



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 18, 2021 – 08:08 am BST

PDB ID : 7AU5
Title : Tubulin-noscapine-analogue-14e complex
Authors : Yong, C.; Devine, S.M.; Abel, A.-C.; Muthiah, D.; Gao, X.; Callaghan, R.;
Capuano, B.; Steinmetz, M.O.; Protá, A.E.; Scammels, P.J.
Deposited on : 2020-11-02
Resolution : 2.20 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)
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
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) : 2.23.1

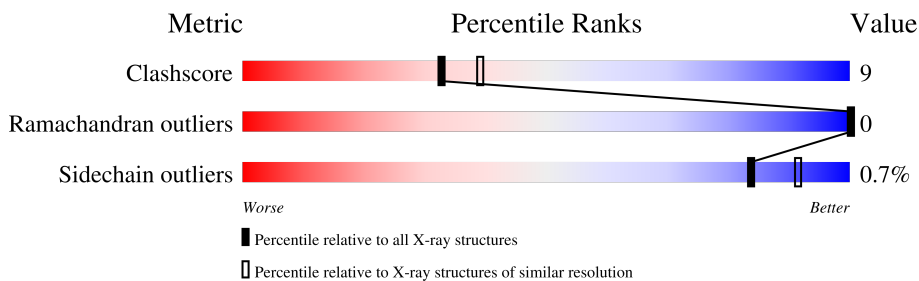
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	451	
1	C	451	
2	B	445	
2	D	445	
3	E	143	
4	F	384	

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 18596 atoms, of which 13 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 Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	438	Total	C	N	O	S	0	12	0
			3497	2214	593	665	25			
1	C	440	Total	C	N	O	S	0	8	0
			3499	2209	595	669	26			

- Molecule 2 is a protein called Tubulin beta-2B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	423	Total	C	N	O	S	0	9	0
			3401	2135	583	656	27			
2	D	426	Total	C	N	O	S	0	3	0
			3368	2112	578	650	28			

- Molecule 3 is a protein called Stathmin-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	124	Total	C	N	O	S	0	4	0
			1059	652	194	207	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	3	MET	-	expression tag	UNP P63043
E	4	ALA	-	expression tag	UNP P63043

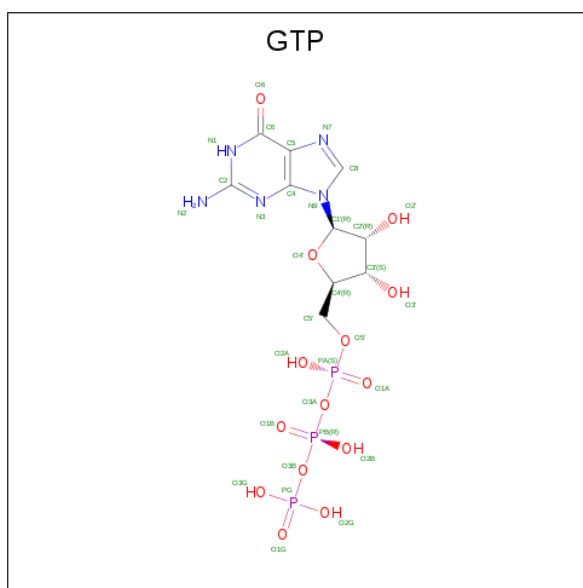
- Molecule 4 is a protein called Tubulin-Tyrosine Ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	F	348	Total	C	N	O	S	0	3	0
			2867	1837	492	523	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	379	HIS	-	expression tag	UNP E1BQ43
F	380	HIS	-	expression tag	UNP E1BQ43
F	381	HIS	-	expression tag	UNP E1BQ43
F	382	HIS	-	expression tag	UNP E1BQ43
F	383	HIS	-	expression tag	UNP E1BQ43
F	384	HIS	-	expression tag	UNP E1BQ43

- Molecule 5 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
5	C	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
6	A	1	Total	Mg	0	0
			1	1		
6	B	1	Total	Mg	0	0
			1	1		
6	C	1	Total	Mg	0	0
			1	1		
6	D	1	Total	Mg	0	0
			1	1		

Continued on next page...

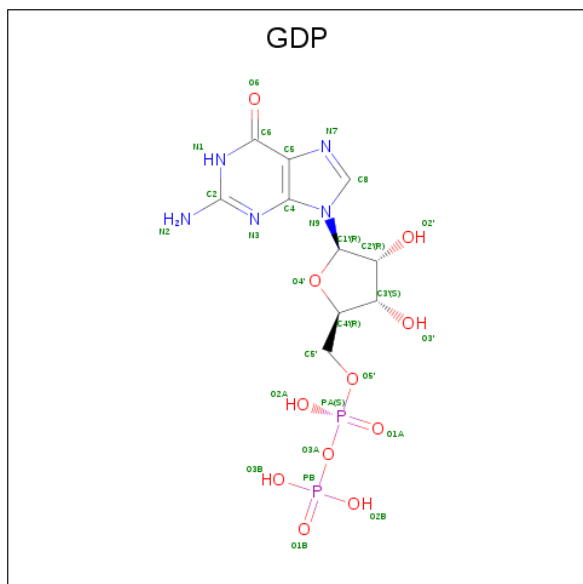
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	F	1	Total Mg 1 1	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

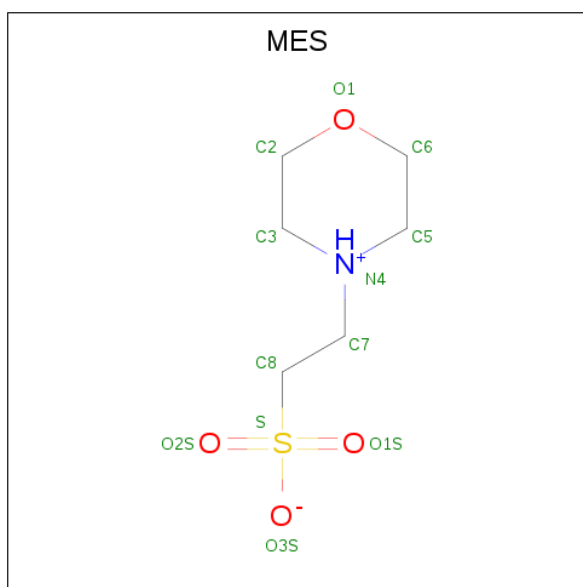
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Ca 1 1	0	0
7	B	2	Total Ca 2 2	0	0
7	C	1	Total Ca 1 1	0	0
7	E	1	Total Ca 1 1	0	0

- Molecule 8 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C₁₀H₁₅N₅O₁₁P₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C N O P 28 10 5 11 2	0	0
8	D	1	Total C N O P 28 10 5 11 2	0	0

- Molecule 9 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).



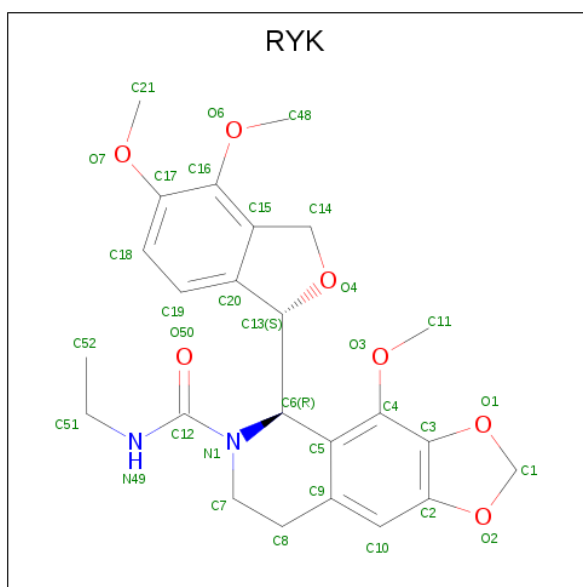
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
9	B	1	12	6	1	4	1	0	0

- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



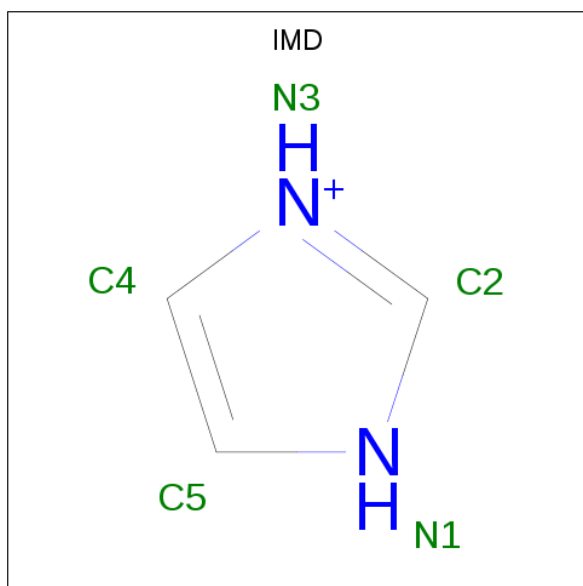
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
10	B	1	14	3	8	3	0	0

- Molecule 11 is (5 {R})-5-[(1 {S})-4,5-dimethoxy-1,3-dihydro-2-benzofuran-1-yl]- {N}-ethyl-4-methoxy-7,8-dihydro-5 {H}-[1,3]dioxolo[4,5-g]isoquinoline-6-carboxamide (three-letter code: RYK) (formula: C₂₄H₂₈N₂O₇) (labeled as "Ligand of Interest" by depositor).



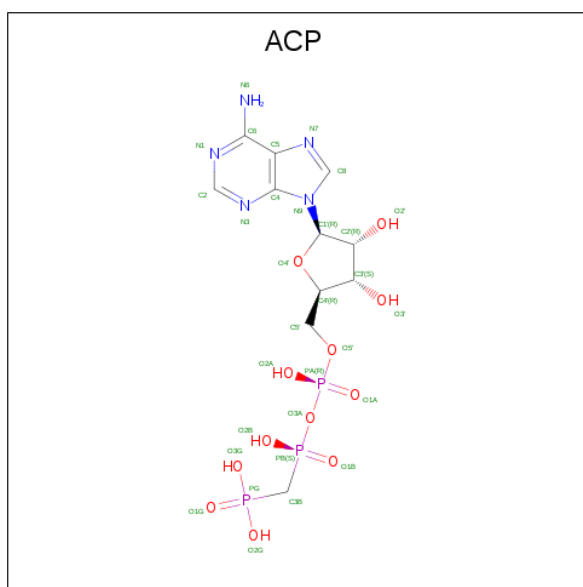
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
11	B	1	33	24	2	7	0	0

- Molecule 12 is IMIDAZOLE (three-letter code: IMD) (formula: $C_3H_5N_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	N		
12	C	1	10	3	5	2	0	0

- Molecule 13 is PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER (three-letter code: ACP) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
13	F	1	31	11	5	12	3	0	0

- Molecule 14 is water.

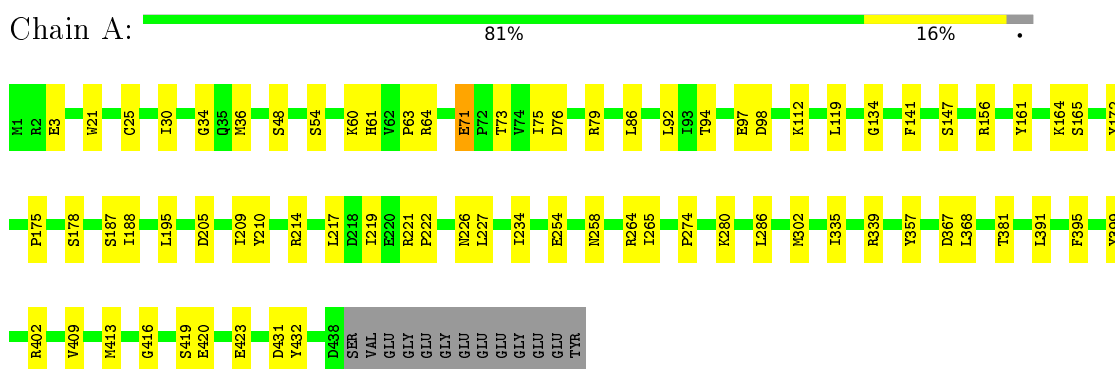
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	122	Total	O	0	0
			122	122		
14	B	150	Total	O	0	0
			150	150		
14	C	278	Total	O	0	0
			278	278		
14	D	62	Total	O	0	0
			62	62		
14	E	23	Total	O	0	0
			23	23		
14	F	40	Total	O	0	0
			40	40		

3 Residue-property plots [i](#)

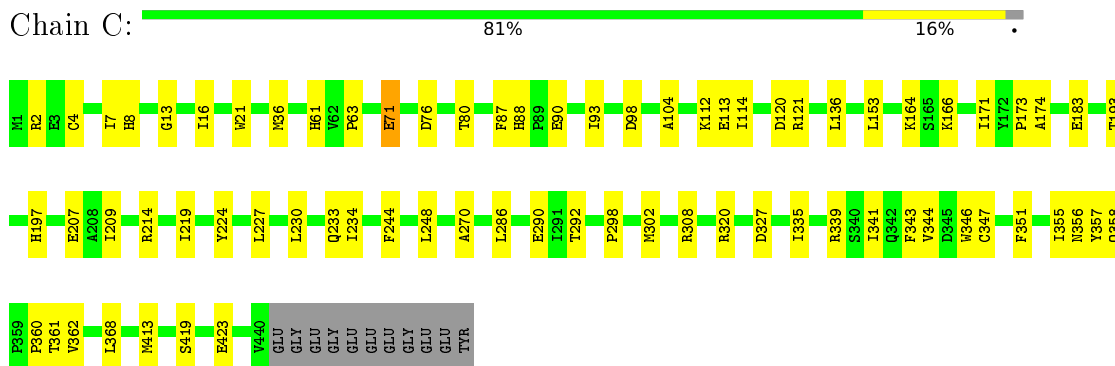
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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 was not executed.

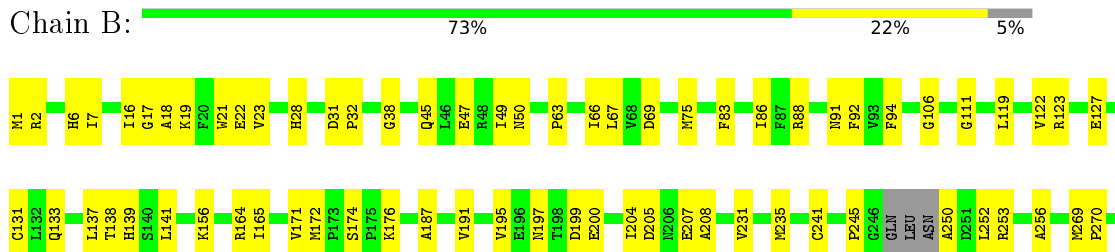
- Molecule 1: Tubulin alpha-1B chain

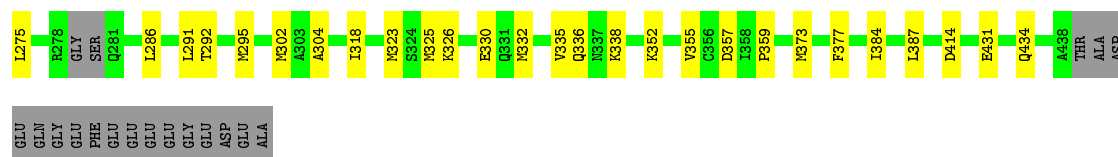


- Molecule 1: Tubulin alpha-1B chain



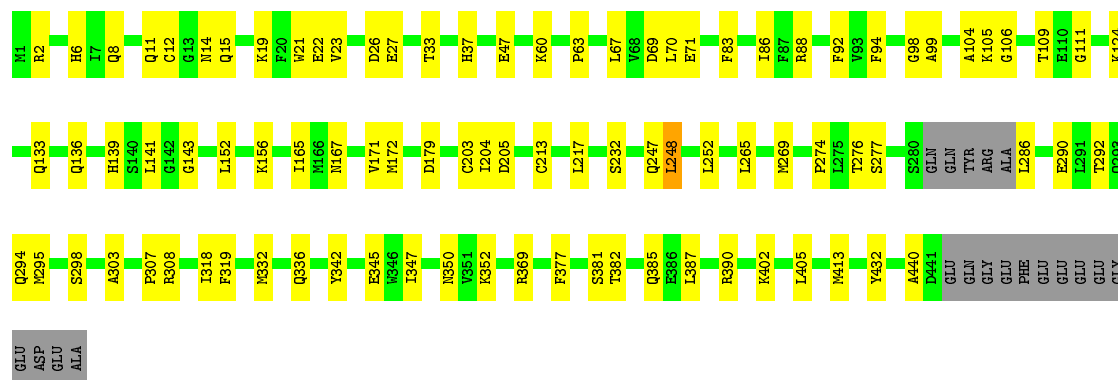
- Molecule 2: Tubulin beta-2B chain





- Molecule 2: Tubulin beta-2B chain

Chain D: 75% 20%



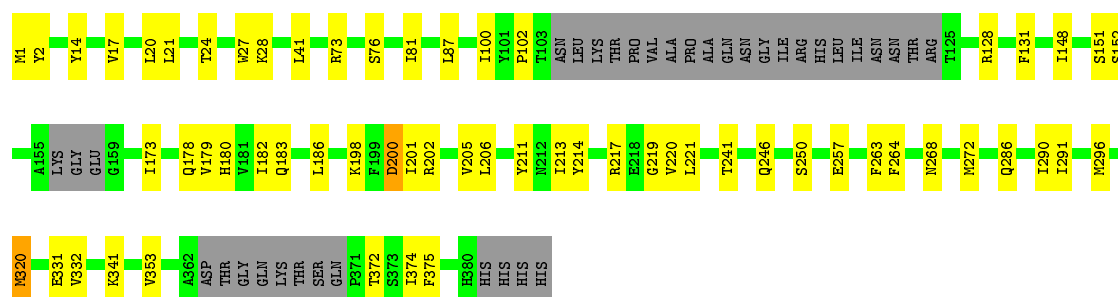
- Molecule 3: Stathmin-4

Chain E: 73% 14% 13%



- Molecule 4: Tubulin-Tyrosine Ligase

Chain F: 75% 15% 9%



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	104.64Å 157.84Å 180.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.88 – 2.20	Depositor
% Data completeness (in resolution range)	98.6 (47.88-2.20)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PHENIX 1.18.2_3874, REFMAC 5.8.0258	Depositor
R, R_{free}	0.178 , 0.212	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	18596	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

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: IMD, GDP, GOL, CA, GTP, ACP, MG, RYK, MES

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	0/3596	0.42	0/4880
1	C	0.26	0/3581	0.43	0/4864
2	B	0.25	0/3486	0.42	0/4717
2	D	0.25	0/3444	0.41	0/4662
3	E	0.24	0/1068	0.35	0/1417
4	F	0.24	0/2932	0.40	0/3960
All	All	0.25	0/18107	0.41	0/24500

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	3497	0	3421	55	0
1	C	3499	0	3396	56	0
2	B	3401	0	3282	72	0
2	D	3368	0	3251	60	0
3	E	1059	0	1071	15	0
4	F	2867	0	2836	51	0
5	A	32	0	12	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	32	0	12	1	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
7	A	1	0	0	0	0
7	B	2	0	0	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
8	B	28	0	12	0	0
8	D	28	0	12	3	0
9	B	12	0	12	3	0
10	B	6	8	8	0	0
11	B	33	0	0	3	0
12	C	5	5	5	0	0
13	F	31	0	14	5	0
14	A	122	0	0	1	0
14	B	150	0	0	2	0
14	C	278	0	0	7	0
14	D	62	0	0	4	0
14	E	23	0	0	3	0
14	F	40	0	0	1	0
All	All	18583	13	17344	308	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 9.

The worst 5 of 308 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)
2:D:248:LEU:HD11	2:D:352:LYS:HB3	1.59	0.84
1:C:335:ILE:HG23	1:C:339:ARG:HG3	1.61	0.82
4:F:200:ASP:HB2	4:F:241:THR:HG21	1.60	0.82
2:B:16[B]:ILE:HD13	2:B:231:VAL:HG11	1.63	0.79
2:B:47:GLU:HG2	2:B:245:PRO:HG3	1.65	0.77

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	448/451 (99%)	438 (98%)	10 (2%)	0	100	100
1	C	447/451 (99%)	436 (98%)	11 (2%)	0	100	100
2	B	426/445 (96%)	417 (98%)	9 (2%)	0	100	100
2	D	425/445 (96%)	415 (98%)	10 (2%)	0	100	100
3	E	124/143 (87%)	123 (99%)	1 (1%)	0	100	100
4	F	343/384 (89%)	334 (97%)	9 (3%)	0	100	100
All	All	2213/2319 (95%)	2163 (98%)	50 (2%)	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	381/379 (100%)	379 (100%)	2 (0%)	88	94
1	C	380/379 (100%)	379 (100%)	1 (0%)	92	97
2	B	375/383 (98%)	372 (99%)	3 (1%)	81	90
2	D	371/383 (97%)	367 (99%)	4 (1%)	73	85
3	E	115/127 (91%)	114 (99%)	1 (1%)	78	88
4	F	314/342 (92%)	311 (99%)	3 (1%)	76	86
All	All	1936/1993 (97%)	1922 (99%)	14 (1%)	84	91

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	139	HIS
2	D	179	ASP
4	F	320	MET
4	F	200	ASP
4	F	257	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	11	GLN
2	B	281	GLN
4	F	178	GLN
4	F	348	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 10 are monoatomic - leaving 9 for Mogul analysis.

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
5	GTP	A	501	6	26,34,34	0.95	1 (3%)	33,54,54	1.75	6 (18%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	GTP	C	501	6	26,34,34	0.93	1 (3%)	33,54,54	1.74	6 (18%)
12	IMD	C	504	-	3,5,5	0.40	0	4,5,5	0.57	0
13	ACP	F	401	6	27,33,33	2.02	7 (25%)	32,52,52	1.27	4 (12%)
8	GDP	D	501	6	24,30,30	1.19	2 (8%)	31,47,47	1.92	7 (22%)
10	GOL	B	506	-	5,5,5	0.81	0	5,5,5	0.90	0
8	GDP	B	501	6	24,30,30	1.14	2 (8%)	31,47,47	1.89	7 (22%)
9	MES	B	504	-	12,12,12	2.19	1 (8%)	14,16,16	1.83	4 (28%)
11	RYK	B	507	-	37,37,37	1.49	4 (10%)	53,54,54	1.64	11 (20%)

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
5	GTP	A	501	6	-	9/18/38/38	0/3/3/3
5	GTP	C	501	6	-	7/18/38/38	0/3/3/3
12	IMD	C	504	-	-	-	0/1/1/1
13	ACP	F	401	6	-	3/15/38/38	0/3/3/3
8	GDP	D	501	6	-	6/12/32/32	0/3/3/3
10	GOL	B	506	-	-	0/4/4/4	-
8	GDP	B	501	6	-	5/12/32/32	0/3/3/3
9	MES	B	504	-	-	0/6/14/14	0/1/1/1
11	RYK	B	507	-	-	5/17/45/45	0/5/5/5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	504	MES	C8-S	-7.32	1.67	1.77
11	B	507	RYK	C14-C15	-5.93	1.44	1.50
13	F	401	ACP	PG-O1G	5.47	1.61	1.50
8	D	501	GDP	C6-C5	4.24	1.48	1.41
13	F	401	ACP	PB-O1B	4.13	1.61	1.51

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	501	GTP	N3-C2-N1	-5.41	120.01	127.22
5	C	501	GTP	N3-C2-N1	-5.39	120.03	127.22

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	B	501	GDP	C2-N3-C4	4.94	121.00	115.36
8	D	501	GDP	C2-N3-C4	4.80	120.83	115.36
8	B	501	GDP	C6-C5-C4	-4.40	116.60	120.80

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

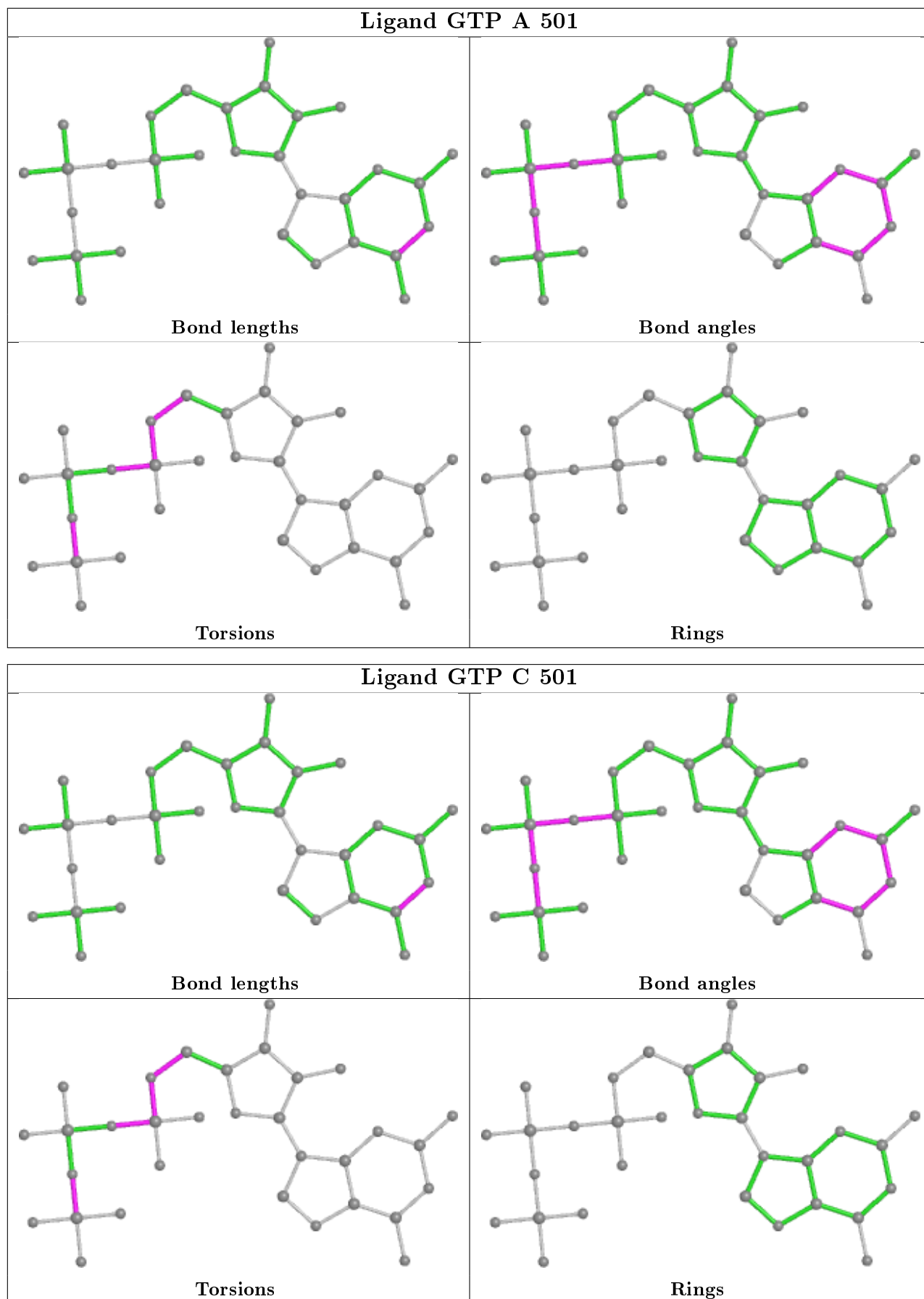
Mol	Chain	Res	Type	Atoms
5	A	501	GTP	C5'-O5'-PA-O1A
5	A	501	GTP	C5'-O5'-PA-O2A
5	C	501	GTP	PB-O3B-PG-O2G
5	C	501	GTP	PB-O3B-PG-O3G
5	C	501	GTP	C5'-O5'-PA-O1A

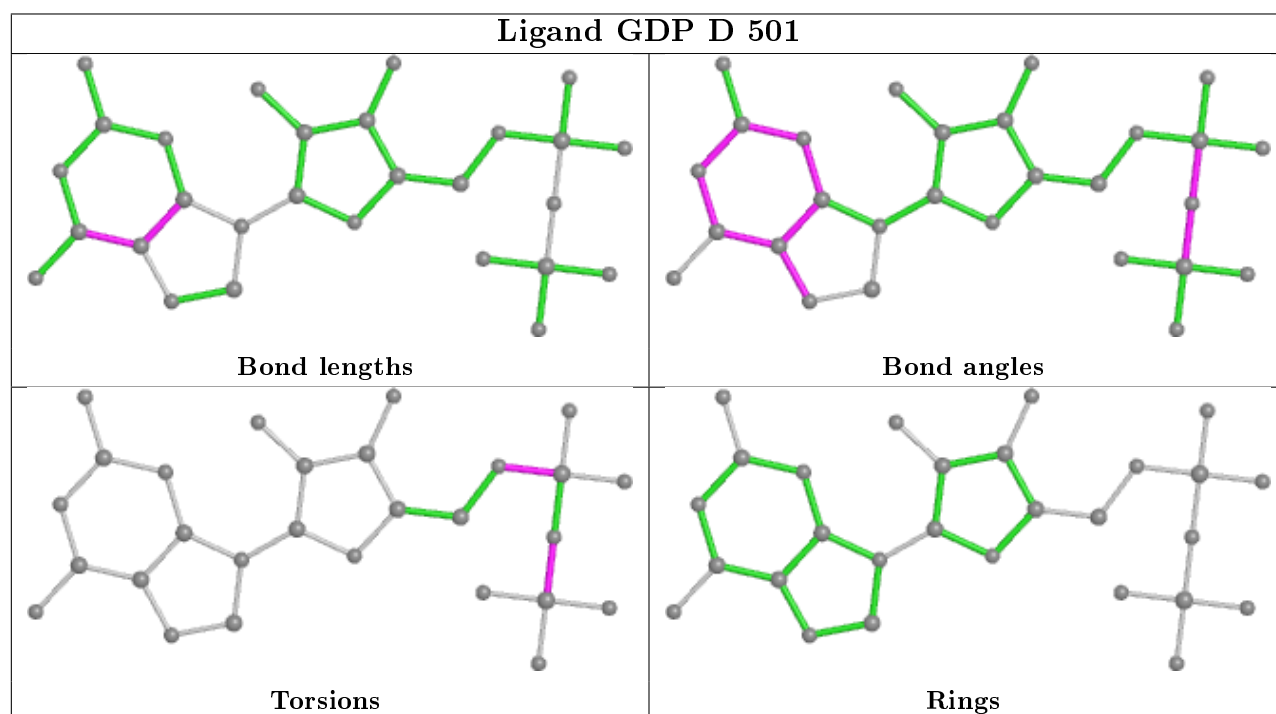
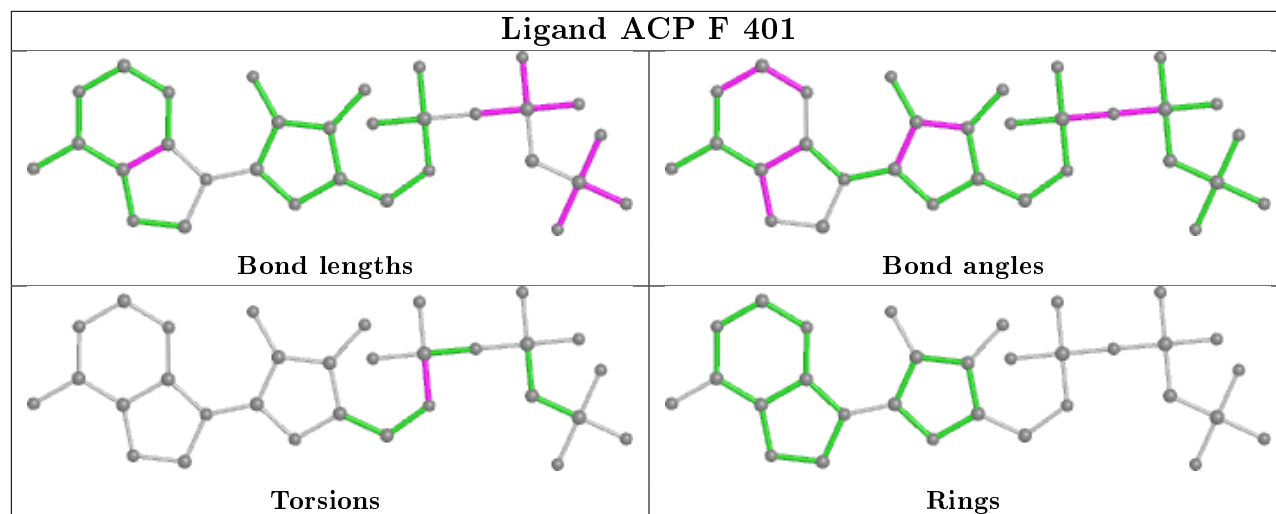
There are no ring outliers.

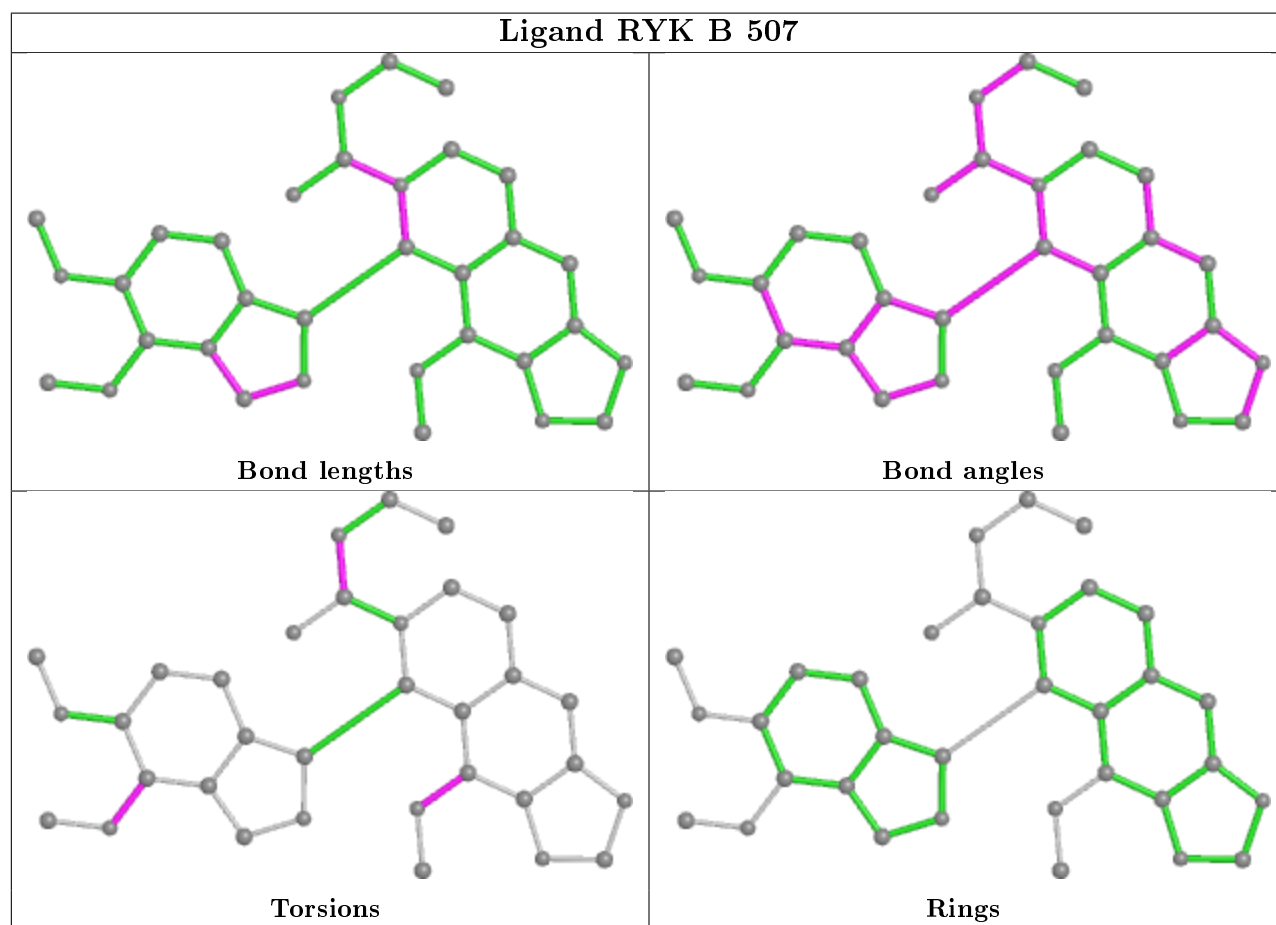
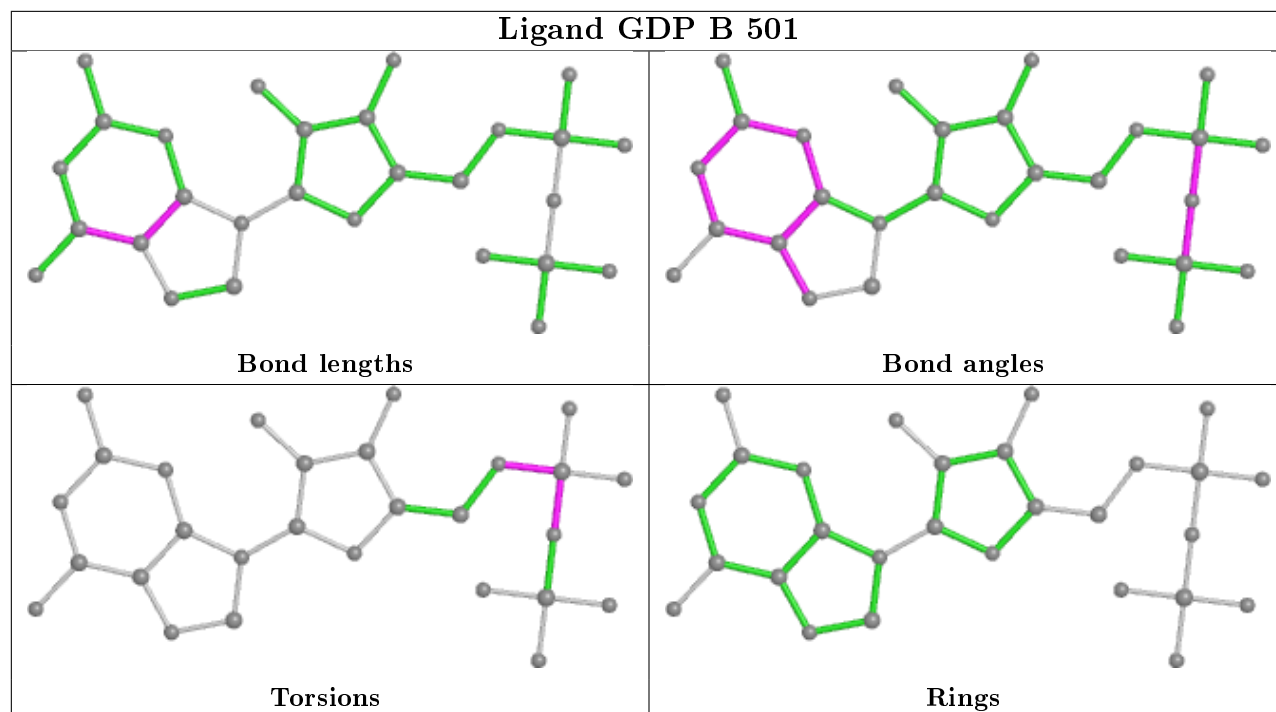
6 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	501	GTP	1	0
5	C	501	GTP	1	0
13	F	401	ACP	5	0
8	D	501	GDP	3	0
9	B	504	MES	3	0
11	B	507	RYK	3	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 was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.