



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 11:50 AM BST

PDB ID : 1GZ5  
Title : Trehalose-6-phosphate synthase. OtsA  
Authors : Gibson, R.P.; Turkenburg, J.P.; Davies, G.J.  
Deposited on : 2002-05-15  
Resolution : 2.43 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](mailto: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) : 2.13.1

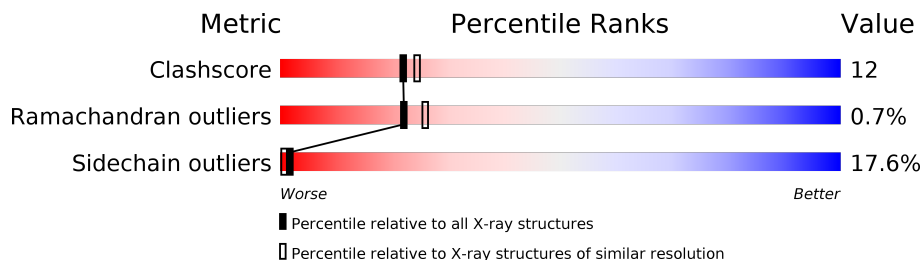
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.43 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	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	456	70% 22% 6% •
1	B	456	68% 25% 5% •
1	C	456	67% 25% 8% •
1	D	456	67% 25% 6% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	IMD	A	902	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	IMD	C	902	-	-	X	-

## 2 Entry composition [i](#)

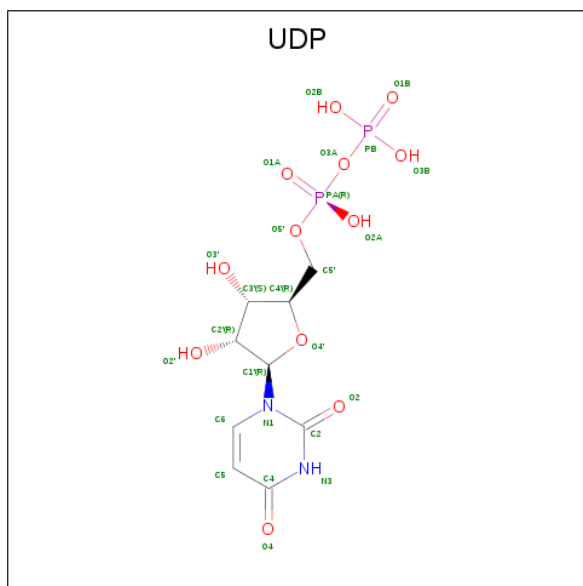
There are 5 unique types of molecules in this entry. The entry contains 15078 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 ALPHA-TREHALOSE-PHOSPHATE SYNTHASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	456	Total	C	N	O	S	Se	0	0	0
			3654	2346	637	664	3	4			
1	B	455	Total	C	N	O	S	Se	0	0	0
			3649	2343	636	663	3	4			
1	C	456	Total	C	N	O	S	Se	0	0	0
			3654	2346	637	664	3	4			
1	D	455	Total	C	N	O	S	Se	0	0	0
			3649	2343	636	663	3	4			

- Molecule 2 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C<sub>9</sub>H<sub>14</sub>N<sub>2</sub>O<sub>12</sub>P<sub>2</sub>).



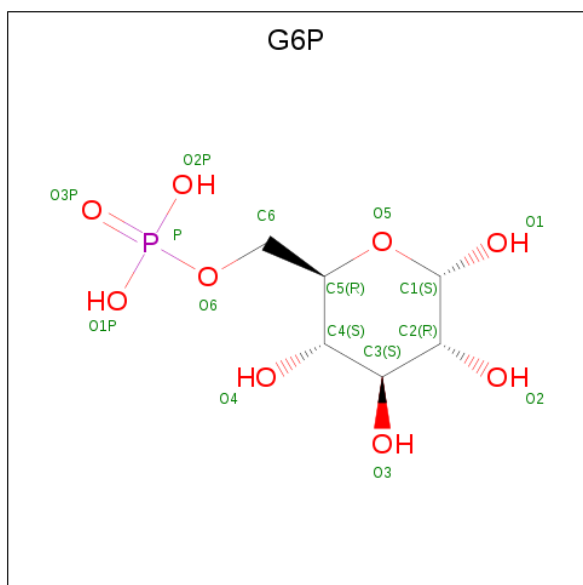
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

*Continued on next page...*

Continued from previous page...

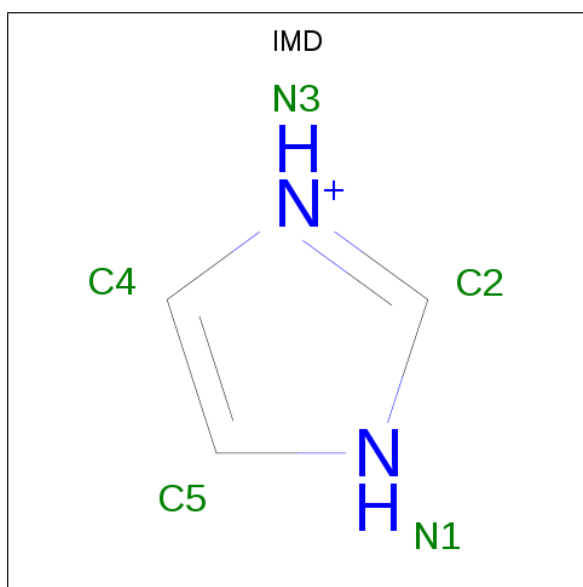
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	D	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 3 is 6-O-phosphono-alpha-D-glucopyranose (three-letter code: G6P) (formula:  $C_6H_{13}O_9P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	P	0	0
			16	6	9	1		
3	B	1	Total	C	O	P	0	0
			16	6	9	1		
3	C	1	Total	C	O	P	0	0
			16	6	9	1		
3	D	1	Total	C	O	P	0	0
			16	6	9	1		

- Molecule 4 is IMIDAZOLE (three-letter code: IMD) (formula:  $C_3H_5N_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	N	0	0
			5	3	2		
4	B	1	Total	C	N	0	0
			5	3	2		
4	C	1	Total	C	N	0	0
			5	3	2		
4	D	1	Total	C	N	0	0
			5	3	2		

- Molecule 5 is water.

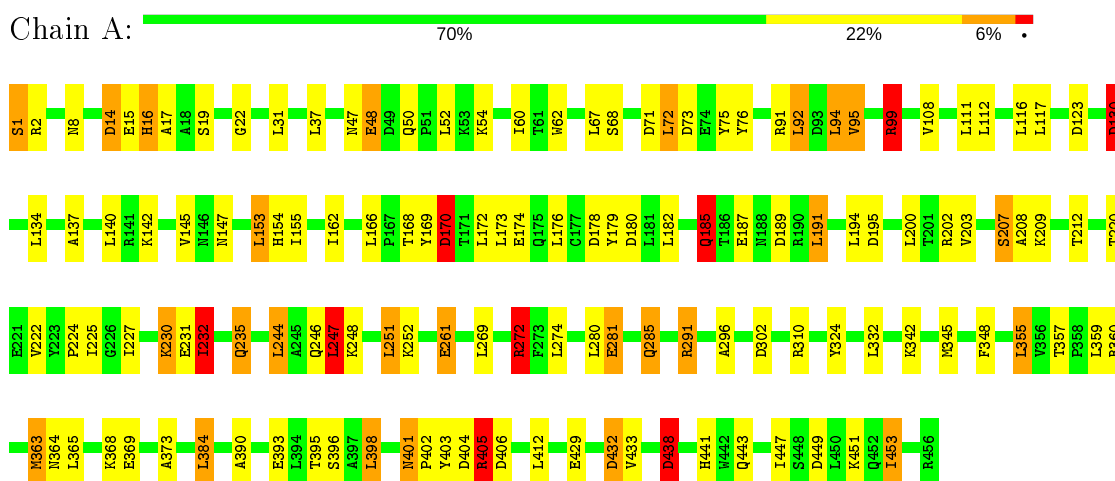
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	74	Total	O	0	0
			74	74		
5	B	74	Total	O	0	0
			74	74		
5	C	72	Total	O	0	0
			72	72		
5	D	68	Total	O	0	0
			68	68		

### 3 Residue-property plots [i](#)

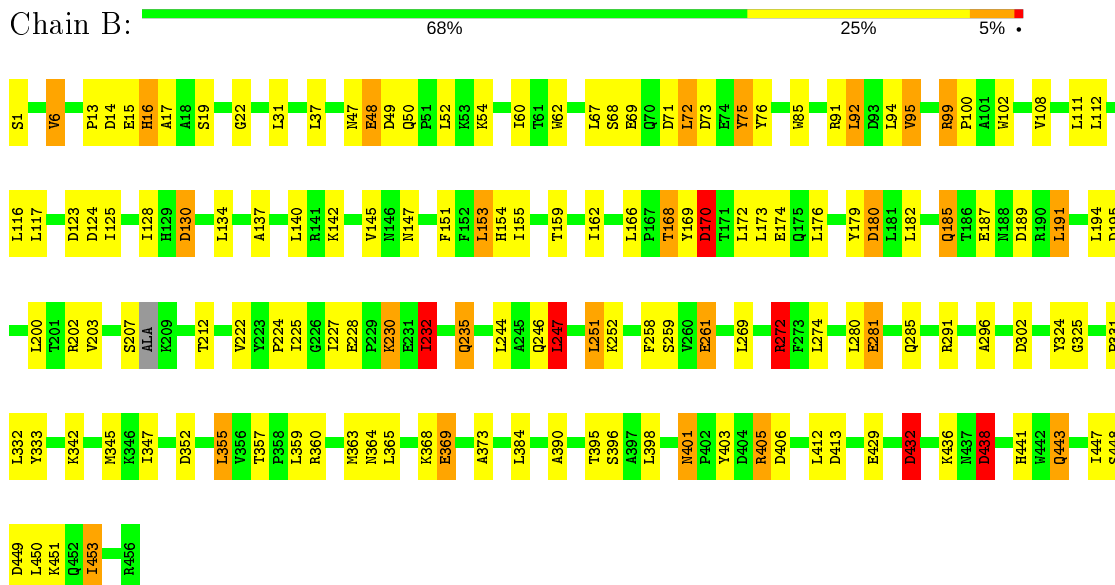
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded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

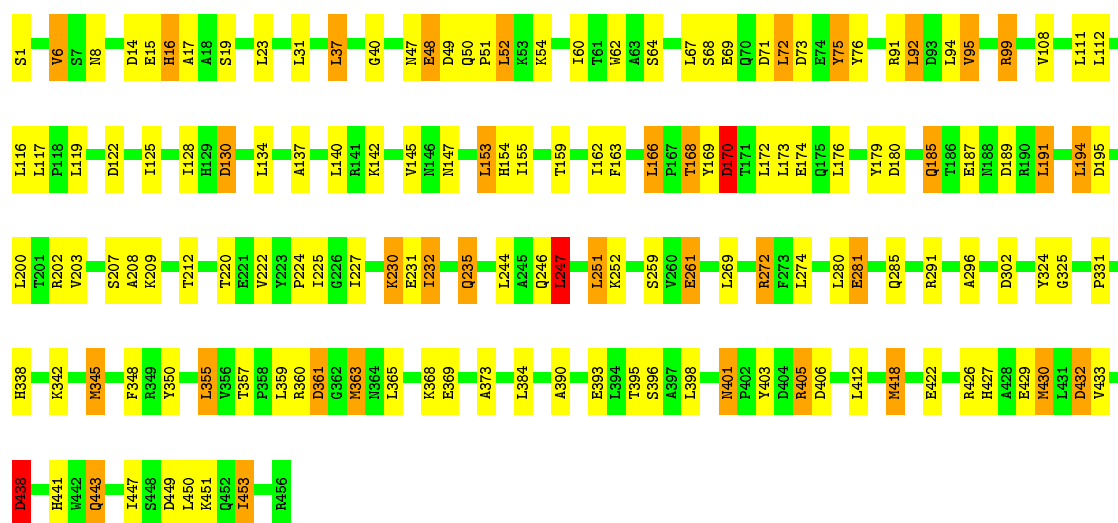


- Molecule 1: ALPHA-TREHALOSE-PHOSPHATE SYNTHASE



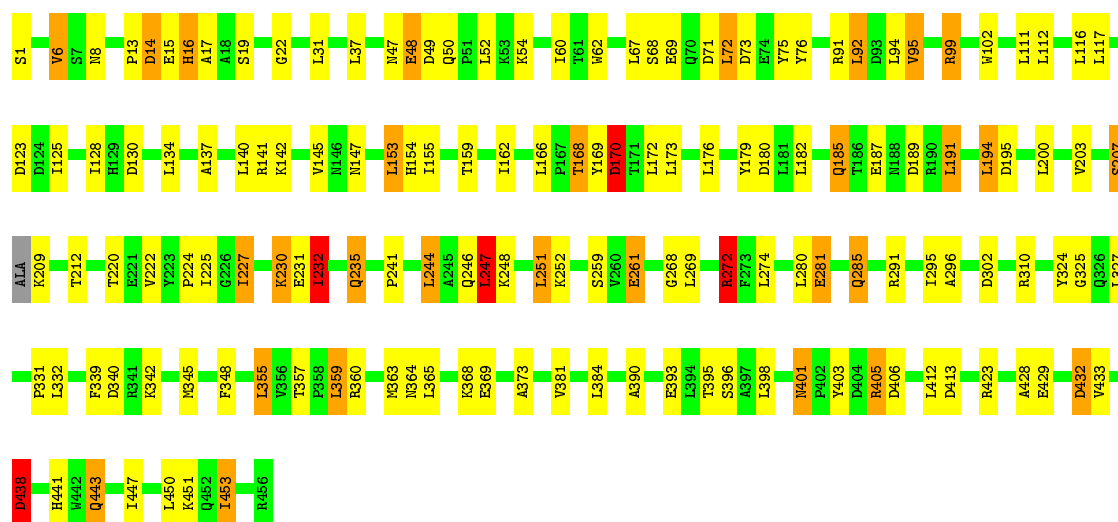
- Molecule 1: ALPHA-TREHALOSE-PHOSPHATE SYNTHASE

Chain C:  67% 25% 8%



• Molecule 1: ALPHA-TREHALOSE-PHOSPHATE SYNTHASE

Chain D:  67% 25% 6%





## 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, $\alpha$ , $\beta$ , $\gamma$	103.61Å 125.40Å 176.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 2.43	Depositor
% Data completeness (in resolution range)	99.3 (100.00-2.43)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.1.19	Depositor
R, $R_{free}$	0.206 , 0.228	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	15078	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	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: UDP, G6P, IMD

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	0.98	5/3743 (0.1%)	1.07	27/5083 (0.5%)
1	B	0.86	4/3737 (0.1%)	1.01	20/5073 (0.4%)
1	C	0.75	4/3743 (0.1%)	0.94	20/5083 (0.4%)
1	D	0.79	1/3737 (0.0%)	0.96	21/5073 (0.4%)
All	All	0.85	14/14960 (0.1%)	1.00	88/20312 (0.4%)

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	363	MSE	CG-SE	14.26	2.44	1.95
1	C	430	MSE	CG-SE	-11.03	1.57	1.95
1	C	418	MSE	SE-CE	9.46	2.51	1.95
1	C	345	MSE	CG-SE	8.43	2.24	1.95
1	A	170	ASP	CB-CG	-7.52	1.35	1.51

The worst 5 of 88 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	170	ASP	CB-CG-OD2	10.52	127.76	118.30
1	A	170	ASP	CB-CG-OD1	-10.18	109.14	118.30
1	B	91	ARG	NE-CZ-NH2	-9.76	115.42	120.30
1	A	406	ASP	CB-CG-OD2	9.76	127.08	118.30
1	B	170	ASP	CB-CG-OD2	9.39	126.75	118.30

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	3654	0	3624	75	0
1	B	3649	0	3618	91	0
1	C	3654	0	3624	93	0
1	D	3649	0	3618	85	0
2	A	25	0	11	1	0
2	B	25	0	11	0	0
2	C	25	0	11	1	0
2	D	25	0	11	0	0
3	A	16	0	10	0	0
3	B	16	0	10	0	0
3	C	16	0	11	4	0
3	D	16	0	11	0	0
4	A	5	0	5	4	0
4	B	5	0	5	3	0
4	C	5	0	5	4	0
4	D	5	0	5	2	0
5	A	74	0	0	2	0
5	B	74	0	0	6	0
5	C	72	0	0	13	0
5	D	68	0	0	2	0
All	All	15078	0	14590	348	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 12.

The worst 5 of 348 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:232:ILE:CD1	1:B:232:ILE:CG1	1.75	1.62
1:C:345:MSE:SE	1:C:345:MSE:CG	2.24	1.36
1:C:363:MSE:SE	1:C:363:MSE:CG	2.43	1.15
1:B:363:MSE:HE2	1:B:368:LYS:HE3	1.24	1.13
1:C:363:MSE:HE2	1:C:368:LYS:HE3	1.31	1.09

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	454/456 (100%)	437 (96%)	13 (3%)	4 (1%)	17	20
1	B	451/456 (99%)	437 (97%)	11 (2%)	3 (1%)	22	26
1	C	454/456 (100%)	438 (96%)	13 (3%)	3 (1%)	22	26
1	D	451/456 (99%)	433 (96%)	15 (3%)	3 (1%)	22	26
All	All	1810/1824 (99%)	1745 (96%)	52 (3%)	13 (1%)	22	26

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	438	ASP
1	B	438	ASP
1	C	438	ASP
1	D	438	ASP
1	A	364	ASN

### 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	384/380 (101%)	319 (83%)	65 (17%)	2	1
1	B	384/380 (101%)	315 (82%)	69 (18%)	1	1
1	C	384/380 (101%)	314 (82%)	70 (18%)	1	0
1	D	384/380 (101%)	317 (83%)	67 (17%)	2	1
All	All	1536/1520 (101%)	1265 (82%)	271 (18%)	2	1

5 of 271 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	359	LEU
1	C	112	LEU
1	D	261	GLU
1	B	398	LEU
1	C	16	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	304	GLN
1	C	8	ASN
1	D	304	GLN
1	B	401	ASN
1	C	16	HIS

### 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](#)

12 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
3	G6P	A	901	-	16,16,16	0.90	0	24,24,24	1.43	5 (20%)
2	UDP	C	900	-	20,26,26	1.32	1 (5%)	25,40,40	1.58	3 (12%)
2	UDP	A	900	-	20,26,26	1.68	3 (15%)	25,40,40	0.97	0
2	UDP	D	900	-	20,26,26	1.10	1 (5%)	25,40,40	1.69	6 (24%)
3	G6P	D	901	-	16,16,16	0.83	0	24,24,24	1.35	3 (12%)
3	G6P	C	901	-	16,16,16	0.88	0	24,24,24	1.91	8 (33%)
3	G6P	B	901	-	16,16,16	0.94	0	24,24,24	1.59	6 (25%)
4	IMD	C	902	-	3,5,5	0.45	0	4,5,5	0.54	0
4	IMD	D	902	-	3,5,5	0.43	0	4,5,5	0.74	0
2	UDP	B	900	-	20,26,26	1.43	2 (10%)	25,40,40	1.44	5 (20%)
4	IMD	B	902	-	3,5,5	0.63	0	4,5,5	0.17	0
4	IMD	A	902	-	3,5,5	0.66	0	4,5,5	0.14	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	G6P	A	901	-	-	2/6/26/26	0/1/1/1
2	UDP	C	900	-	-	7/14/32/32	0/2/2/2
2	UDP	A	900	-	-	3/14/32/32	0/2/2/2
2	UDP	D	900	-	-	4/14/32/32	0/2/2/2
3	G6P	D	901	-	-	1/6/26/26	0/1/1/1
3	G6P	C	901	-	-	4/6/26/26	0/1/1/1
3	G6P	B	901	-	-	2/6/26/26	0/1/1/1
4	IMD	C	902	-	-	-	0/1/1/1
4	IMD	D	902	-	-	-	0/1/1/1
2	UDP	B	900	-	-	4/14/32/32	0/2/2/2
4	IMD	B	902	-	-	-	0/1/1/1
4	IMD	A	902	-	-	-	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	900	UDP	C2'-C1'	-4.31	1.47	1.53
2	C	900	UDP	C4-N3	4.25	1.40	1.33
2	A	900	UDP	C4-N3	3.68	1.39	1.33
2	B	900	UDP	PB-O2B	-2.58	1.44	1.54

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	900	UDP	C4-N3	2.55	1.37	1.33

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	901	G6P	C1-O5-C5	4.88	122.86	113.66
2	C	900	UDP	O2B-PB-O3A	4.05	118.23	104.64
2	C	900	UDP	PA-O3A-PB	-3.98	119.18	132.83
2	D	900	UDP	O4'-C1'-C2'	-3.98	101.12	106.93
3	D	901	G6P	O2P-P-O1P	3.84	122.33	107.64

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	900	UDP	C2'-C1'-N1-C6
2	C	900	UDP	O4'-C1'-N1-C6
2	A	900	UDP	PB-O3A-PA-O5'
2	D	900	UDP	PB-O3A-PA-O5'
3	C	901	G6P	C4-C5-C6-O6

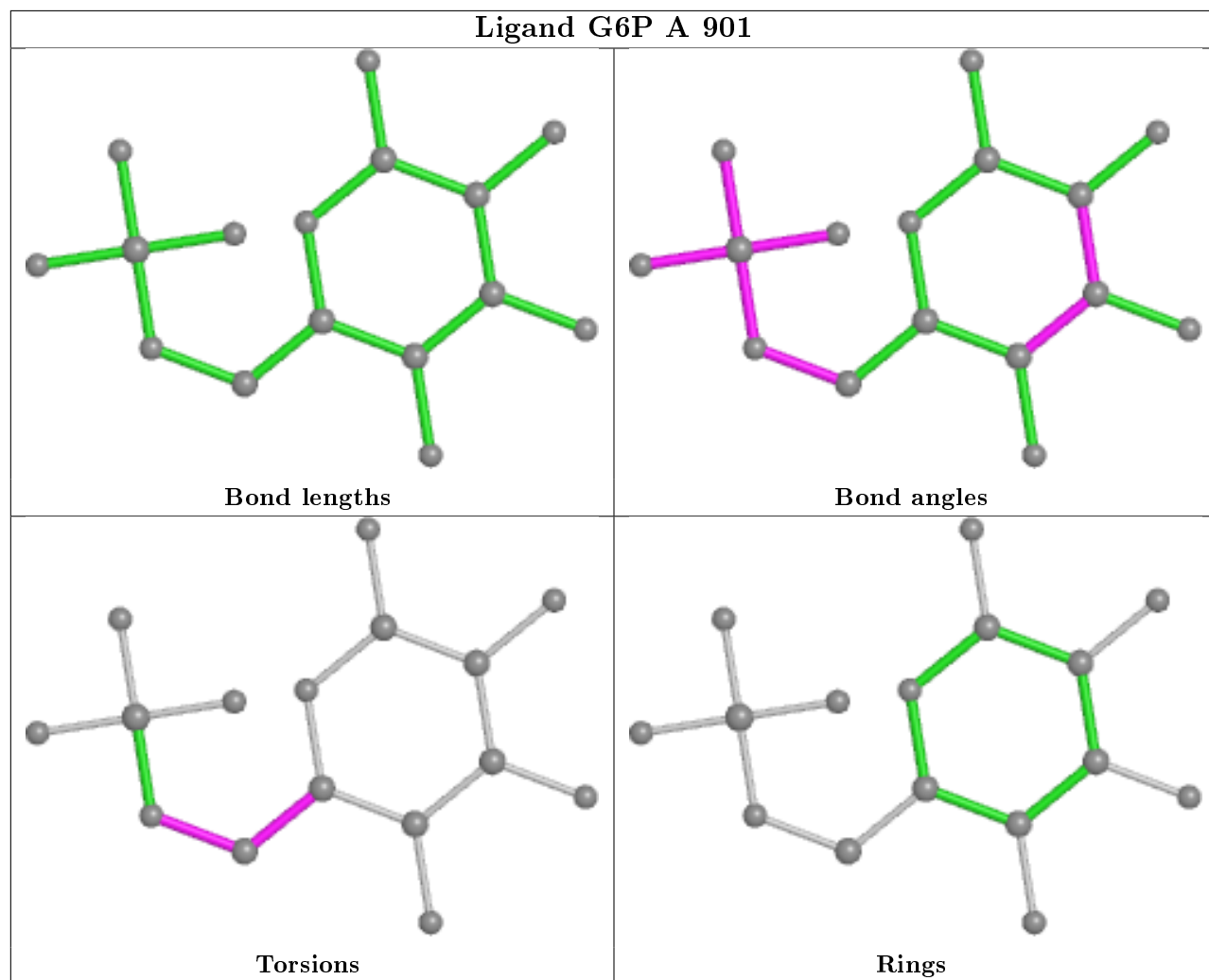
There are no ring outliers.

7 monomers are involved in 17 short contacts:

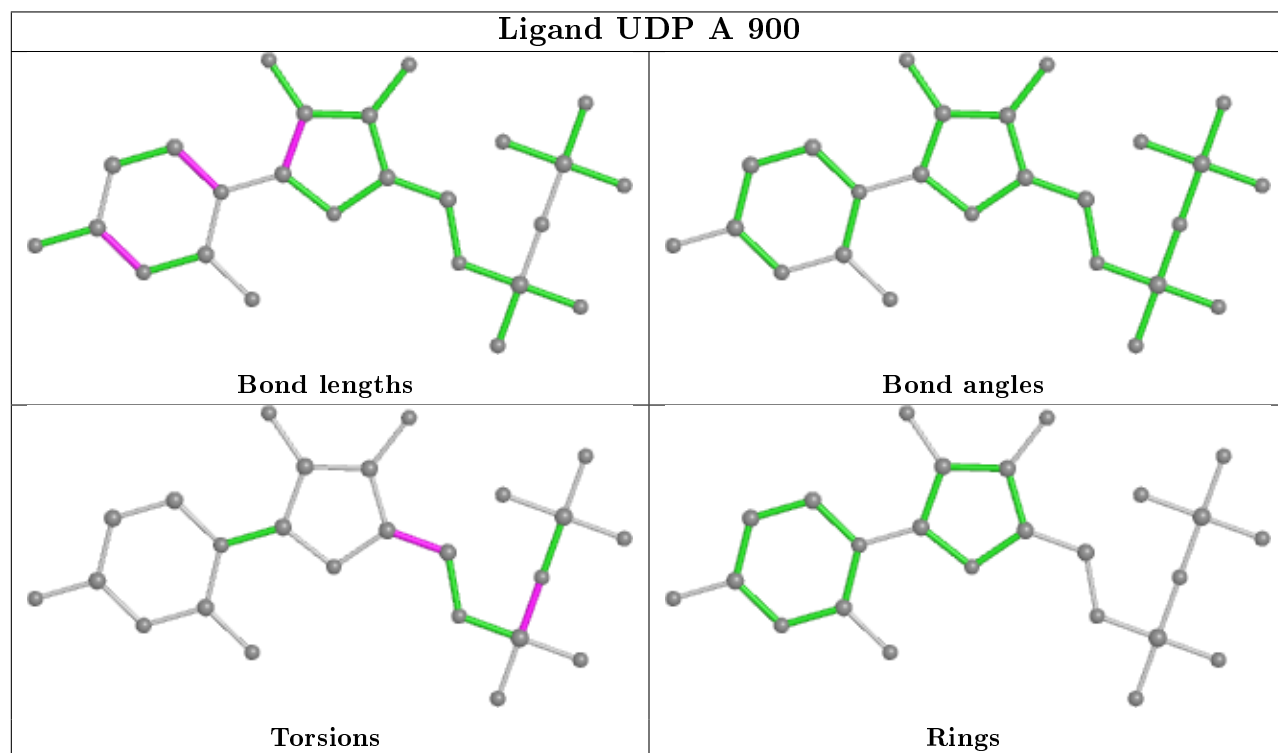
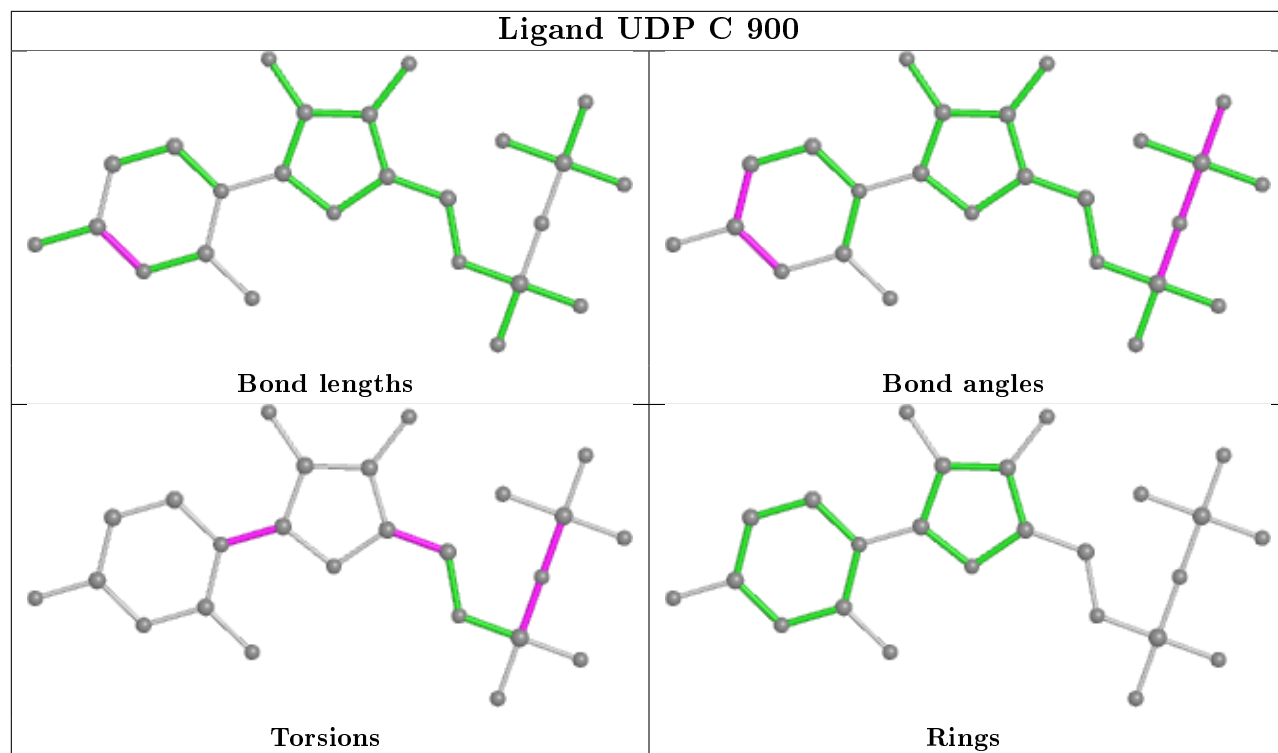
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	900	UDP	1	0
2	A	900	UDP	1	0
3	C	901	G6P	4	0
4	C	902	IMD	4	0
4	D	902	IMD	2	0
4	B	902	IMD	3	0
4	A	902	IMD	4	0

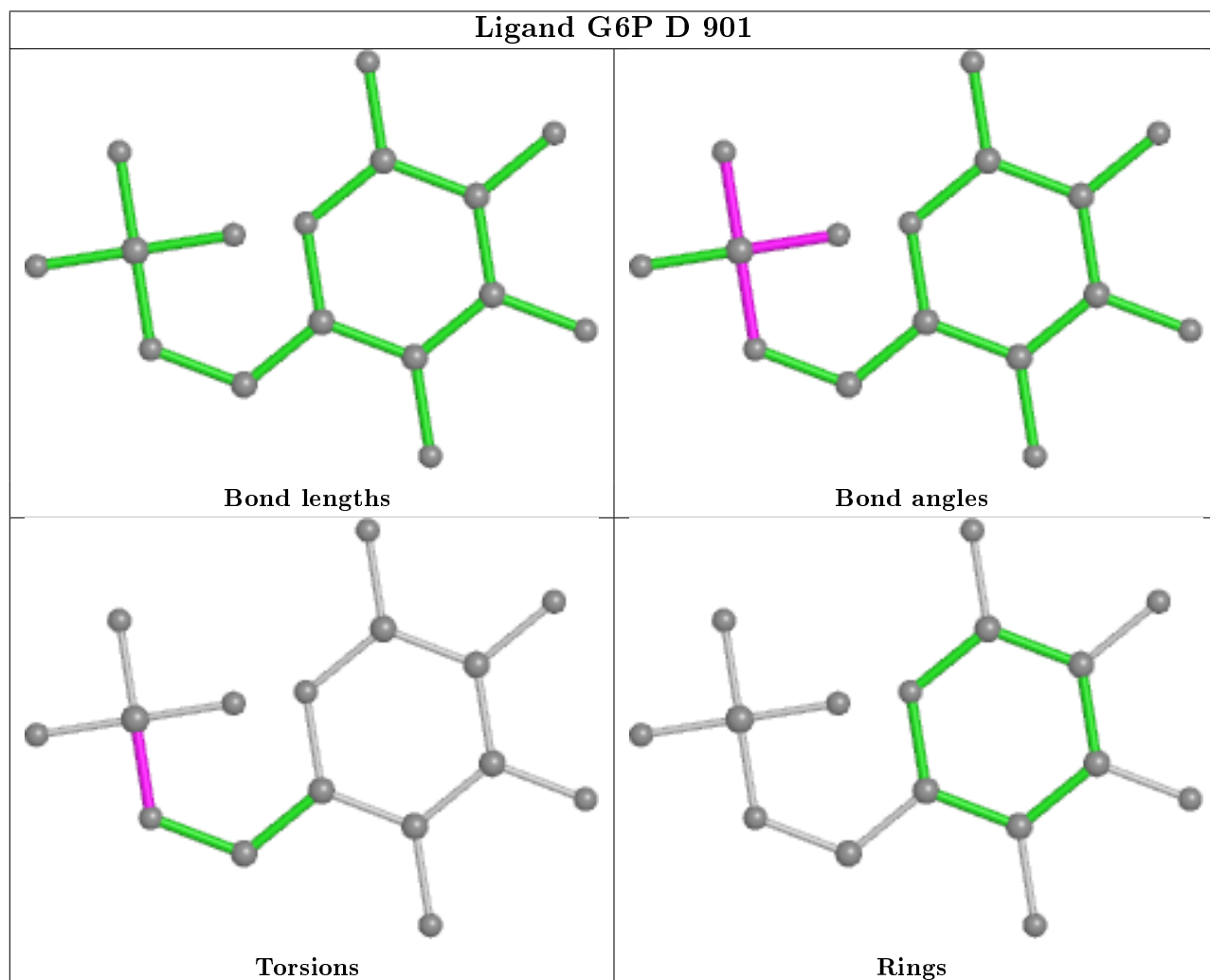
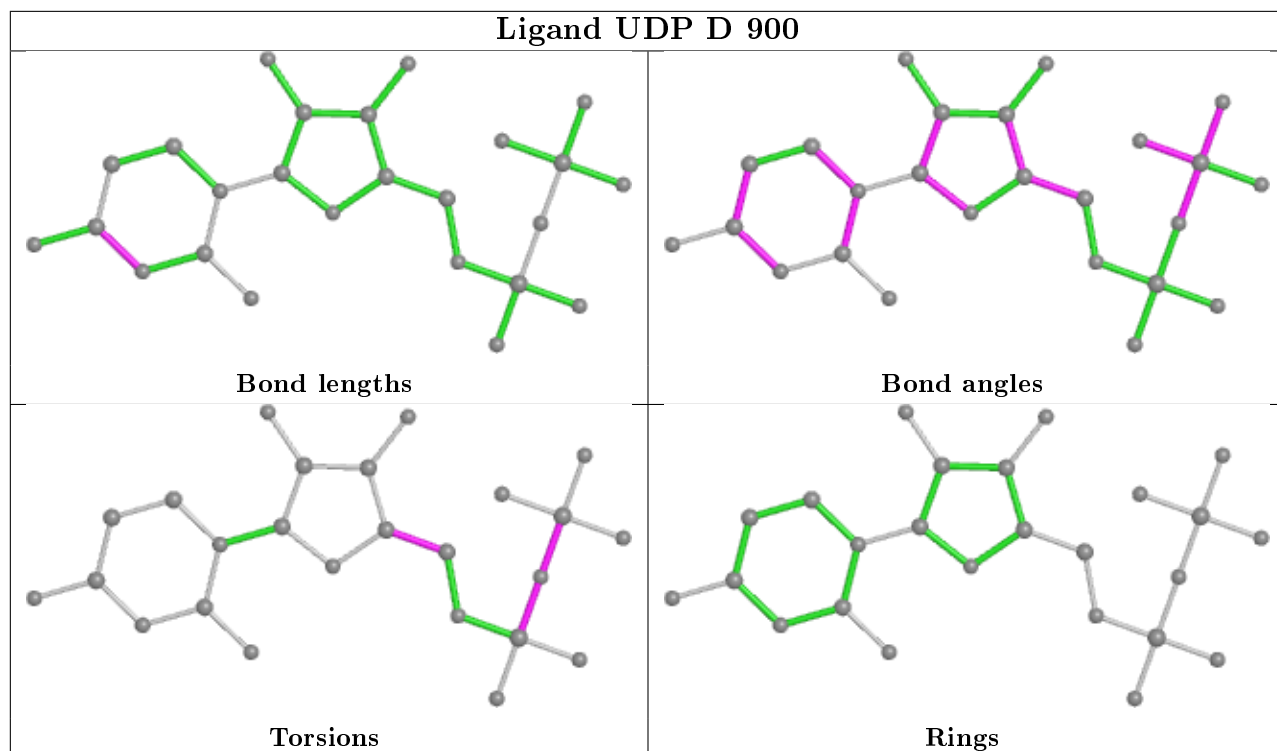
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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

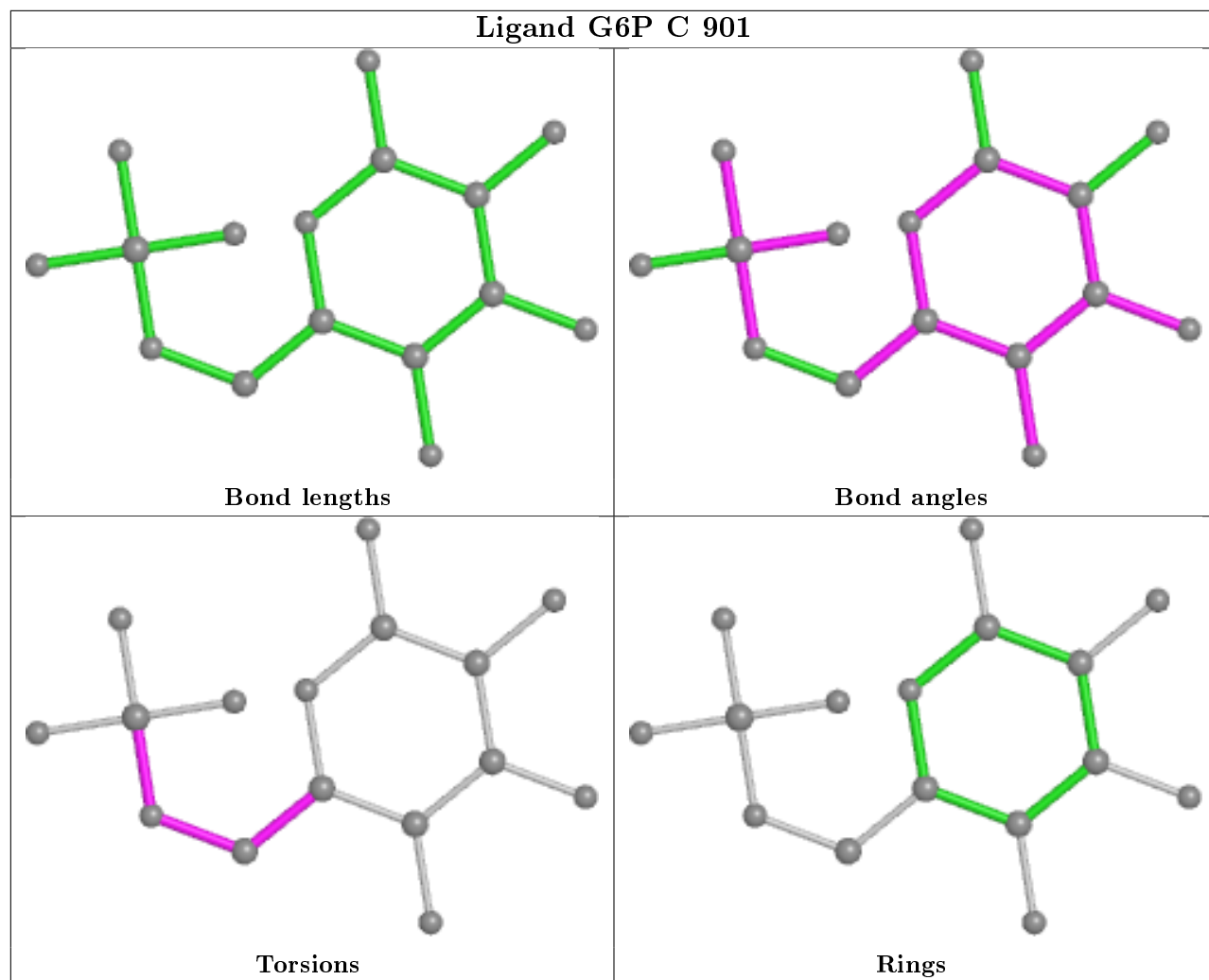
any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

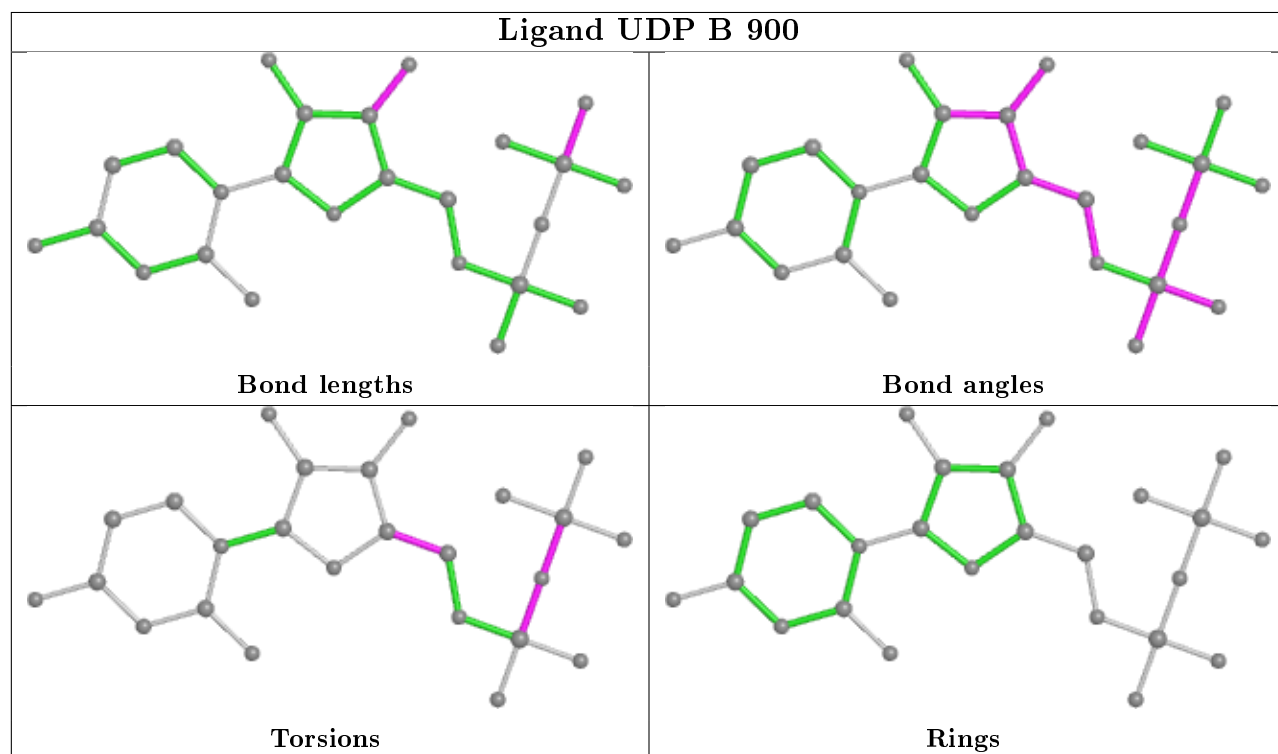
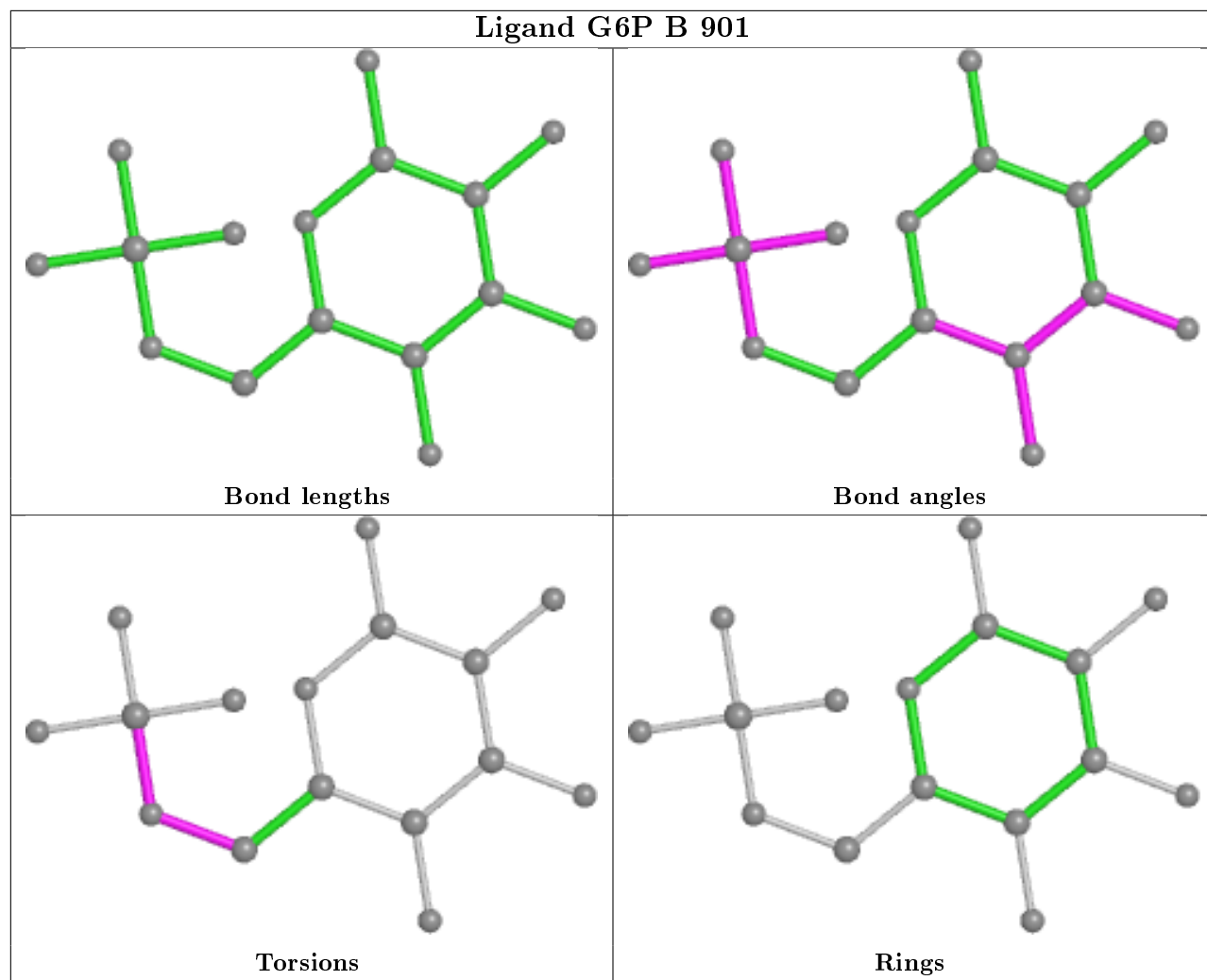












## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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.