0
4	D	30	0	0	0	0
4	E	30	0	0	0	0
5	A	33	0	20	0	0
5	B	33	0	20	1	0
5	C	33	0	20	2	0
5	D	33	0	20	2	0
5	E	33	0	20	1	0
6	A	5	0	0	0	0
6	B	5	0	0	0	0
6	C	5	0	0	0	0
6	D	5	0	0	0	0
7	A	55	0	0	0	0
7	B	27	0	0	4	0
7	C	28	0	0	1	0
7	D	25	0	0	1	0
7	E	4	0	0	0	0
All	All	21040	0	19691	181	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 4.

The worst 5 of 181 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:A:341:ILE:HG22	1:A:397:MET:HE3	1.66	0.78
1:A:225:ASN:O	1:A:233:ARG:NH2	2.20	0.74
1:C:333:ARG:HG3	1:C:337:ASP:HB3	1.72	0.72
1:C:207:PHE:HB3	1:C:210:ARG:HB2	1.73	0.71
1:A:333:ARG:HG3	1:A:337:ASP:HB3	1.73	0.69

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	503/521 (96%)	480 (95%)	23 (5%)	0	100	100
1	B	503/521 (96%)	478 (95%)	25 (5%)	0	100	100
1	C	504/521 (97%)	480 (95%)	24 (5%)	0	100	100
1	D	503/521 (96%)	480 (95%)	23 (5%)	0	100	100
1	E	504/521 (97%)	480 (95%)	24 (5%)	0	100	100
All	All	2517/2605 (97%)	2398 (95%)	119 (5%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	446/470 (95%)	443 (99%)	3 (1%)	84	95
1	B	442/470 (94%)	439 (99%)	3 (1%)	84	95
1	C	434/470 (92%)	431 (99%)	3 (1%)	84	95
1	D	437/470 (93%)	435 (100%)	2 (0%)	88	96
1	E	415/470 (88%)	414 (100%)	1 (0%)	93	98
All	All	2174/2350 (92%)	2162 (99%)	12 (1%)	86	96

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

Mol	Chain	Res	Type
1	C	100	ASN
1	C	235	HIS
1	E	235	HIS
1	D	104	ASP
1	B	99	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	126	ASN
1	A	345	GLN
1	B	345	GLN
1	C	345	GLN
1	D	345	GLN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

24 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

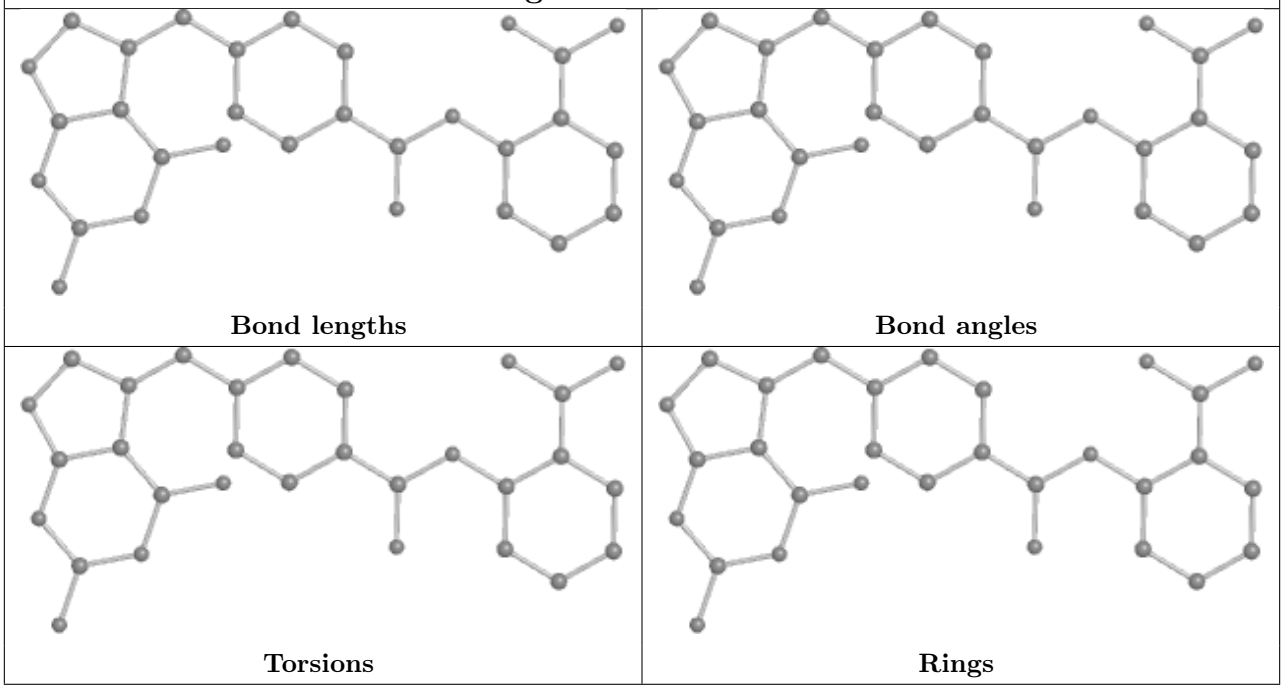
There are no torsion outliers.

There are no ring outliers.

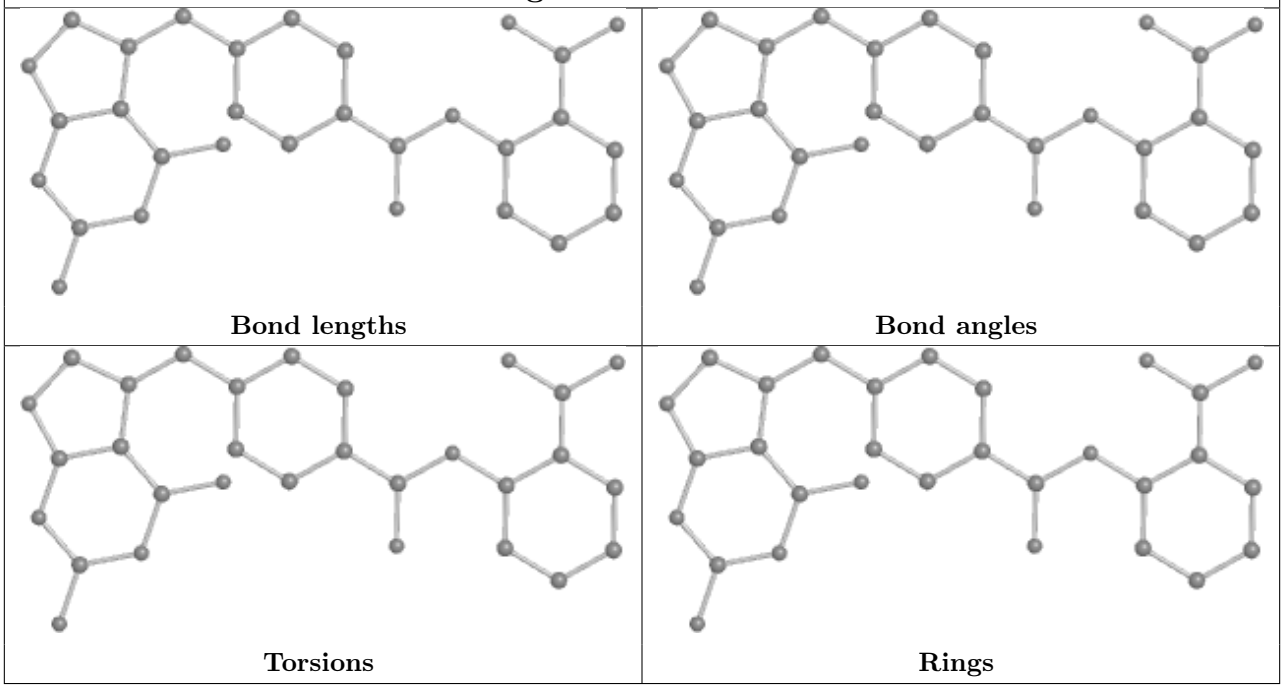
No monomer is involved in short contacts.

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.

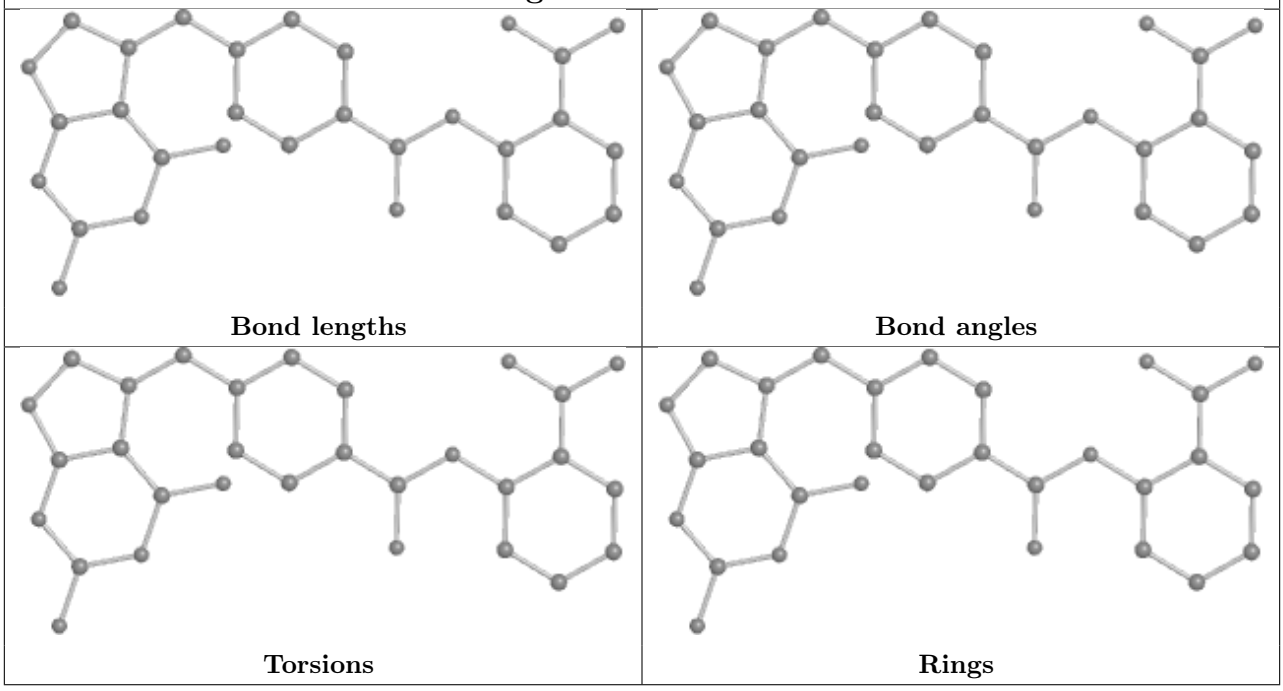
Ligand OG7 A 603



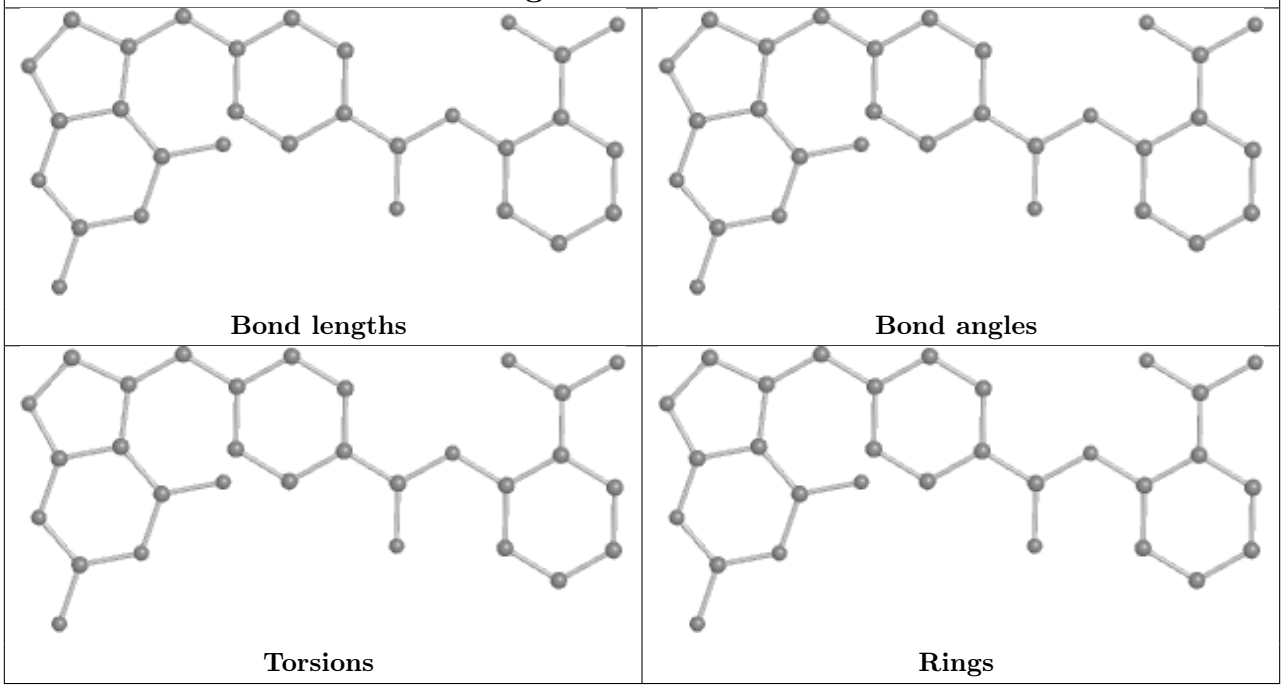
Ligand OG7 B 603



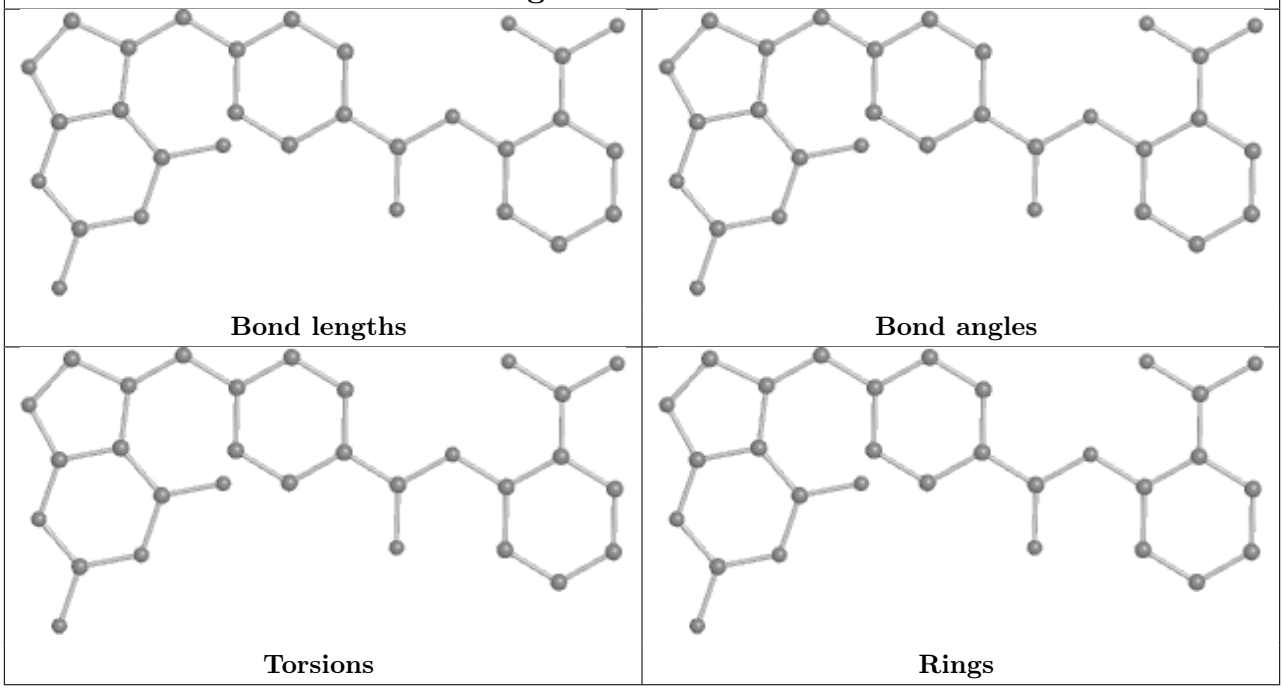
Ligand OG7 C 603



Ligand OG7 D 603



Ligand OG7 E 603



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

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

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

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

5.3 Carbohydrates (i)

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

5.4 Ligands (i)

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

5.5 Other polymers (i)

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