



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 02:38 AM BST

PDB ID : 1QHA  
Title : HUMAN HEXOKINASE TYPE I COMPLEXED WITH ATP ANALOGUE  
AMP-PNP  
Authors : Rosano, C.; Sabini, E.; Deriu, D.; Magnani, M.; Bolognesi, M.  
Deposited on : 1999-05-11  
Resolution : 2.25 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)  
Xtriage (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

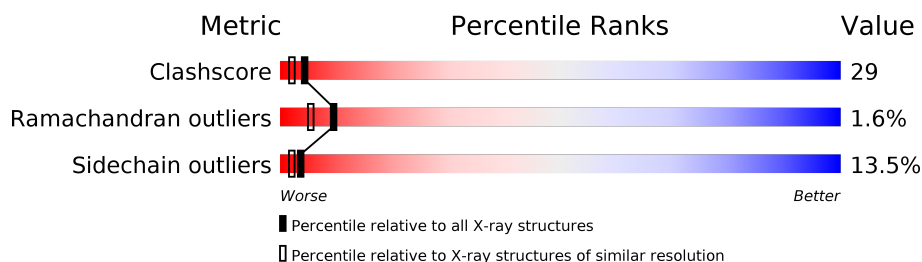
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

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	917	
1	B	917	

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
3	G6P	A	919	X	-	X	-
3	G6P	A	921	X	-	-	-
3	G6P	B	1916	X	-	-	-
3	G6P	B	1918	X	-	-	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 14737 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.

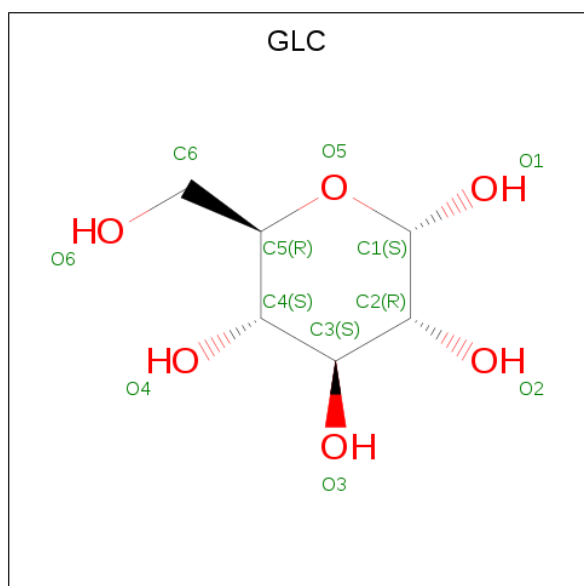
- Molecule 1 is a protein called PROTEIN (HEXOKINASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	903	Total	C	N	O	S	67	0	0
			7065	4428	1246	1338	53			
1	B	899	Total	C	N	O	S	62	0	0
			7032	4407	1241	1331	53			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	776	LEU	MET	conflict	UNP P19367
B	776	LEU	MET	conflict	UNP P19367

- Molecule 2 is alpha-D-glucopyranose (three-letter code: GLC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



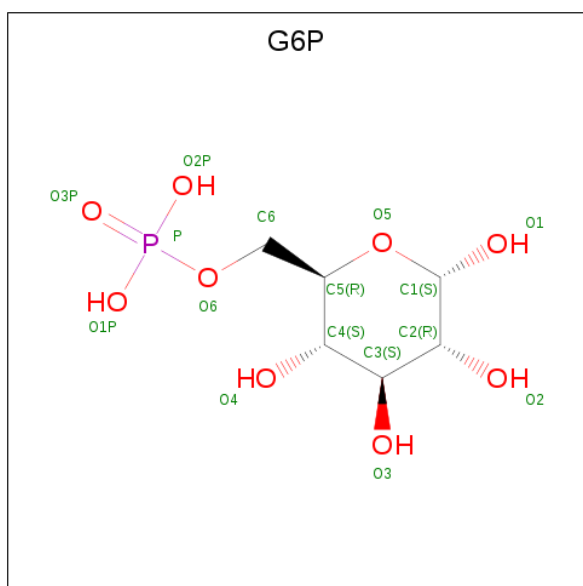
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		

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

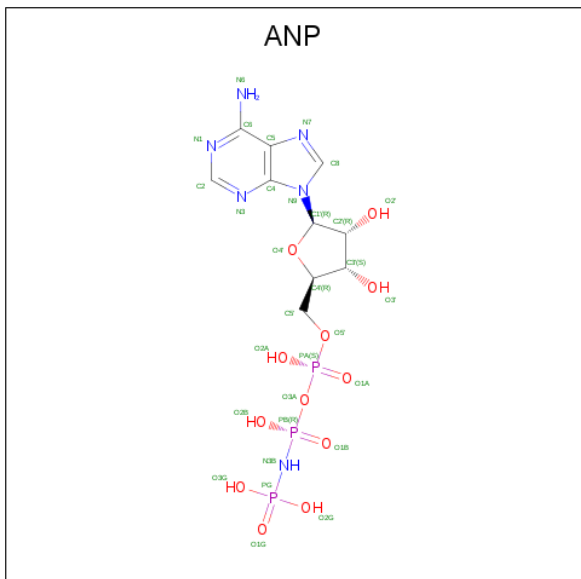


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	P	0	0
			16	6	9	1		
3	A	1	Total	C	O	P	0	0
			16	6	9	1		
3	B	1	Total	C	O	P	0	0
			16	6	9	1		
3	B	1	Total	C	O	P	0	0
			16	6	9	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	3	Total	Mg	0	0
			3	3		
4	A	3	Total	Mg	0	0
			3	3		

- Molecule 5 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula:  $\text{C}_{10}\text{H}_{17}\text{N}_6\text{O}_{12}\text{P}_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total 31	C 10	N 6	O 12	P 3	0	0
5	B	1	Total 31	C 10	N 6	O 12	P 3	0	0

- Molecule 6 is water.

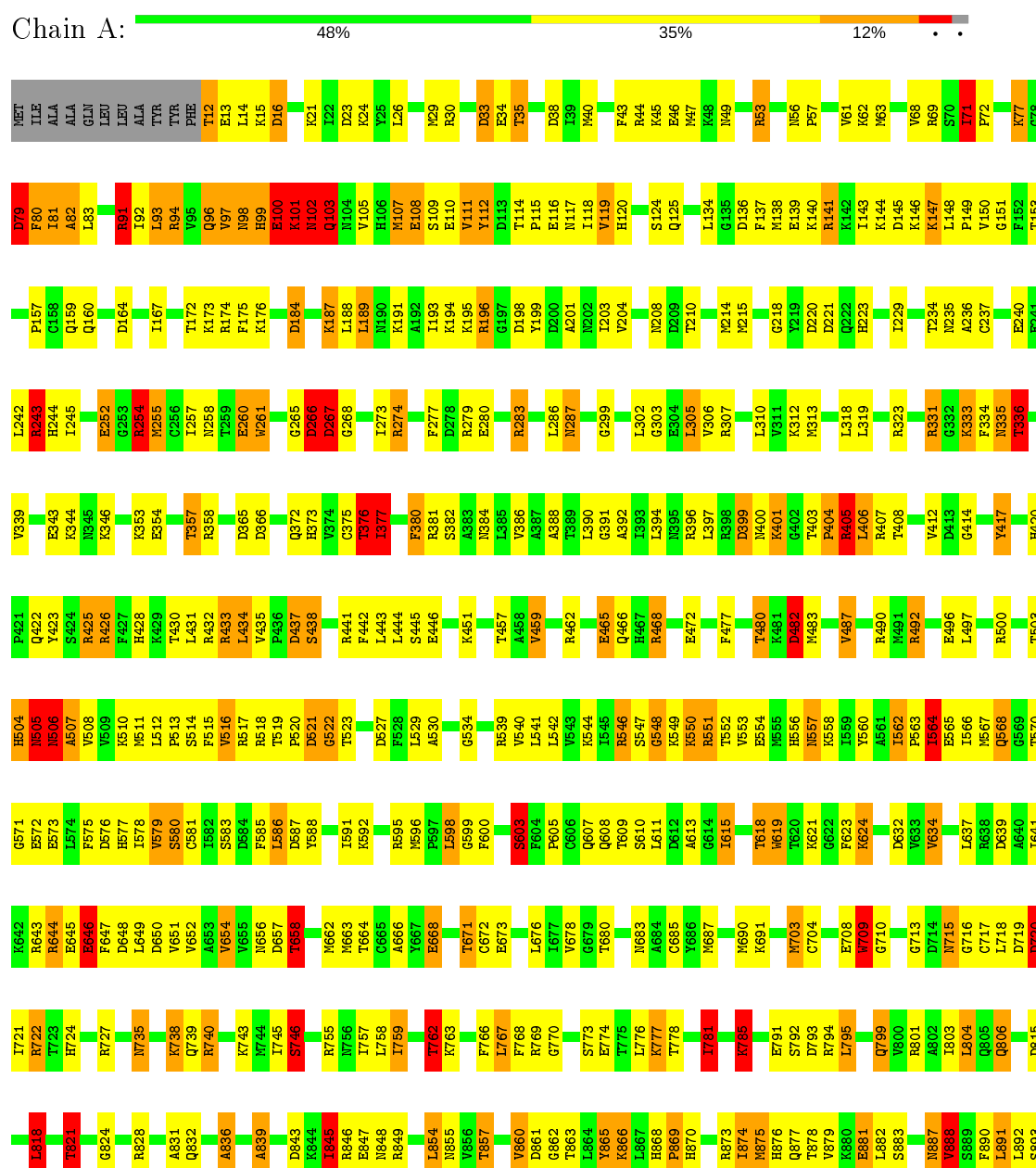
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	188	Total O 188 188	0	0
6	B	272	Total O 272 272	0	0

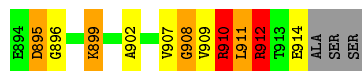
### 3 Residue-property plots

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.

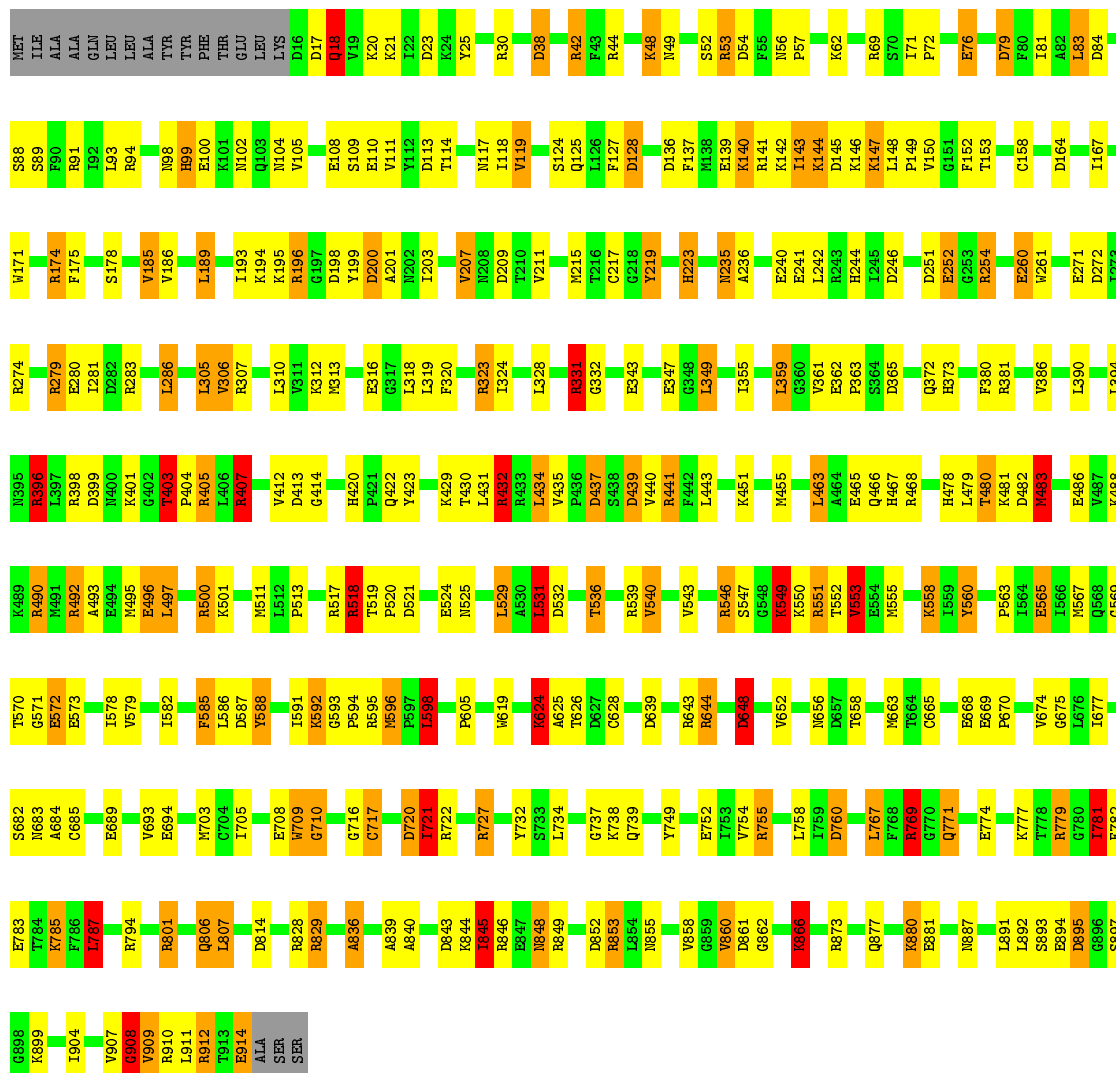
#### • Molecule 1: PROTEIN (HEXOKINASE)





• Molecule 1: PROTEIN (HEXOKINASE)

Chain B: 61% 26% 9%



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.59 Å   176.17 Å   86.76 Å 90.00°   90.53°   90.00°	Depositor
Resolution (Å)	12.00 – 2.25	Depositor
% Data completeness (in resolution range)	85.0 (12.00-2.25)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.10	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.208 , 0.279	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	14737	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, GLC, ANP, G6P

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	0.72	8/7170 (0.1%)	1.75	146/9647 (1.5%)
1	B	0.83	6/7134 (0.1%)	1.89	162/9594 (1.7%)
All	All	0.78	14/14304 (0.1%)	1.82	308/19241 (1.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	28
1	B	0	16
All	All	0	44

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	914	GLU	CB-CG	-32.64	0.90	1.52
1	A	144	LYS	CG-CD	-17.85	0.91	1.52
1	A	147	LYS	CG-CD	16.68	2.09	1.52
1	B	99	HIS	C-N	14.83	1.68	1.34
1	B	62	LYS	CG-CD	-13.55	1.06	1.52

The worst 5 of 308 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	102	ASN	N-CA-CB	-39.95	38.70	110.60
1	B	829	ARG	NE-CZ-NH2	-27.14	106.73	120.30
1	B	829	ARG	CD-NE-CZ	25.57	159.40	123.60
1	B	100	GLU	CB-CA-C	-25.01	60.37	110.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	755	ARG	NE-CZ-NH1	-24.72	107.94	120.30

There are no chirality outliers.

5 of 44 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	GLU	Mainchain
1	A	101	LYS	Mainchain
1	A	68	VAL	Mainchain
1	A	79	ASP	Mainchain
1	A	99	HIS	Mainchain

## 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	7065	0	7124	483	0
1	B	7032	0	7087	334	0
2	A	24	0	24	0	0
2	B	24	0	24	0	0
3	A	32	0	22	7	0
3	B	32	0	21	8	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
5	A	31	0	13	1	0
5	B	31	0	13	1	0
6	A	188	0	0	33	0
6	B	272	0	0	37	0
All	All	14737	0	14328	823	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 29.

The worst 5 of 823 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:505:ASN:O	1:A:507:ALA:N	1.57	1.31
1:A:521:ASP:HB3	6:A:1169:HOH:O	1.34	1.24
1:A:29:MET:HB2	1:A:377:ILE:HD11	1.27	1.17
1:B:441:ARG:HG3	1:B:441:ARG:HH11	1.13	1.11
1:A:62:LYS:CE	6:A:1057:HOH:O	1.89	1.11

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	899/917 (98%)	825 (92%)	51 (6%)	23 (3%)	5	2
1	B	891/917 (97%)	861 (97%)	24 (3%)	6 (1%)	22	21
All	All	1790/1834 (98%)	1686 (94%)	75 (4%)	29 (2%)	9	5

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ALA
1	A	100	GLU
1	A	111	VAL
1	A	504	HIS
1	A	506	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	778/788 (99%)	657 (84%)	121 (16%)	2	1
1	B	774/788 (98%)	686 (89%)	88 (11%)	5	3
All	All	1552/1576 (98%)	1343 (86%)	209 (14%)	4	2

5 of 209 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	735	ASN
1	A	895	ASP
1	B	781	ILE
1	A	762	THR
1	A	821	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	735	ASN
1	B	49	ASN
1	B	771	GLN
1	A	771	GLN
1	A	806	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 16 ligands modelled in this entry, 6 are monoatomic - leaving 10 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	$\# Z  > 2$
2	GLC	A	918	-	12,12,12	0.86	1 (8%)	17,17,17	1.75	4 (23%)
2	GLC	B	1915	-	12,12,12	0.54	0	17,17,17	1.45	2 (11%)
2	GLC	A	920	-	12,12,12	0.96	1 (8%)	17,17,17	2.02	4 (23%)
3	G6P	B	1918	-	16,16,16	1.40	3 (18%)	24,24,24	2.68	10 (41%)
5	ANP	B	1999	4	29,33,33	1.54	5 (17%)	31,52,52	1.45	4 (12%)
3	G6P	B	1916	-	16,16,16	0.98	1 (6%)	24,24,24	2.17	10 (41%)
2	GLC	B	1917	-	12,12,12	1.14	1 (8%)	17,17,17	2.02	6 (35%)
5	ANP	A	999	-	29,33,33	1.54	6 (20%)	31,52,52	1.56	6 (19%)
3	G6P	A	919	-	16,16,16	1.30	3 (18%)	24,24,24	2.46	11 (45%)
3	G6P	A	921	-	16,16,16	1.00	0	24,24,24	2.20	6 (25%)

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
2	GLC	A	918	-	-	0/2/22/22	0/1/1/1
2	GLC	B	1915	-	-	0/2/22/22	0/1/1/1
2	GLC	A	920	-	-	0/2/22/22	0/1/1/1
3	G6P	B	1918	-	1/1/6/6	3/6/26/26	0/1/1/1
5	ANP	B	1999	4	-	7/14/38/38	0/3/3/3
3	G6P	B	1916	-	1/1/6/6	2/6/26/26	0/1/1/1
2	GLC	B	1917	-	-	0/2/22/22	0/1/1/1
5	ANP	A	999	-	-	5/14/38/38	0/3/3/3
3	G6P	A	919	-	1/1/6/6	4/6/26/26	0/1/1/1
3	G6P	A	921	-	1/1/6/6	5/6/26/26	0/1/1/1

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	999	ANP	C8-N7	-3.44	1.28	1.34
5	B	1999	ANP	C8-N7	-3.30	1.28	1.34
5	A	999	ANP	PG-O3G	-3.03	1.48	1.56
5	A	999	ANP	PG-O2G	-3.01	1.48	1.56
5	B	1999	ANP	PG-O3G	-2.98	1.48	1.56

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1918	G6P	O5-C5-C6	-6.82	92.91	106.67
3	A	919	G6P	C6-C5-C4	5.38	123.32	112.09
3	B	1918	G6P	C6-C5-C4	5.07	122.68	112.09
3	A	921	G6P	O5-C5-C4	5.01	118.79	109.69
3	A	921	G6P	O5-C1-C2	4.84	118.93	110.28

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	1916	G6P	C1
3	B	1918	G6P	C1
3	A	919	G6P	C1
3	A	921	G6P	C1

5 of 26 torsion outliers are listed below:

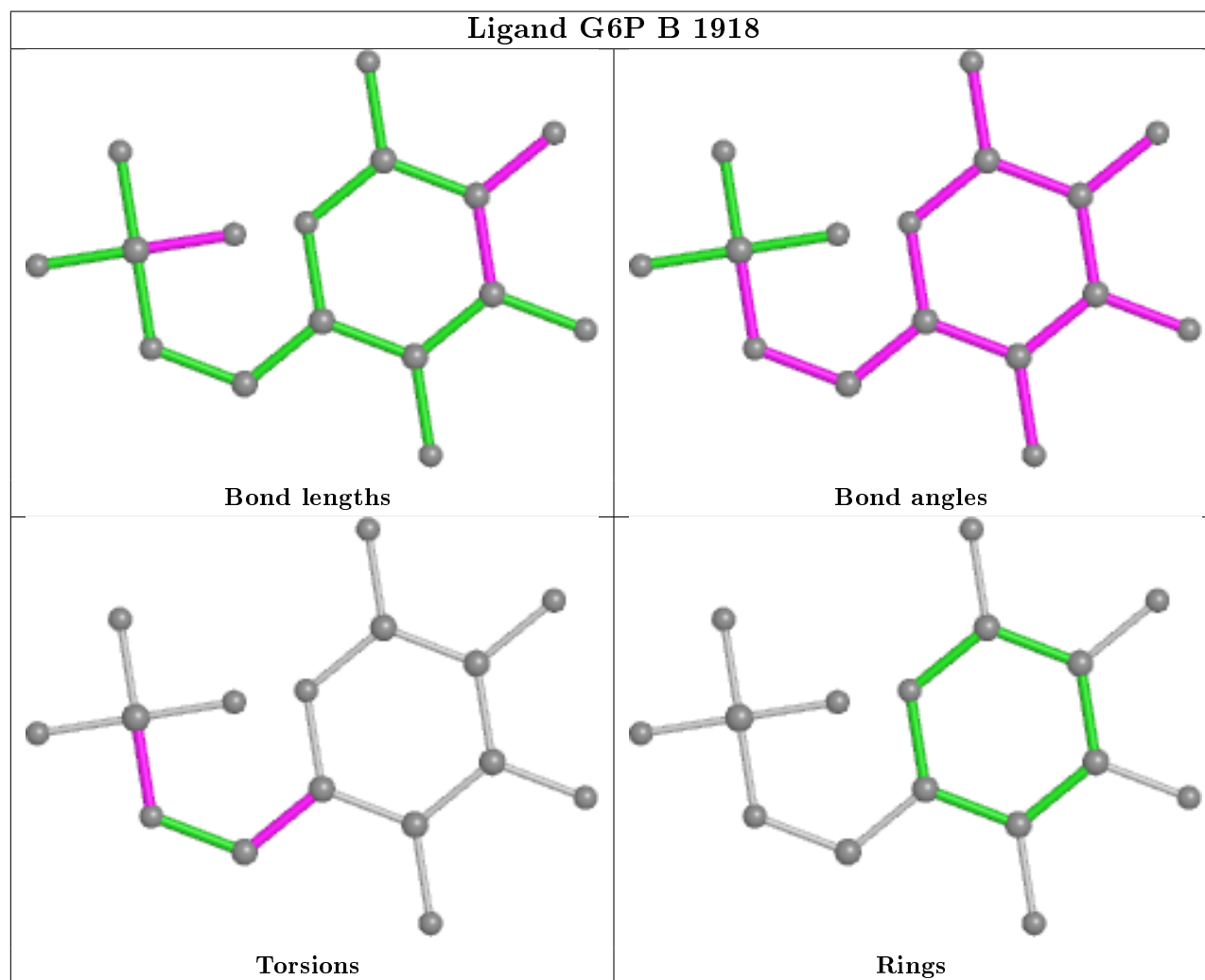
Mol	Chain	Res	Type	Atoms
3	B	1916	G6P	C4-C5-C6-O6
3	B	1916	G6P	O5-C5-C6-O6
3	B	1918	G6P	C4-C5-C6-O6
3	B	1918	G6P	O5-C5-C6-O6
5	A	999	ANP	PB-N3B-PG-O1G

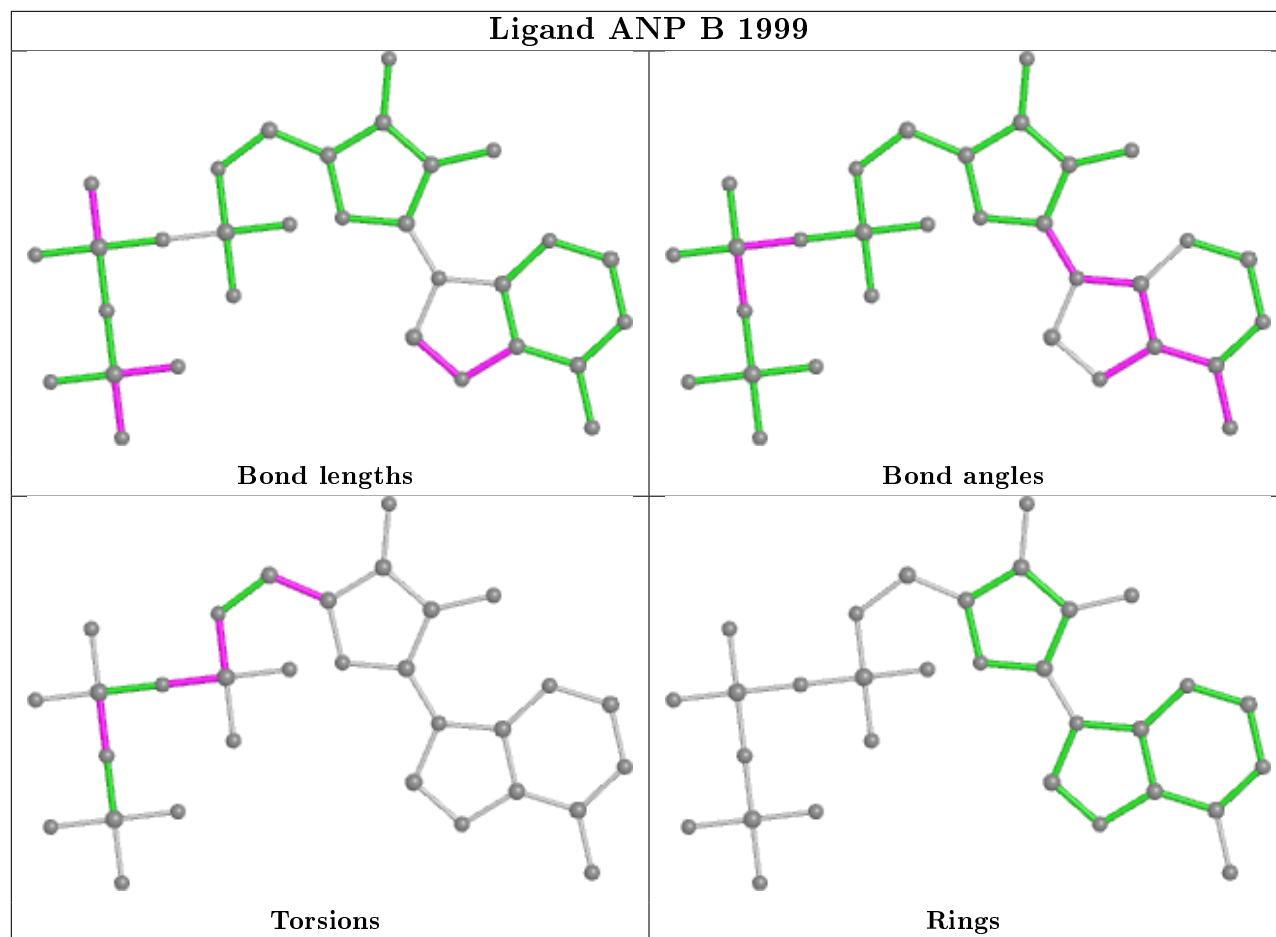
There are no ring outliers.

6 monomers are involved in 17 short contacts:

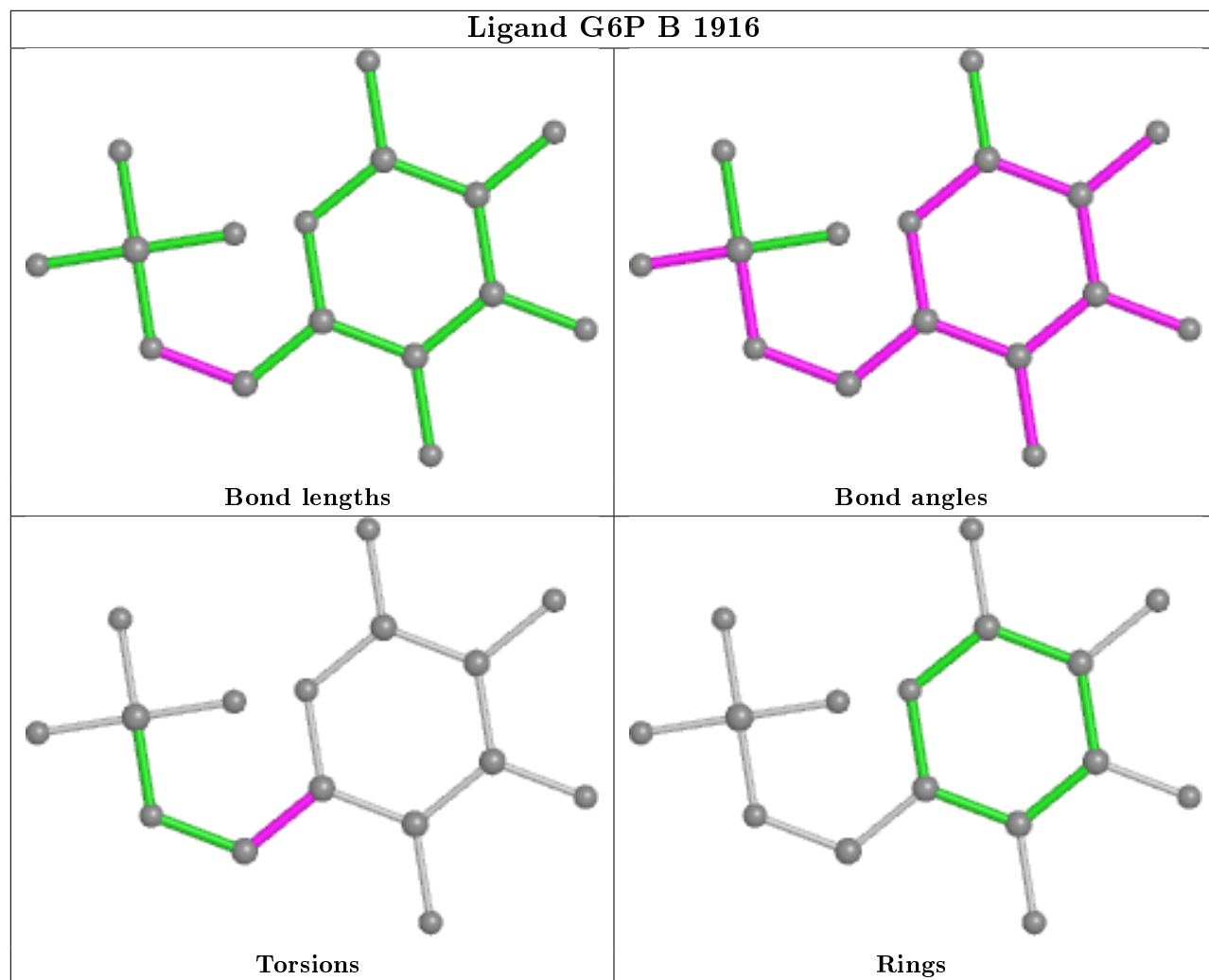
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1918	G6P	5	0
5	B	1999	ANP	1	0
3	B	1916	G6P	3	0
5	A	999	ANP	1	0
3	A	919	G6P	6	0
3	A	921	G6P	1	0

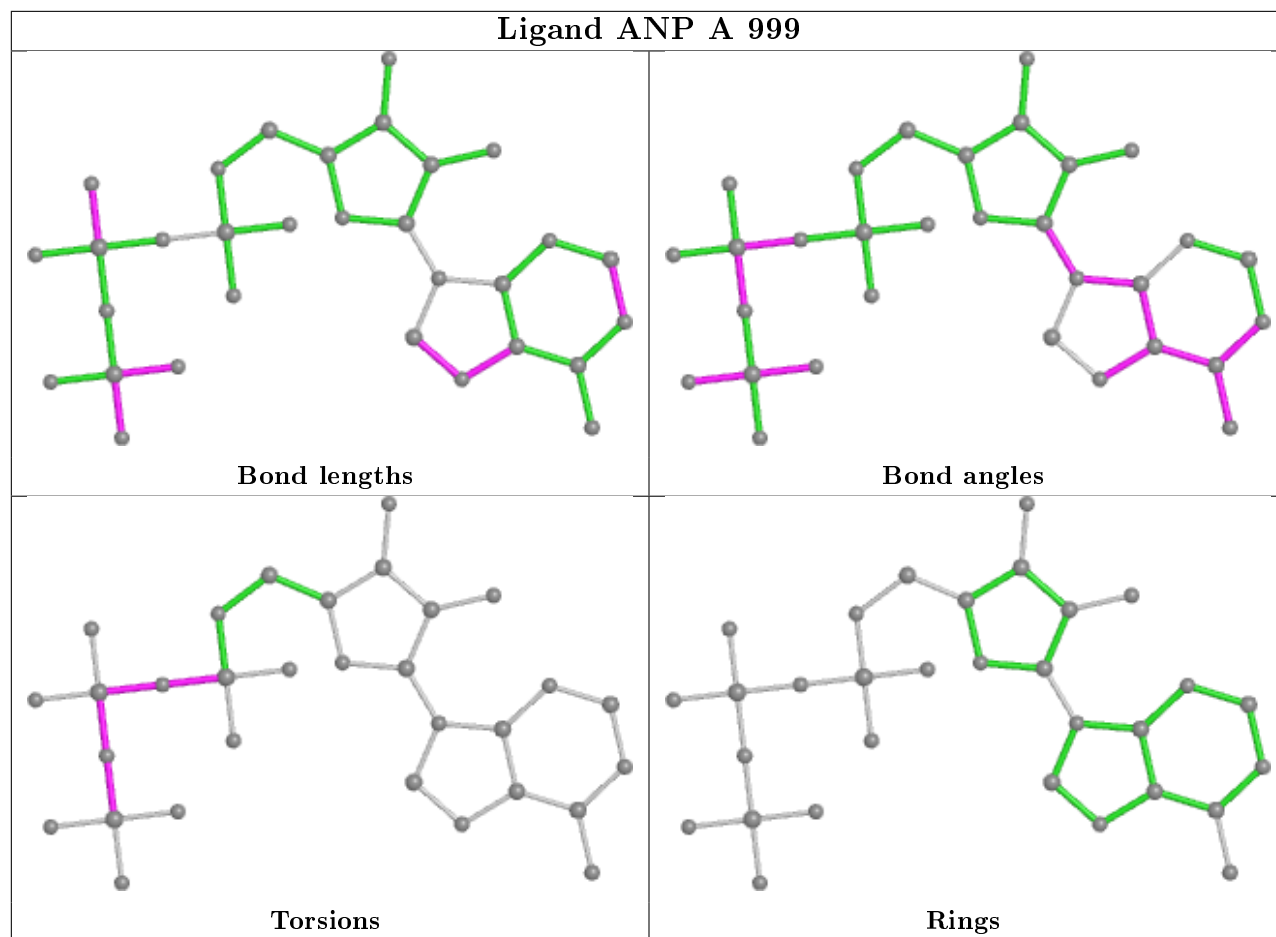
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



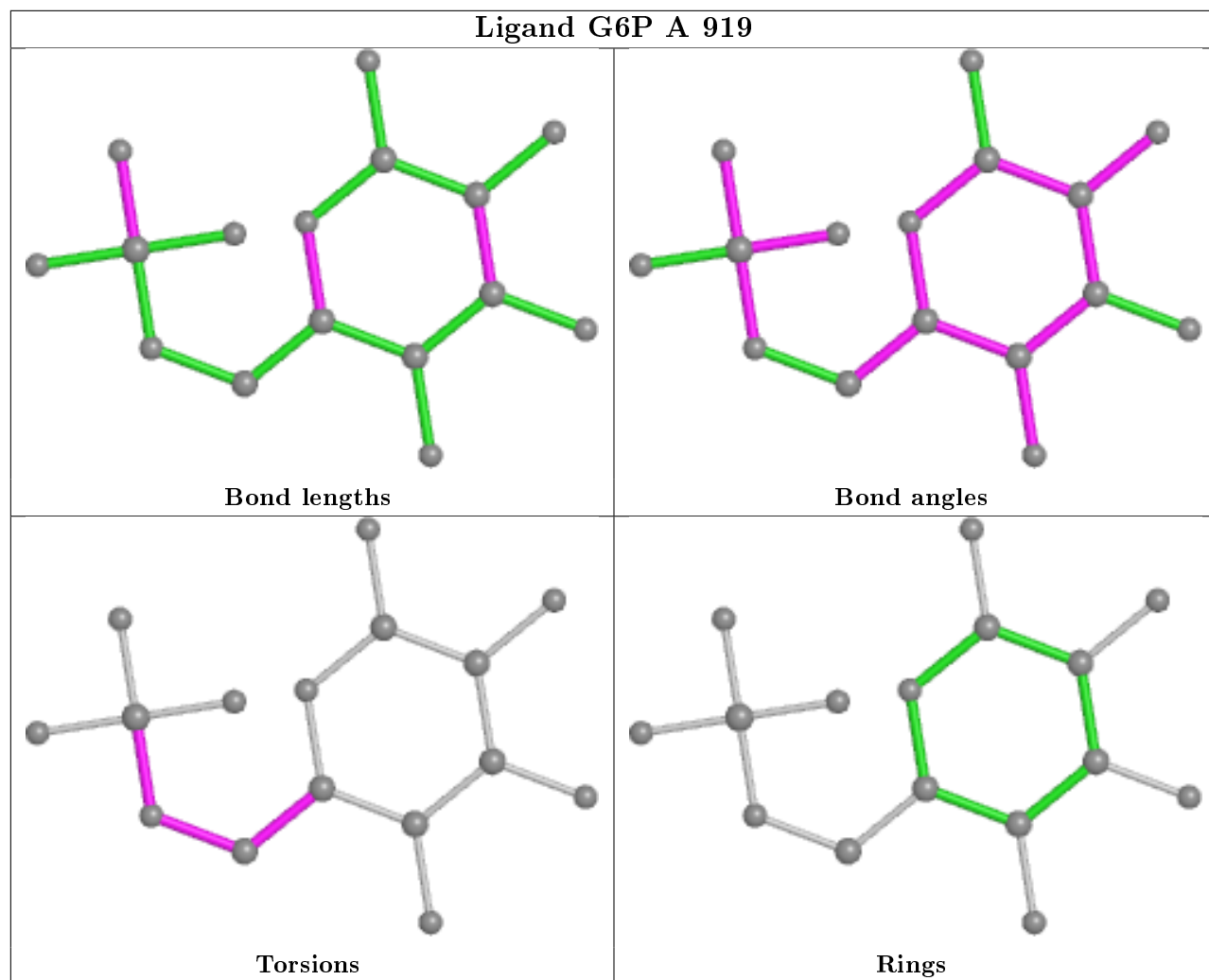


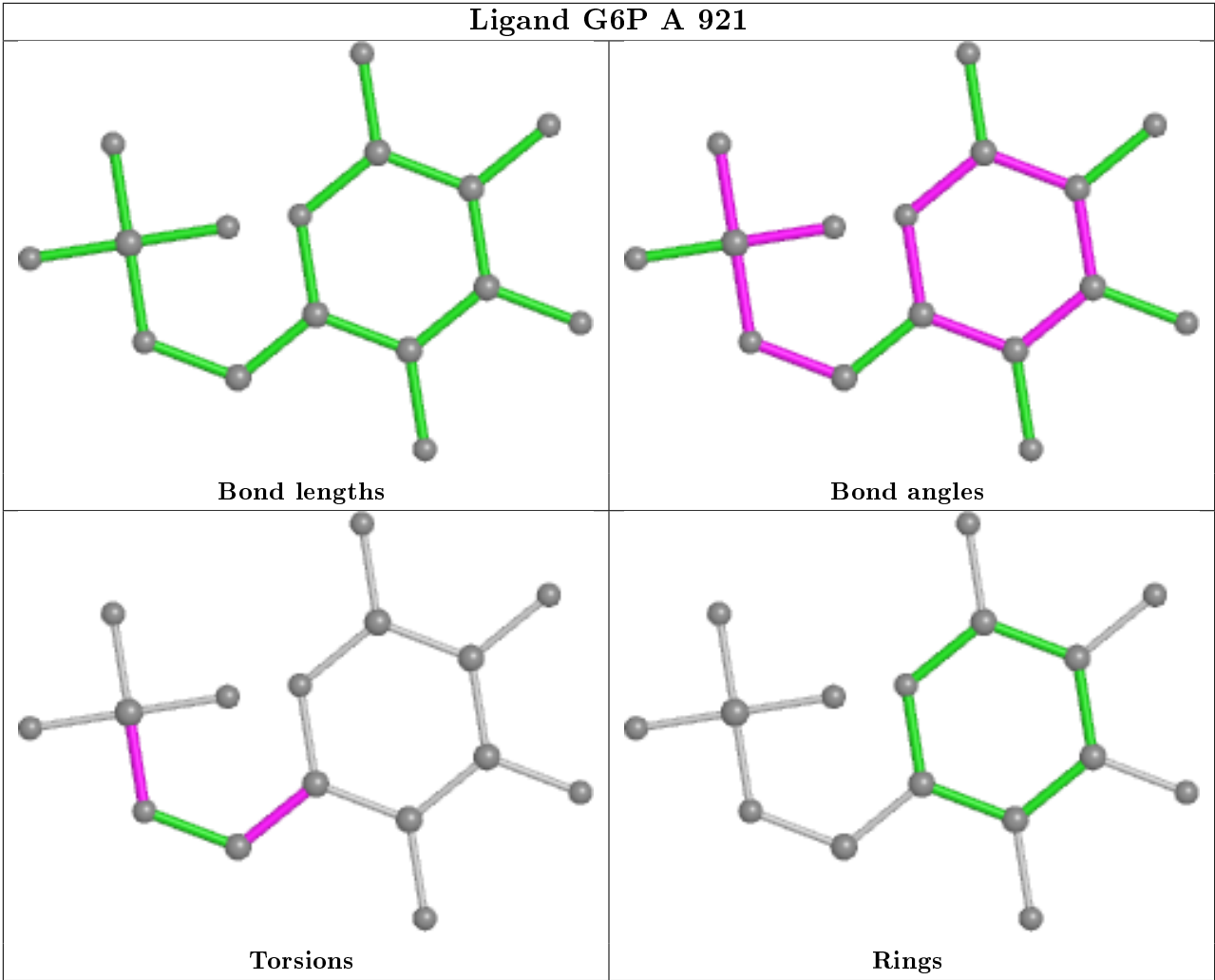






## Ligand G6P A 919





5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	5
1	A	2

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	101:LYS	C	102:ASN	N	2.86

*Continued on next page...*

*Continued from previous page...*

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	100:GLU	C	101:LYS	N	2.68
1	B	16:ASP	C	17:ASP	N	2.25
1	B	103:GLN	C	104:ASN	N	2.23
1	A	912:ARG	C	913:THR	N	2.03

## 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.