wwPDB X-ray Structure Validation Summary Report (i)

Jan 27, 2024 – 01:31 PM EST

PDB ID	:	1BG3
Title	:	RAT BRAIN HEXOKINASE TYPE I COMPLEX WITH GLUCOSE AND INHIBITOR GLUCOSE-6-PHOSPHATE
Authors	:	Mulichak, A.M.; Garavito, R.M.
Deposited on	:	1998-06-04
Resolution	:	2.80 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:

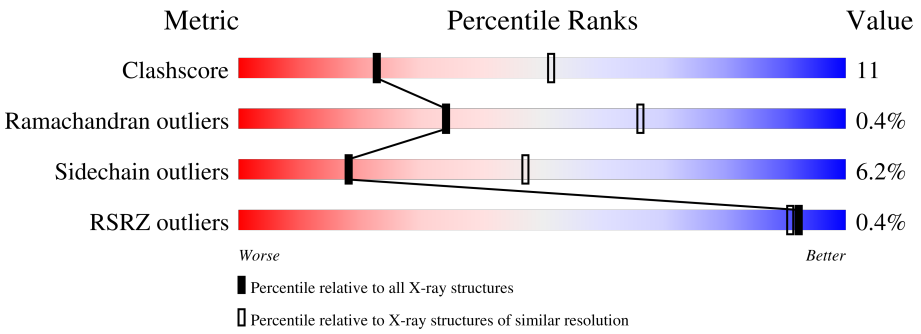
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.80 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)
RSRZ outliers	127900	3078 (2.80-2.80)

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 <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	918	68% 27% • •
1	B	918	71% 26% • •

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	1002	X	-	-	-
3	G6P	A	1004	X	-	-	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	G6P	B	1002	X	-	-	-
3	G6P	B	1004	X	-	-	-

2 Entry composition (i)

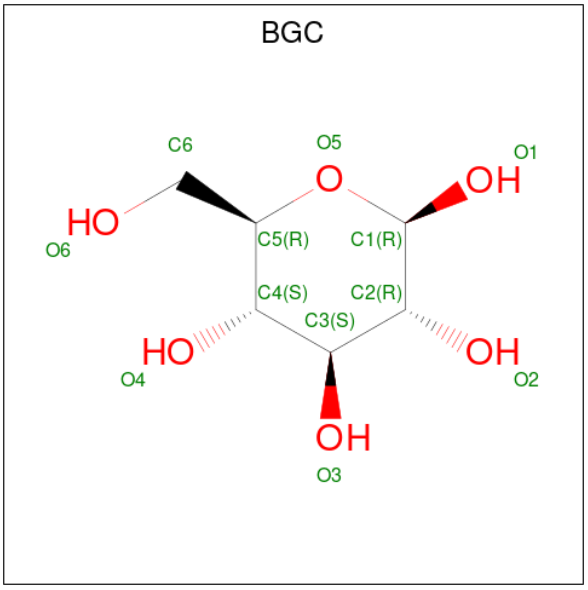
There are 5 unique types of molecules in this entry. The entry contains 13962 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HEXOKINASE.

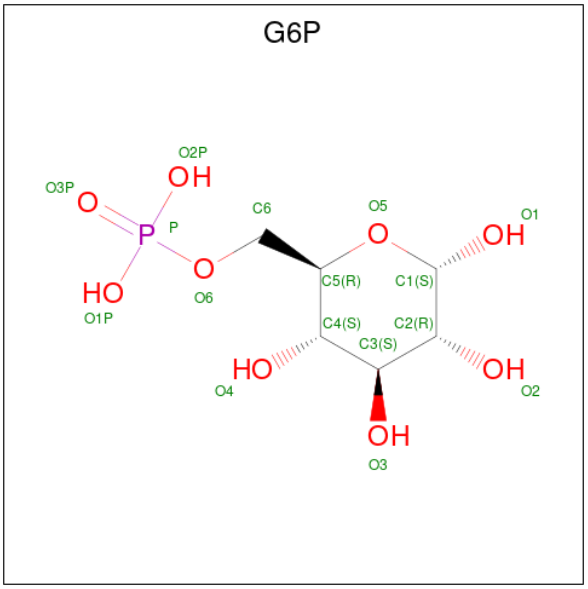
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	902	6832	4304	1177	1296	55	0	0	0
1	B	902	6783	4276	1157	1295	55	0	0	0

- Molecule 2 is beta-D-glucopyranose (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



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

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

- Molecule 5 is water.

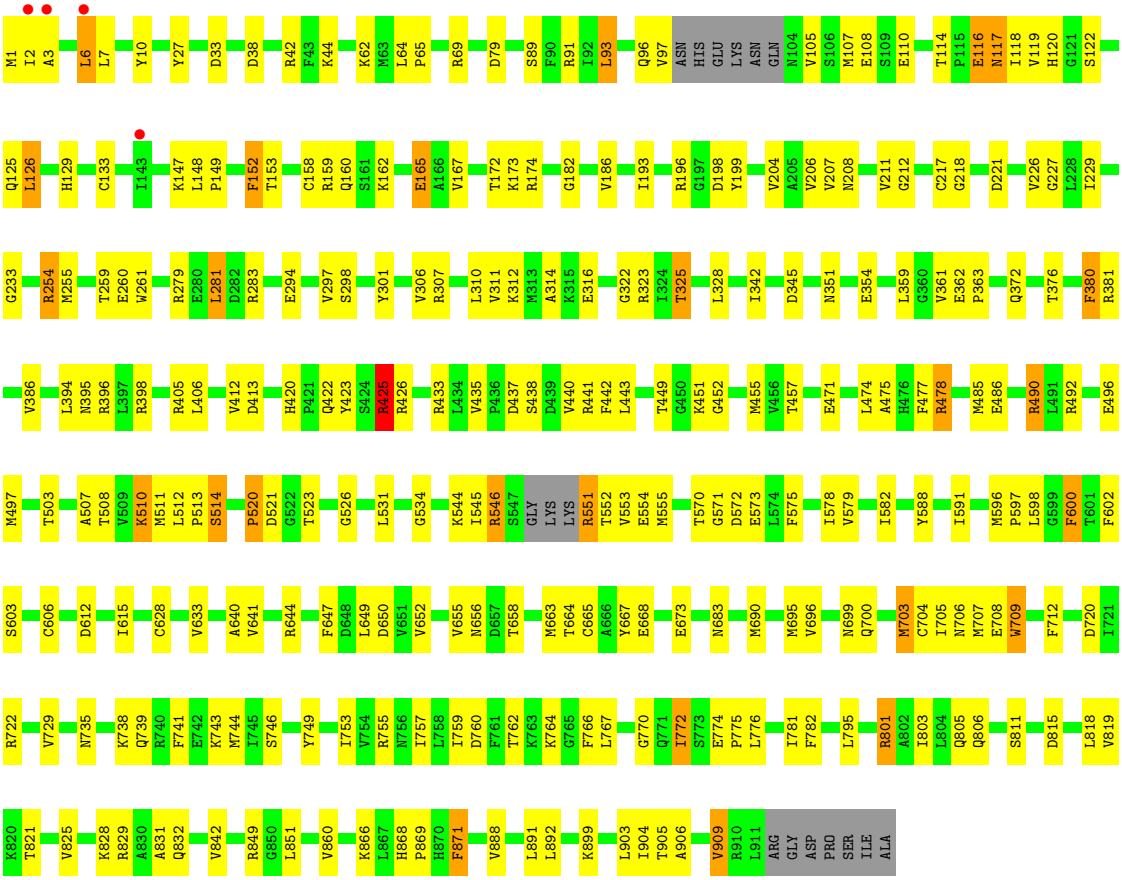
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	136	Total	O	0	0
			136	136		
5	B	98	Total	O	0	0
			98	98		

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

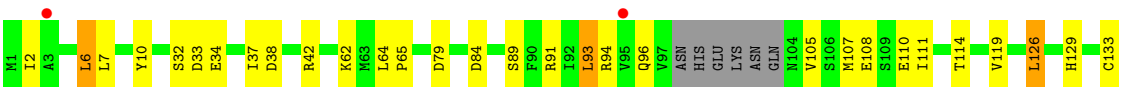
• Molecule 1: HEXOKINASE

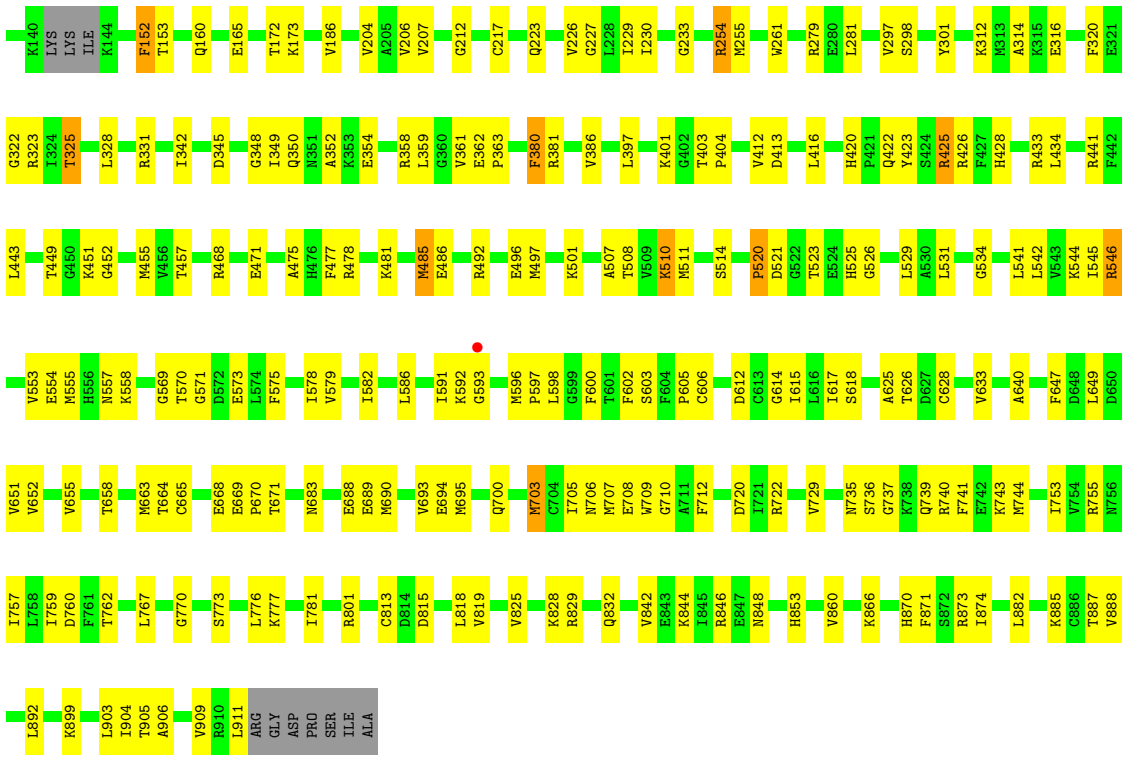
Chain A: 68% 27%



• Molecule 1: HEXOKINASE

Chain B: 71% 26%




K140 LYS LYS ILE K144 F152 T153 Q160 E165 T172 K173 V186 V204 A205 V206 V207 G212 C217 Q223 V226 G227 L228 I229 I230 G233 R254 M255 W261 R279 E280 L281 V297 S298 Y301 K312 M313 A314 K315 E316 F320 E321
G322 R323 I324 T325 L328 R331 I342 D345 G348 I349 Q350 N351 A352 K353 E354 R358 L359 G360 V361 E362 P363 F380 R381 V386 L397 K401 G402 T403 P404 V412 D413 L416 H420 P421 Q422 Y423 S424 R425 R426 F427 H428 R433 L434 R441 F442
L443 T449 G450 K451 G452 M455 V456 T457 R468 E471 A475 H476 F477 R478 K481 M485 E486 R492 E496 M497 K501 A507 T508 V509 K510 M511 S514 P520 D521 G522 T523 E524 H525 G526 L529 A530 L531 G534 L541 L542 V543 K544 I545 R546
V553 E554 M555 H556 N557 K558 G569 T570 G571 D572 E573 L574 F575 I578 V579 I582 L586 I591 K592 G593 M596 P597 L598 G599 F600 T601 F602 S603 F604 P605 C606 D612 C613 G614 I615 L616 I617 S618 A625 T626 D627 C628 V633 A640 F647 D648 L649 D650
V651 V652 V655 T658 M663 T664 C665 E668 E669 P670 T671 N683 E688 E689 M690 V693 E694 M695 Q700 M703 G704 I705 N706 M707 E708 W709 G710 A711 F712 D720 I721 R722 V729 N735 S736 G737 K738 Q739 R740 F741 E742 K743 M744 I753 V754 R755 N756
I757 L758 I759 D760 F761 T762 L767 G770 S773 L776 K777 I781 R801 C813 D814 D815 L818 V819 V825 K828 R829 Q832 V842 E843 K844 I845 R846 E847 N848 H853 V860 K866 H870 F871 S872 R873 I874 L882 K885 C886 T887 V888
L892 K899 L903 I904 T905 A906 V909 R910 L911 ARG GLY ASP PRO SER ILE ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	132.10Å 77.10Å 137.10Å 90.00° 96.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80 38.48 – 2.80	Depositor EDS
% Data completeness (in resolution range)	71.0 (20.00-2.80) 70.6 (38.48-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.83 (at 2.81Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.196 , 0.251 0.198 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	35.9	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 74.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.045 for l,-k,h	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	13962	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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.

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: G6P, CA, BGC

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	0.62	1/6935 (0.0%)	0.89	15/9367 (0.2%)
1	B	0.62	0/6885	0.90	15/9310 (0.2%)
All	All	0.62	1/13820 (0.0%)	0.90	30/18677 (0.2%)

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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	158	CYS	CB-SG	-5.23	1.73	1.81

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	425	ARG	NE-CZ-NH2	-14.98	112.81	120.30
1	B	279	ARG	NE-CZ-NH2	-14.97	112.81	120.30
1	A	279	ARG	NE-CZ-NH2	-14.74	112.93	120.30
1	B	425	ARG	NE-CZ-NH2	-14.09	113.26	120.30
1	B	279	ARG	NE-CZ-NH1	14.07	127.34	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	490	ARG	Sidechain

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	6832	0	6711	177	0
1	B	6783	0	6623	136	0
2	A	24	0	23	2	0
2	B	24	0	24	1	0
3	A	32	0	22	0	0
3	B	32	0	22	0	0
4	A	1	0	0	0	0
5	A	136	0	0	6	0
5	B	98	0	0	0	0
All	All	13962	0	13425	311	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 11.

The worst 5 of 311 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:361:VAL:O	1:A:363:PRO:HD3	1.62	1.00
1:B:361:VAL:O	1:B:363:PRO:HD3	1.67	0.93
1:A:312:LYS:O	1:A:316:GLU:HG3	1.80	0.81
1:A:162:LYS:HG3	1:A:165:GLU:HB2	1.62	0.81
1:A:849:ARG:HB2	1:A:851:LEU:HD11	1.62	0.81

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	896/918 (98%)	829 (92%)	65 (7%)	2 (0%)	47	78
1	B	896/918 (98%)	835 (93%)	55 (6%)	6 (1%)	22	53
All	All	1792/1836 (98%)	1664 (93%)	120 (7%)	8 (0%)	34	66

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	345	ASP
1	A	345	ASP
1	A	871	PHE
1	B	871	PHE
1	B	404	PRO

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	722/791 (91%)	676 (94%)	46 (6%)	17	45
1	B	714/791 (90%)	671 (94%)	43 (6%)	19	48
All	All	1436/1582 (91%)	1347 (94%)	89 (6%)	18	47

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	380	PHE
1	B	555	MET
1	B	425	ARG
1	B	492	ARG
1	B	626	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	125	GLN
1	B	384	ASN
1	B	853	HIS
1	B	692	ASN
1	A	692	ASN

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 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	$\#\|Z\| > 2$
3	G6P	A	1004	-	16,16,16	0.69	0	24,24,24	0.79	1 (4%)
3	G6P	B	1002	-	16,16,16	0.73	0	24,24,24	0.92	1 (4%)
2	BGC	B	1003	-	12,12,12	0.38	0	17,17,17	0.64	0
2	BGC	A	1001	-	12,12,12	0.61	0	17,17,17	1.02	1 (5%)
3	G6P	A	1002	-	16,16,16	0.74	0	24,24,24	0.74	0
2	BGC	B	1001	-	12,12,12	0.53	0	17,17,17	0.97	1 (5%)
2	BGC	A	1003	-	12,12,12	0.32	0	17,17,17	0.54	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	G6P	B	1004	-	16,16,16	0.73	0	24,24,24	0.83	1 (4%)

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	1004	-	1/1/6/6	4/6/26/26	0/1/1/1
3	G6P	B	1002	-	1/1/6/6	3/6/26/26	0/1/1/1
2	BGC	B	1003	-	-	2/2/22/22	0/1/1/1
2	BGC	A	1001	-	-	0/2/22/22	0/1/1/1
3	G6P	A	1002	-	1/1/6/6	3/6/26/26	0/1/1/1
2	BGC	B	1001	-	-	0/2/22/22	0/1/1/1
2	BGC	A	1003	-	-	2/2/22/22	0/1/1/1
3	G6P	B	1004	-	1/1/6/6	4/6/26/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1001	BGC	C1-C2-C3	-3.00	104.09	110.31
2	B	1001	BGC	C1-C2-C3	-2.63	104.85	110.31
3	B	1002	G6P	P-O6-C6	2.13	124.15	118.30
3	B	1004	G6P	P-O6-C6	2.09	124.05	118.30
3	A	1004	G6P	P-O6-C6	2.01	123.82	118.30

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	1002	G6P	C1
3	A	1004	G6P	C1
3	B	1002	G6P	C1
3	B	1004	G6P	C1

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1004	G6P	C5-C6-O6-P
3	A	1004	G6P	C6-O6-P-O1P

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	B	1004	G6P	C5-C6-O6-P
3	B	1004	G6P	C6-O6-P-O1P
3	B	1002	G6P	C5-C6-O6-P

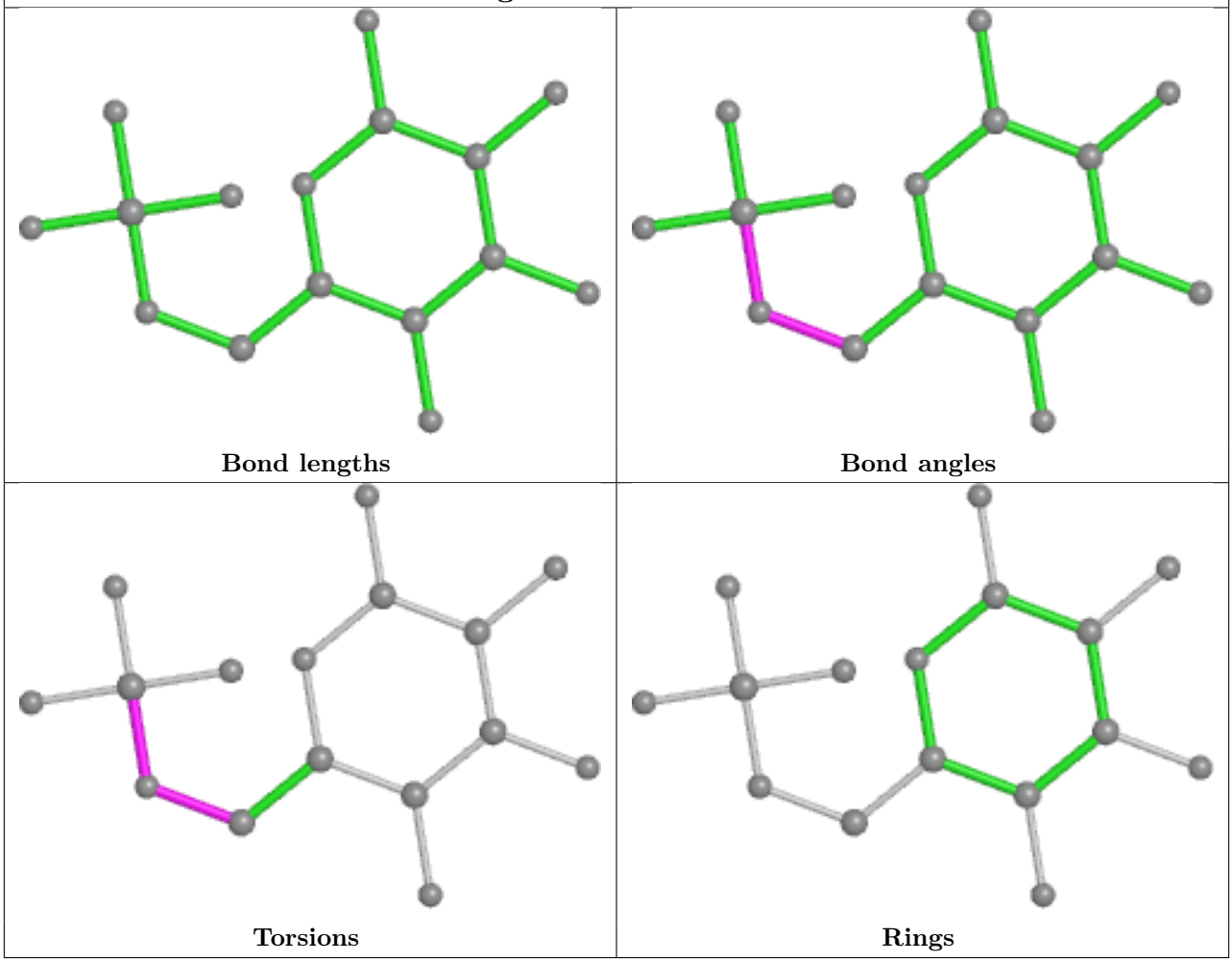
There are no ring outliers.

2 monomers are involved in 3 short contacts:

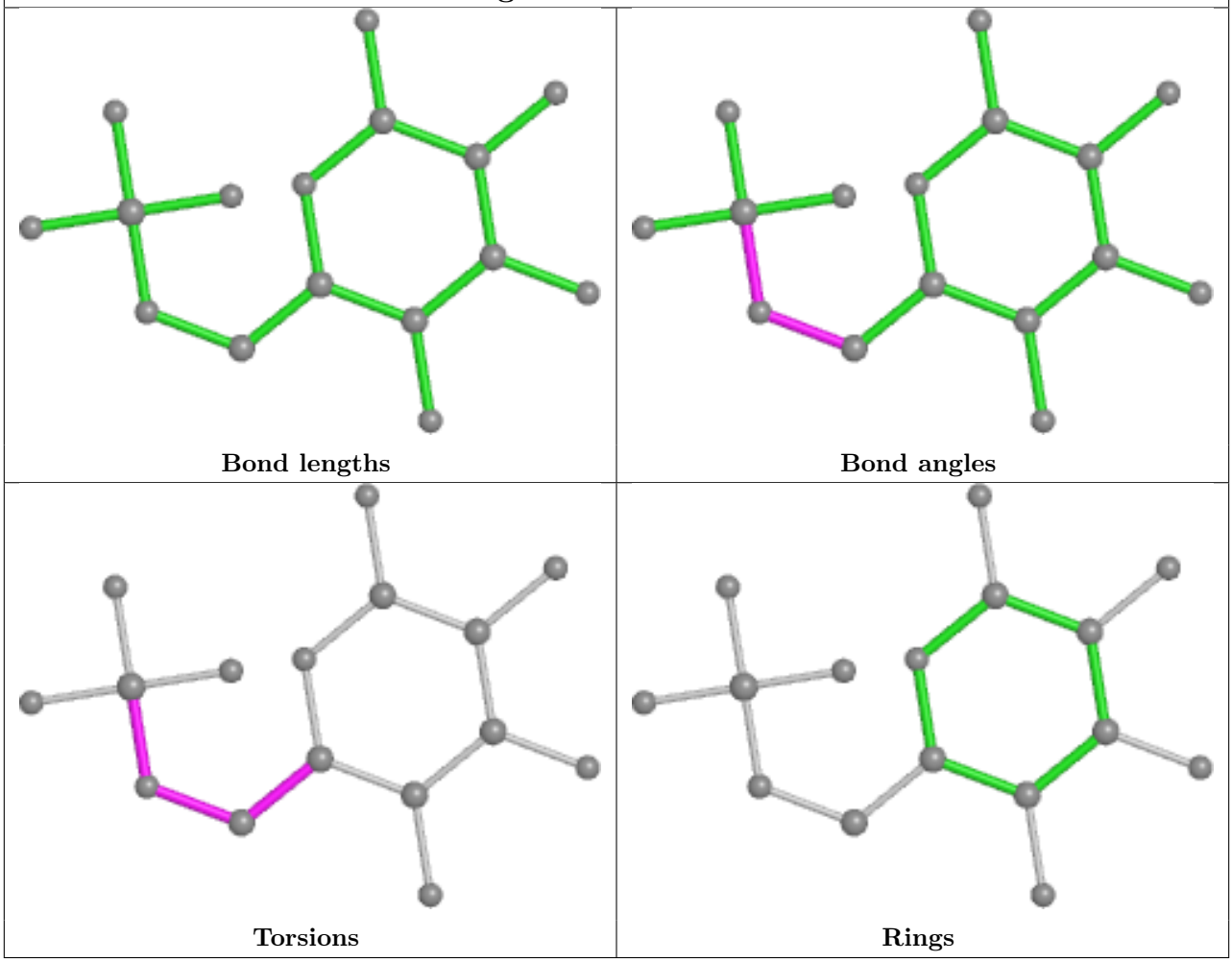
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	BGC	2	0
2	B	1001	BGC	1	0

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 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.

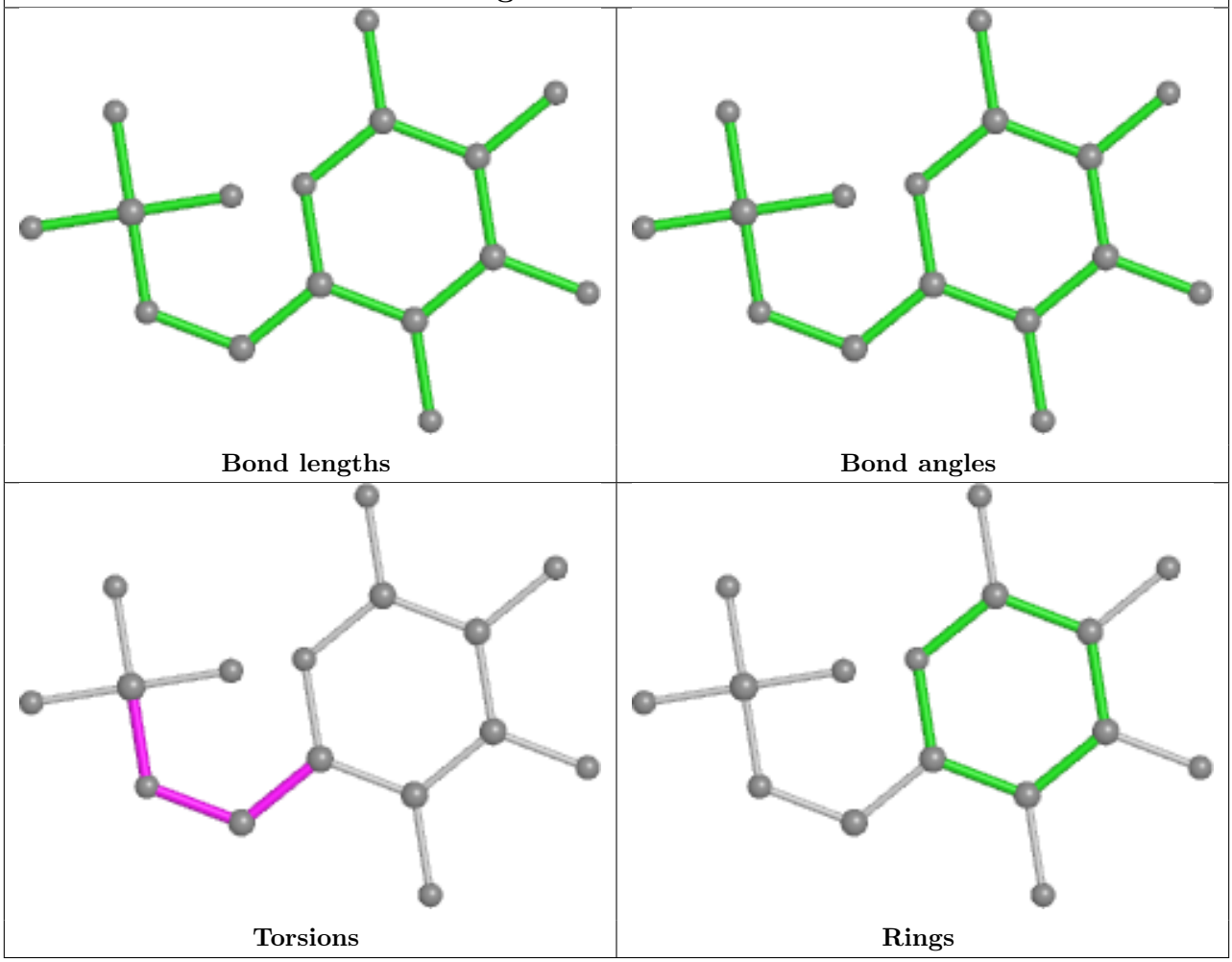
Ligand G6P A 1004



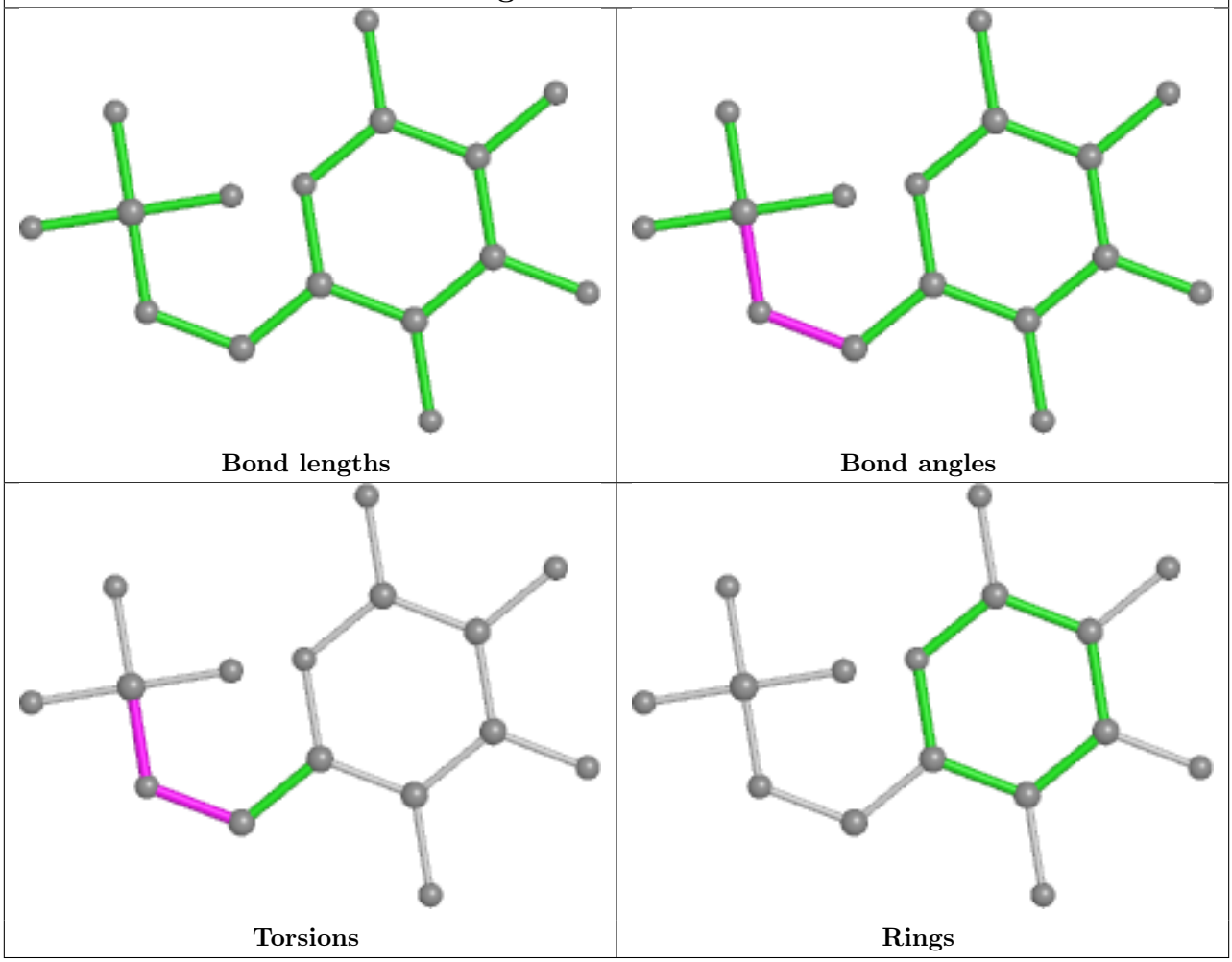
Ligand G6P B 1002



Ligand G6P A 1002



Ligand G6P B 1004



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 (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	902/918 (98%)	-0.44	4 (0%)	92	91	14, 33, 47, 59	0
1	B	902/918 (98%)	-0.46	3 (0%)	94	93	15, 33, 47, 59	0
All	All	1804/1836 (98%)	-0.45	7 (0%)	92	91	14, 33, 47, 59	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3	ALA	3.3
1	B	95	VAL	3.2
1	A	3	ALA	3.0
1	A	2	ILE	2.9
1	A	6	LEU	2.7

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

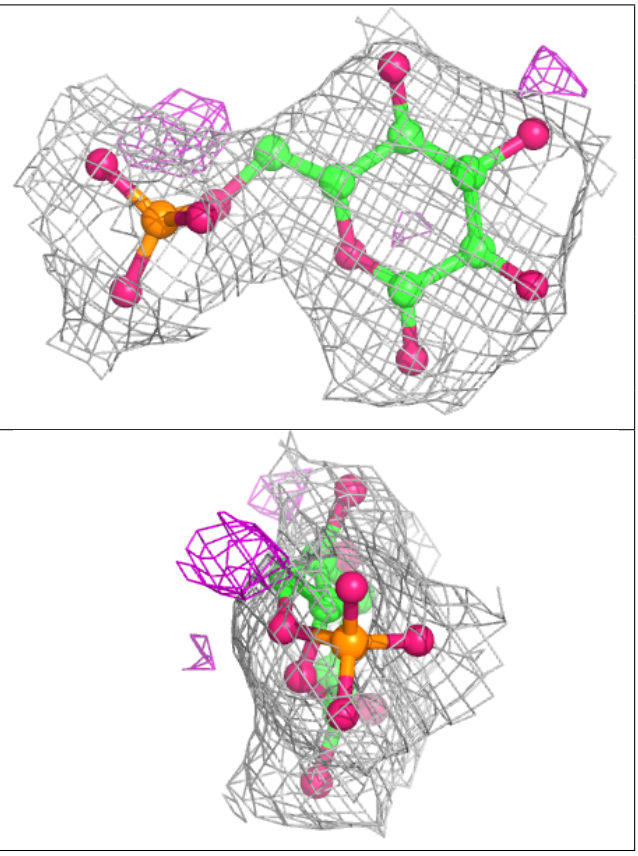
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BGC	A	1003	12/12	0.96	0.25	21,26,31,33	0
2	BGC	B	1001	12/12	0.96	0.18	23,29,31,31	0
4	CA	A	1005	1/1	0.96	0.18	10,10,10,10	0
2	BGC	B	1003	12/12	0.97	0.23	21,26,29,31	0
3	G6P	A	1002	16/16	0.97	0.14	28,31,36,37	0
3	G6P	A	1004	16/16	0.97	0.19	36,38,41,45	0
3	G6P	B	1002	16/16	0.97	0.11	30,34,42,44	0
2	BGC	A	1001	12/12	0.97	0.18	25,27,29,29	0
3	G6P	B	1004	16/16	0.98	0.18	35,37,42,43	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

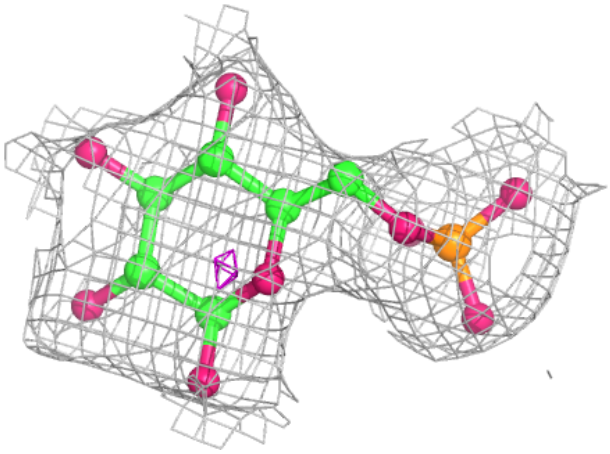
Electron density around G6P A 1002:

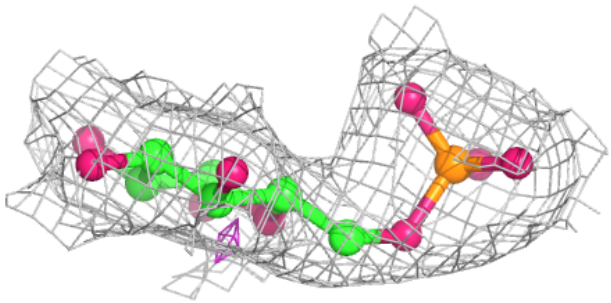
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

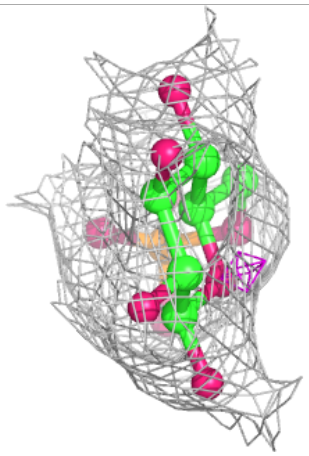


Electron density around G6P A 1004:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

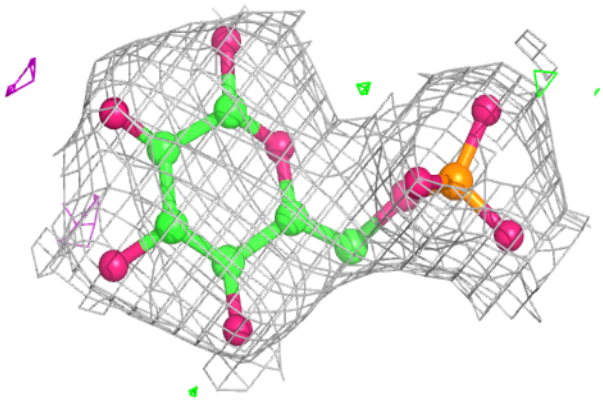


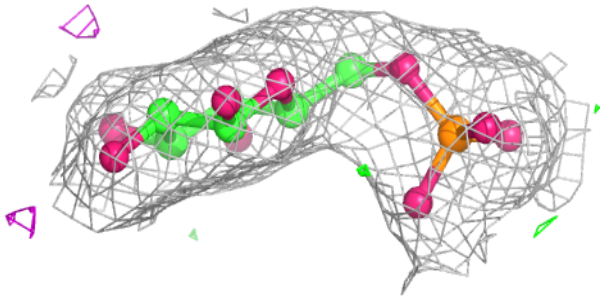


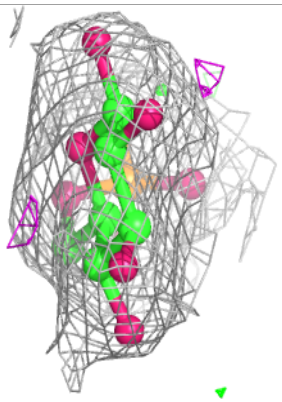


Electron density around G6P B 1002:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

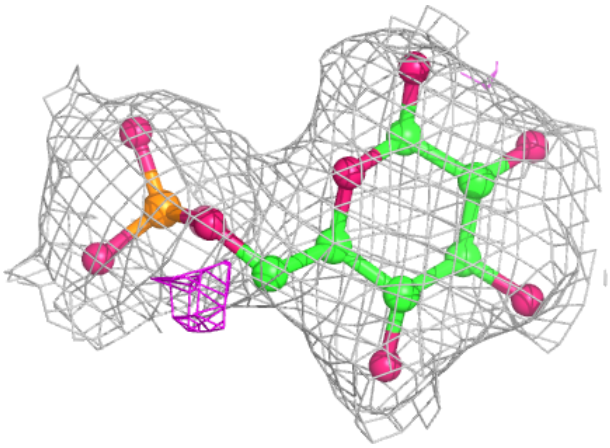


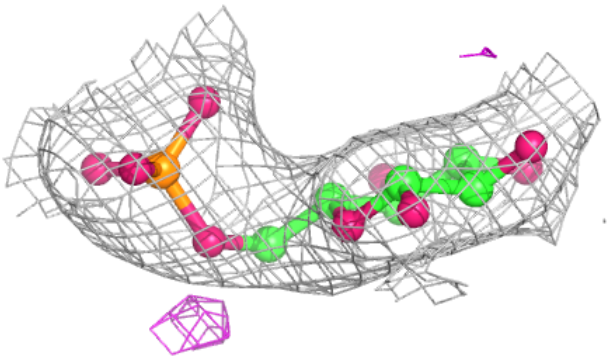


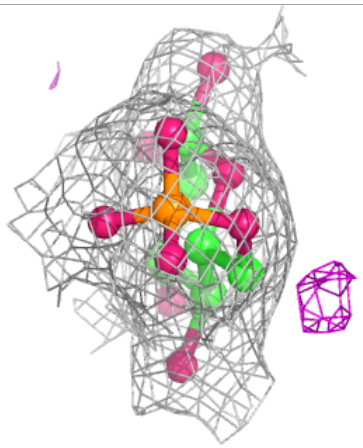


Electron density around G6P B 1004:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.