wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2024 – 05:14 PM EST

PDB ID	:	1HKB
Title	:	CRYSTAL STRUCTURE OF RECOMBINANT HUMAN BRAIN HEXOKINASE TYPE I COMPLEXED WITH GLUCOSE AND GLUCOSE-6-PHOSPHATE
Authors	:	Aleshin, A.E.; Zeng, C.; Burenkov, G.P.; Bartunik, H.D.; Fromm, H.J.; Honzatko, R.B.
Deposited on	:	1997-12-01
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 (1)) were used in the production of this report:

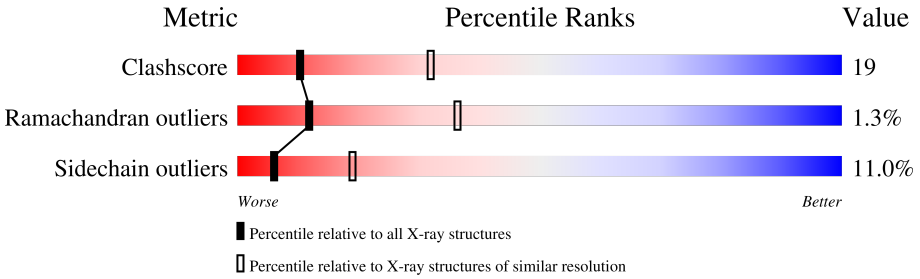
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

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%

Mol	Chain	Length	Quality of chain
1	A	917	59% 33% 5% ·
1	B	917	61% 32% 5% ·

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

2 Entry composition (i)

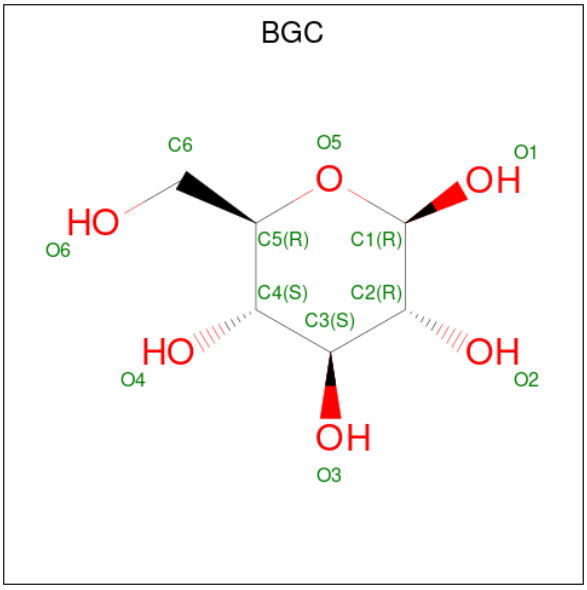
There are 5 unique types of molecules in this entry. The entry contains 14322 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 D-GLUCOSE 6-PHOSPHOTRANSFERASE.

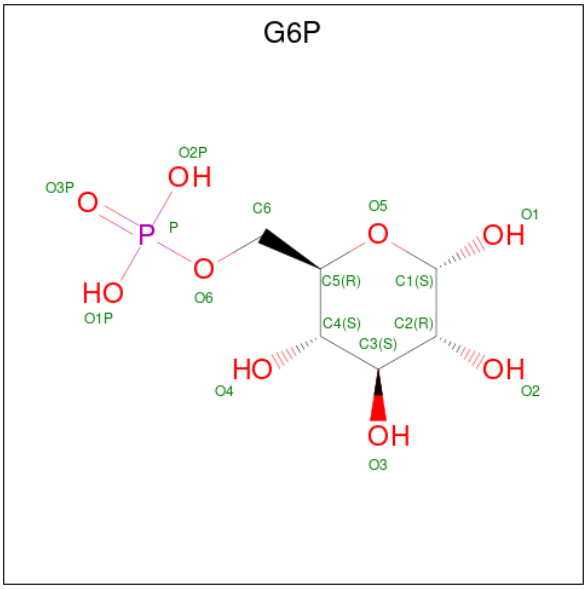
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	899	Total	C	N	O	S	0	0	0
			7032	4407	1240	1332	53			
1	B	899	Total	C	N	O	S	0	0	0
			7032	4407	1240	1332	53			

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
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 16 6 9 1	0	0
3	A	1	Total C O P 16 6 9 1	0	0
3	B	1	Total C O P 16 6 9 1	0	0
3	B	1	Total C O P 16 6 9 1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	B	2	Total Ca 2 2	0	0

- Molecule 5 is water.

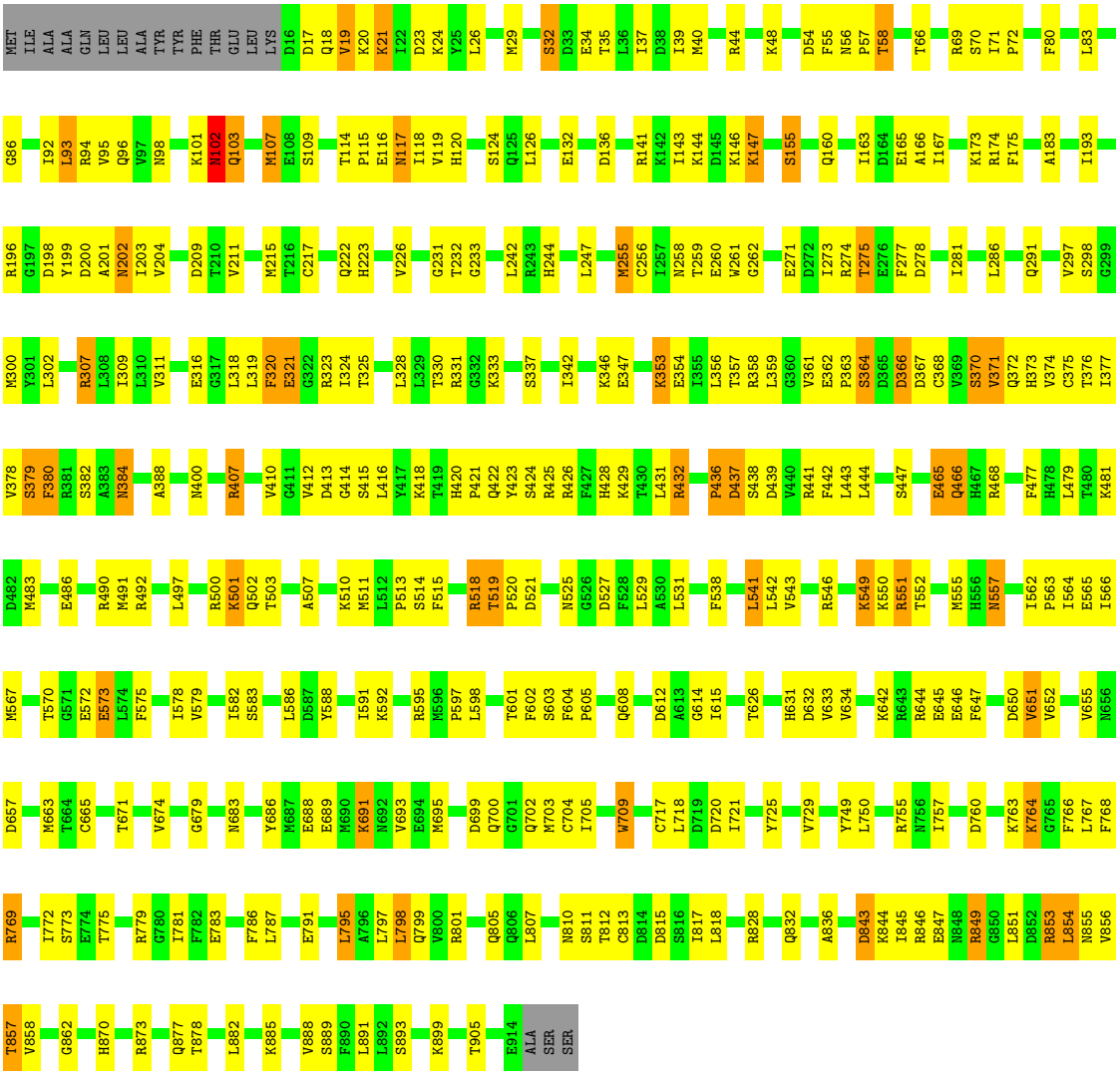
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	56	Total O 56 56	0	0
5	B	86	Total O 86 86	0	0

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.

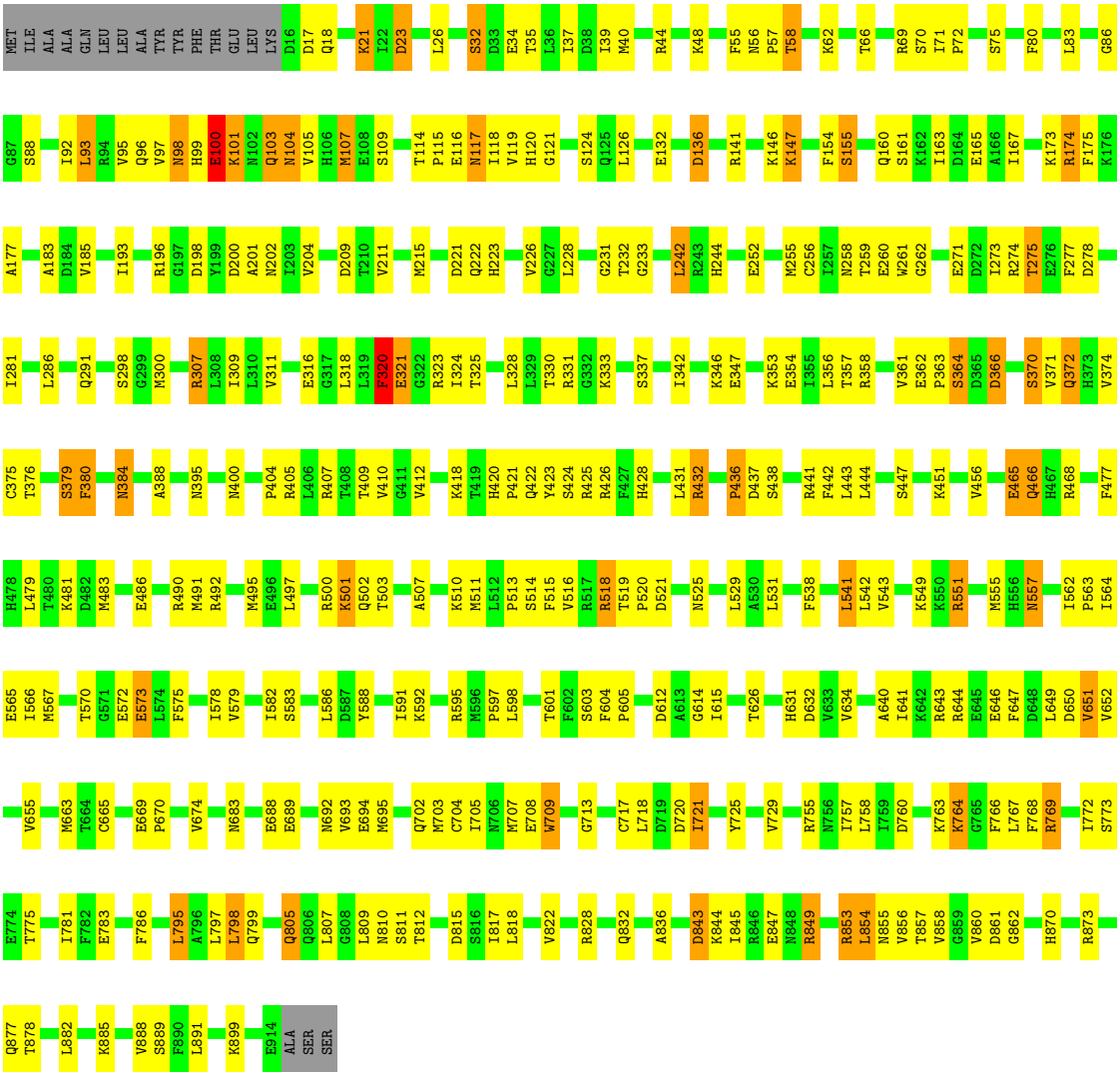
• Molecule 1: D-GLUCOSE 6-PHOSPHOTRANSFERASE

Chain A:



• Molecule 1: D-GLUCOSE 6-PHOSPHOTRANSFERASE

Chain B:


MET ILE ALA ALA GLN LEU LEU ALA TYR TYR PHE THR GLU LEU LYS D16 D17 Q18 K21 I22 D23 L26 S32 D33 E34 T35 L36 I37 D38 I39 M40 R44 K48 F55 N56 P57 T58 K62 T66 R69 S70 I71 P72 S75 F80 L83 G86
G87 S88 I92 L93 R94 V95 Q96 V97 N98 H99 E100 K101 N102 Q103 N104 V105 H106 M107 E108 S109 T114 P115 E116 N117 I118 V119 H120 G121 S124 Q125 L126 E132 D136 R141 K146 K147 F154 S155 Q160 S161 K162 I163 D164 E165 A166 I167 K173 R174 F175 K176
A177 A183 D184 V185 I193 R196 G197 D198 Y199 D200 A201 N202 I203 V204 D209 T210 V211 M215 D221 Q222 H223 V226 G227 L228 G231 T232 G233 L242 R243 H244 E252 M255 C256 I257 N258 T259 E260 W261 G262 E271 D272 I273 R274 T275 E276 F277 D278
I281 L286 Q291 S298 G299 M300 R307 L308 I309 L310 V311 E316 G317 L318 L319 F320 E321 G322 R323 I324 T325 L328 L329 T330 R331 G332 K333 S337 I342 K346 E347 K353 E354 I355 L356 T357 R358 V361 E362 P363 S364 D365 D366 S370 V371 Q372 H373 V374
C375 T376 S379 F380 N384 A388 N395 N400 P404 R405 L406 R407 T408 T409 V410 G411 V412 K418 T419 H420 P421 Q422 Y423 S424 R425 R426 F427 H428 L431 R432 P436 D437 S438 R441 F442 L443 L444 S447 K451 V456 E465 Q466 H467 R468 F477
H478 L479 T480 K481 D482 M483 E486 R490 M491 R492 M495 E496 L497 R500 K501 Q502 T503 A507 K510 M511 L512 P513 S514 F515 V516 R517 R518 T519 P520 D521 N525 L529 A530 L531 F538 L541 L542 V543 K549 K550 R551 M555 H556 N557 I562 P563 I564
E565 I566 M567 T570 G571 E572 E573 L574 F575 I578 V579 I582 S583 L586 D587 Y588 I591 K592 R596 M596 P597 L598 T601 F602 S603 F604 P605 D612 A613 G614 I615 T626 H631 D632 V633 V634 A640 I641 K642 R643 R644 E645 E646 F647 D648 L649 D650 V651 V652
V655 M663 T664 C665 E669 P670 V674 N683 E688 E689 N692 V693 E694 M695 Q702 M703 C704 I705 N706 M707 E708 W709 G713 C717 L718 D719 D720 I721 Y725 V729 R755 N756 I757 L758 I759 D760 K763 K764 G765 F766 L767 F768 R769 I772 S773
E774 T775 I781 F782 E783 F786 L795 A796 L797 L798 Q799 Q805 Q806 L807 G808 L809 N810 S811 T812 D815 S816 I817 L818 V822 R828 Q832 A836 D843 K844 I845 R846 E847 N848 R849 R853 L854 N855 V856 T857 V858 G859 V860 D861 G862 H870 R873
Q877 T878 L882 K885 V888 S889 F890 L891 K899 E914 ALA SER SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	83.00Å 122.00Å 123.00Å 90.00° 92.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.80 39.27 – 2.79	Depositor EDS
% Data completeness (in resolution range)	96.4 (8.00-2.80) 96.0 (39.27-2.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	3.79 (at 2.81Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.200 , 0.270 0.268 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	52.7	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 69.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.007 for -h,-l,-k 0.019 for h,-k,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	14322	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% 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

5.1 Standard geometry

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	0.38	0/7138	0.60	0/9606
1	B	0.39	0/7138	0.61	2/9606 (0.0%)
All	All	0.39	0/14276	0.61	2/19212 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	103	GLN	N-CA-C	-6.91	92.33	111.00
1	B	516	VAL	N-CA-C	-5.07	97.31	111.00

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	7032	0	7090	282	0
1	B	7032	0	7090	259	0
2	A	24	0	24	3	0
2	B	24	0	24	2	0
3	A	32	0	22	6	0
3	B	32	0	22	5	0
4	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	2	0	0	0	0
5	A	56	0	0	2	0
5	B	86	0	0	2	0
All	All	14322	0	14272	531	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 19.

The worst 5 of 531 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:691:LYS:H	1:A:691:LYS:HD2	1.01	1.13
1:A:307:ARG:HH21	1:A:331:ARG:HA	1.22	1.00
1:A:550:LYS:HE3	1:A:552:THR:HG21	1.47	0.97
1:B:307:ARG:HH21	1:B:331:ARG:HA	1.31	0.95
1:A:689:GLU:HB3	1:A:691:LYS:HD3	1.47	0.95

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	897/917 (98%)	810 (90%)	76 (8%)	11 (1%)	13	39
1	B	897/917 (98%)	811 (90%)	73 (8%)	13 (1%)	11	34
All	All	1794/1834 (98%)	1621 (90%)	149 (8%)	24 (1%)	12	36

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	346	LYS
1	B	100	GLU
1	B	104	ASN
1	B	346	LYS

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	774/788 (98%)	687 (89%)	87 (11%)	6	18
1	B	774/788 (98%)	691 (89%)	83 (11%)	6	20
All	All	1548/1576 (98%)	1378 (89%)	170 (11%)	6	19

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	286	LEU
1	B	557	ASN
1	B	337	SER
1	B	424	SER
1	B	603	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	525	ASN
1	B	806	GLN
1	B	556	HIS
1	B	702	GLN
1	B	870	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)

Of 12 ligands modelled in this entry, 4 are 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	$\#\|Z\| > 2$
2	BGC	A	920	-	12,12,12	0.43	0	17,17,17	0.79	0
2	BGC	B	918	-	12,12,12	0.40	0	17,17,17	0.73	0
3	G6P	B	919	-	16,16,16	0.81	1 (6%)	24,24,24	1.04	2 (8%)
3	G6P	B	921	-	16,16,16	0.80	0	24,24,24	0.98	2 (8%)
2	BGC	A	918	-	12,12,12	0.44	0	17,17,17	0.85	1 (5%)
3	G6P	A	919	-	16,16,16	0.72	0	24,24,24	0.91	1 (4%)
3	G6P	A	921	-	16,16,16	0.77	0	24,24,24	1.24	4 (16%)
2	BGC	B	920	-	12,12,12	0.36	0	17,17,17	0.58	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
2	BGC	A	920	-	-	1/2/22/22	0/1/1/1
2	BGC	B	918	-	-	2/2/22/22	0/1/1/1
3	G6P	B	921	-	1/1/6/6	1/6/26/26	0/1/1/1
3	G6P	B	919	-	1/1/6/6	2/6/26/26	0/1/1/1
2	BGC	A	918	-	-	0/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	G6P	A	919	-	1/1/6/6	3/6/26/26	0/1/1/1
3	G6P	A	921	-	1/1/6/6	0/6/26/26	0/1/1/1
2	BGC	B	920	-	-	2/2/22/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	919	G6P	P-O1P	-2.08	1.46	1.54

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	921	G6P	O1P-P-O6	2.58	113.60	106.73
3	A	921	G6P	C6-C5-C4	2.49	117.30	112.09
3	A	919	G6P	O2P-P-O6	2.43	113.19	106.73
3	A	921	G6P	O2P-P-O6	2.42	113.17	106.73
3	B	921	G6P	O1P-P-O6	2.33	112.94	106.73

All (4) chirality outliers are listed below:

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

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	919	G6P	C6-O6-P-O1P
3	A	919	G6P	C6-O6-P-O2P
3	A	919	G6P	C6-O6-P-O3P
2	B	920	BGC	C4-C5-C6-O6
2	B	918	BGC	C4-C5-C6-O6

There are no ring outliers.

8 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	920	BGC	1	0
2	B	918	BGC	1	0

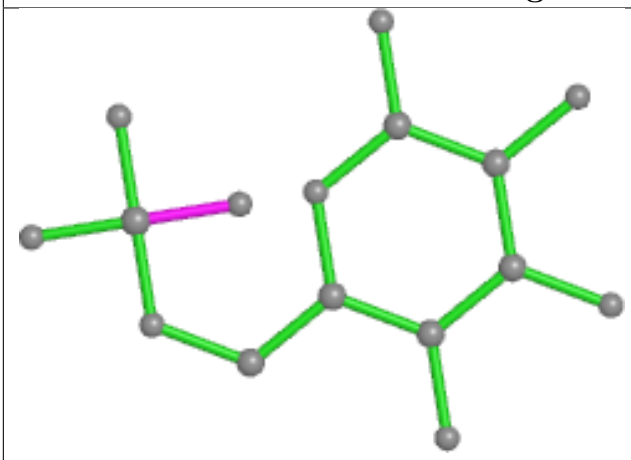
Continued on next page...

Continued from previous page...

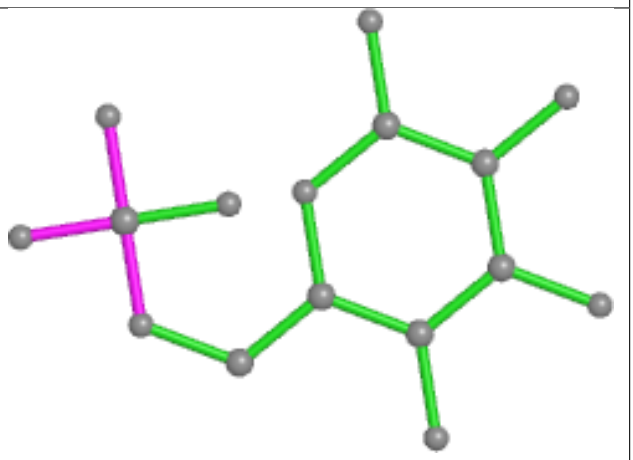
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	919	G6P	4	0
3	B	921	G6P	1	0
2	A	918	BGC	2	0
3	A	919	G6P	3	0
3	A	921	G6P	3	0
2	B	920	BGC	1	0

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.

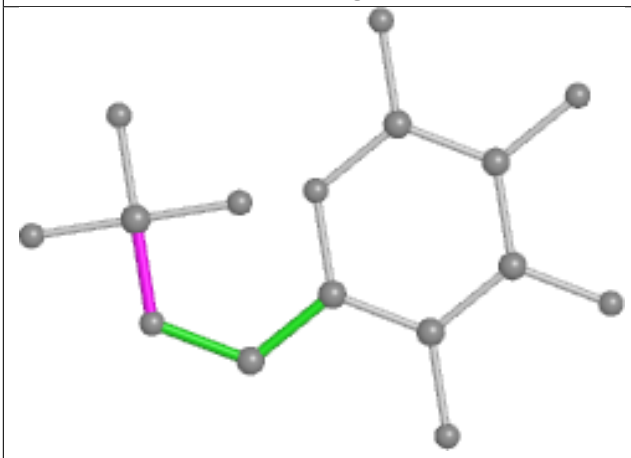
Ligand G6P B 919



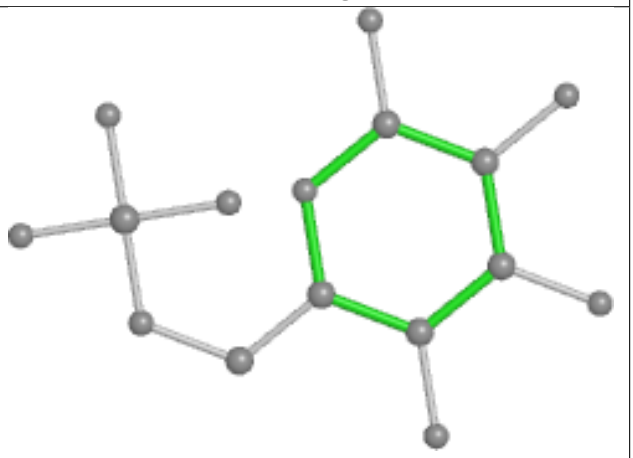
Bond lengths



Bond angles

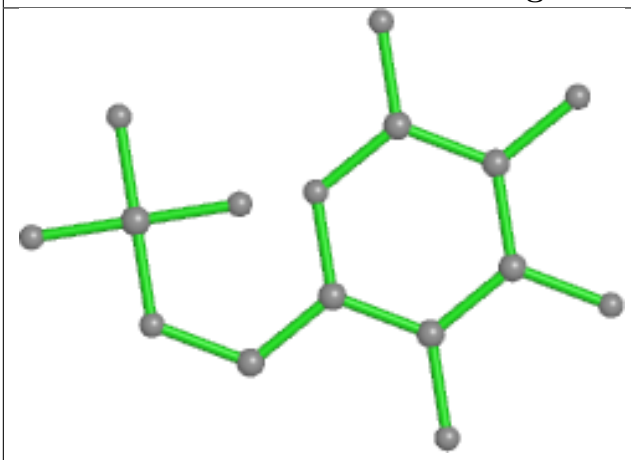


Torsions

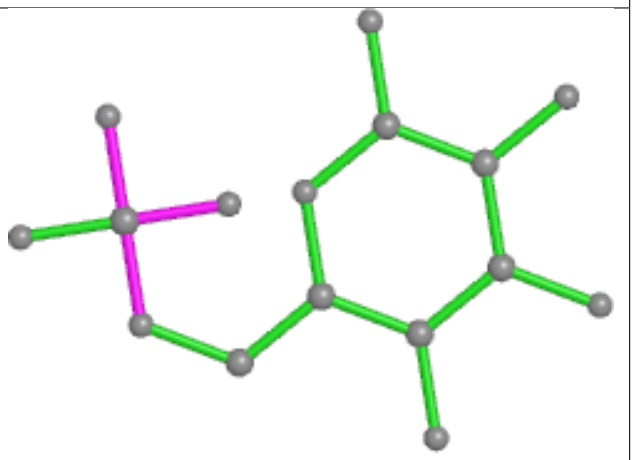


Rings

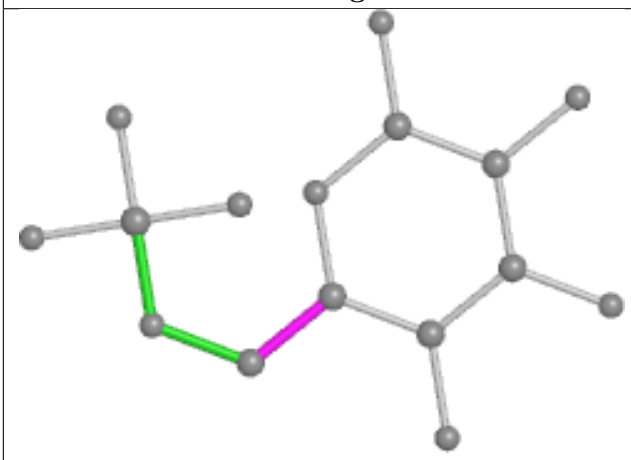
Ligand G6P B 921



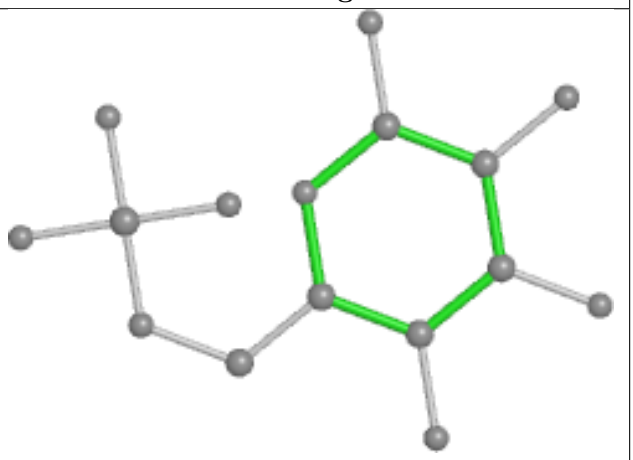
Bond lengths



Bond angles

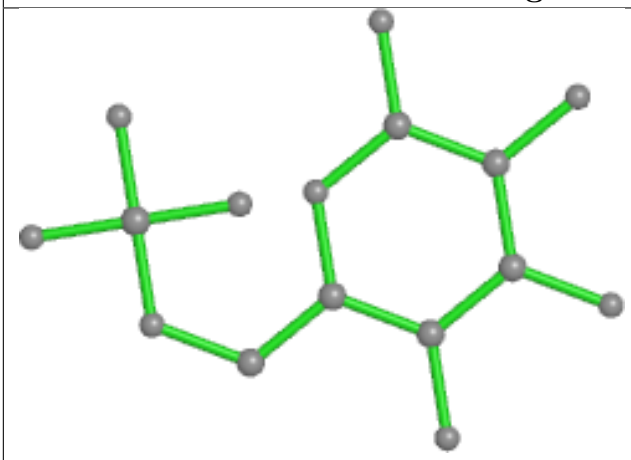


Torsions

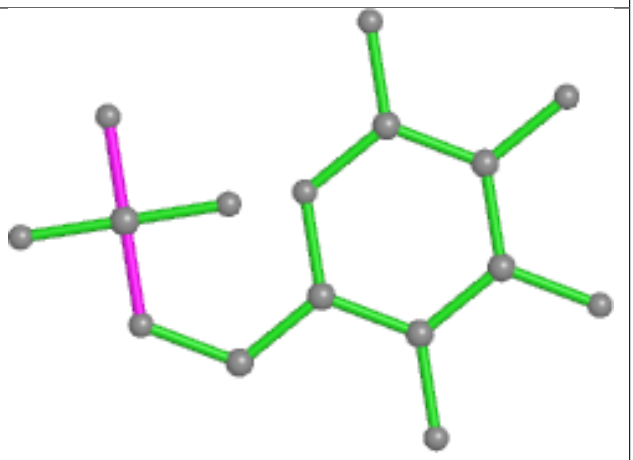


Rings

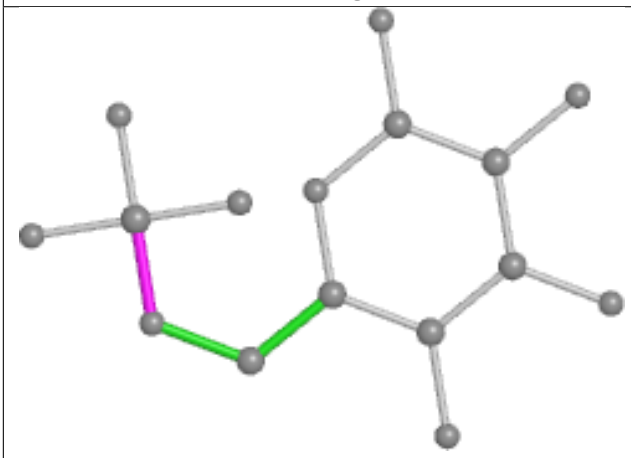
Ligand G6P A 919



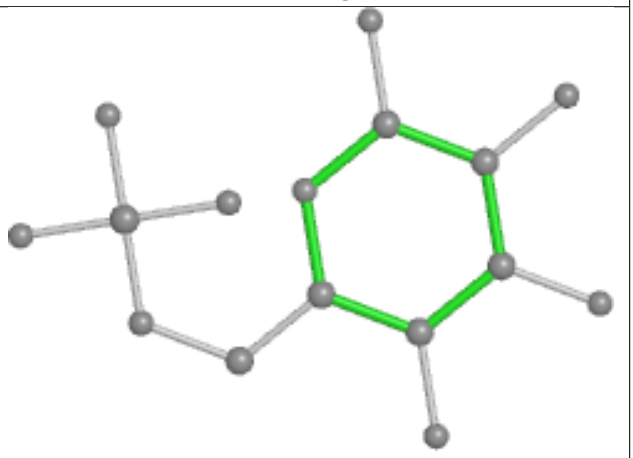
Bond lengths



Bond angles

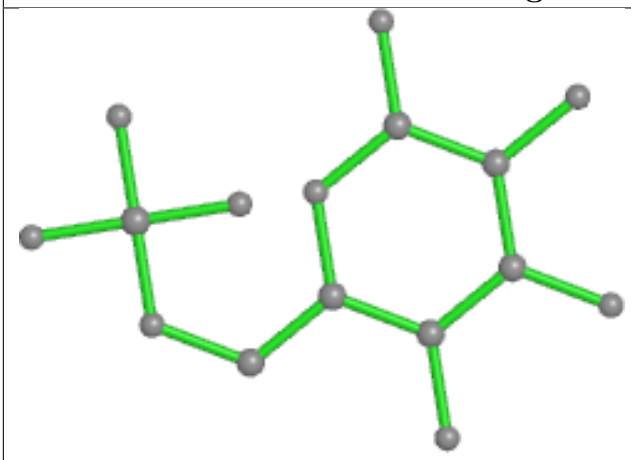


Torsions

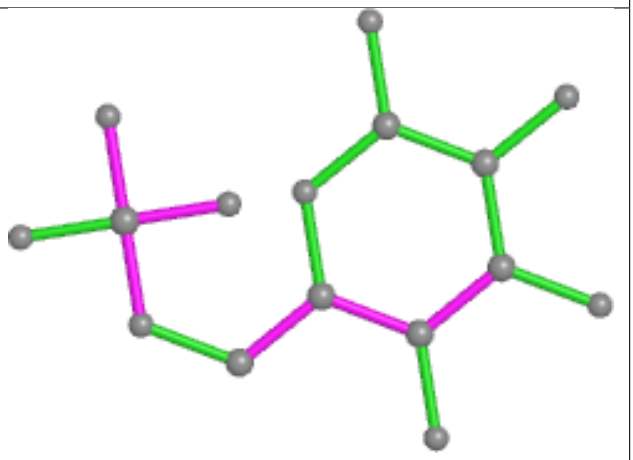


Rings

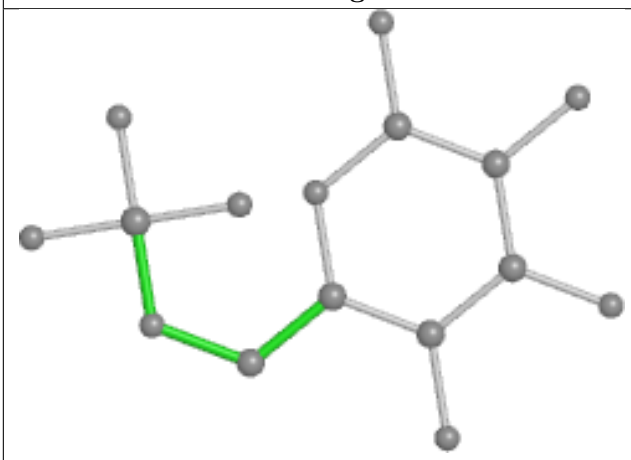
Ligand G6P A 921



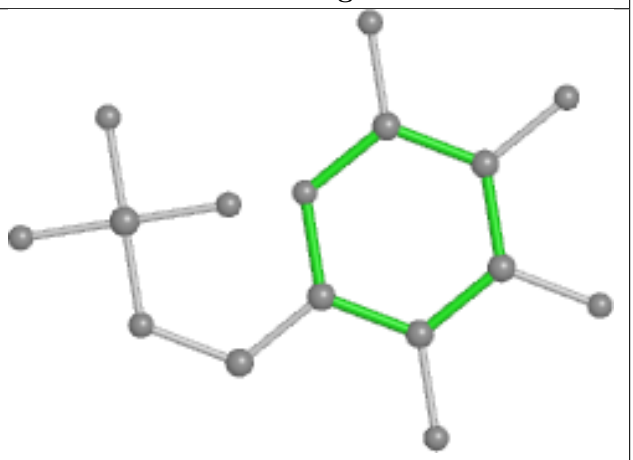
Bond lengths



Bond angles



Torsions



Rings

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

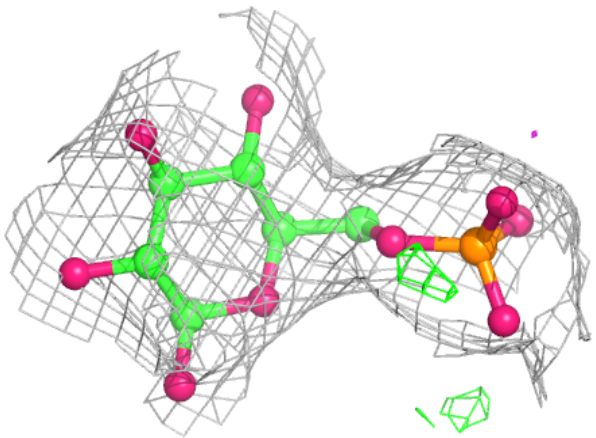
6.4 Ligands

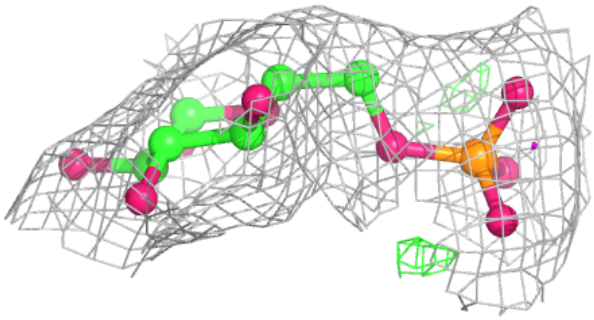
Unable to reproduce the depositors R factor - this section is therefore empty.

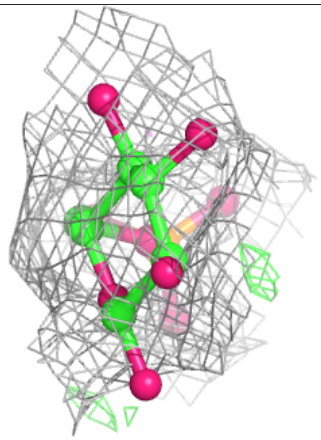
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around G6P A 919:

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

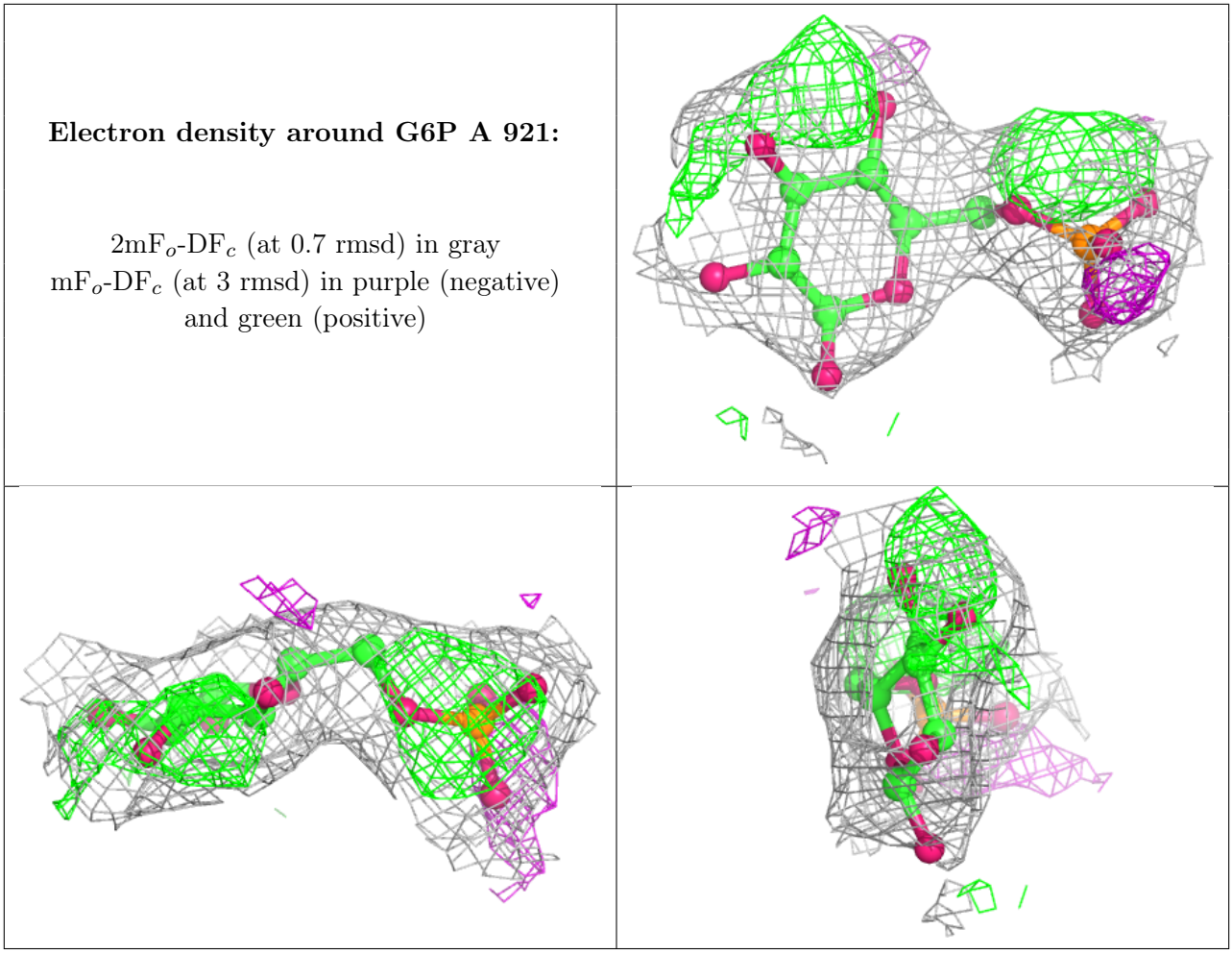






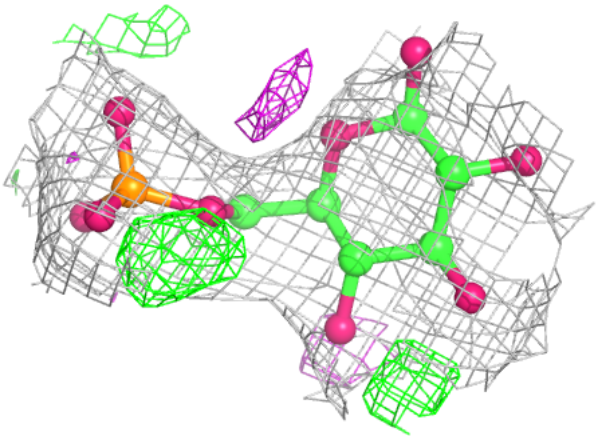
Electron density around G6P A 921:

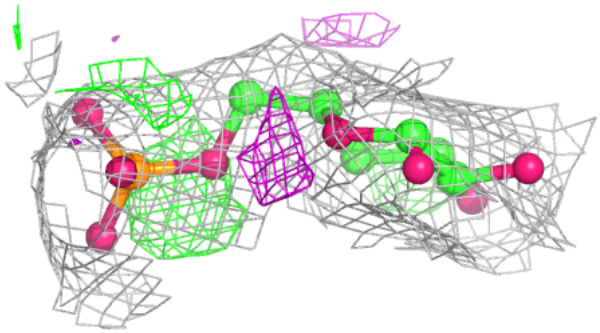
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

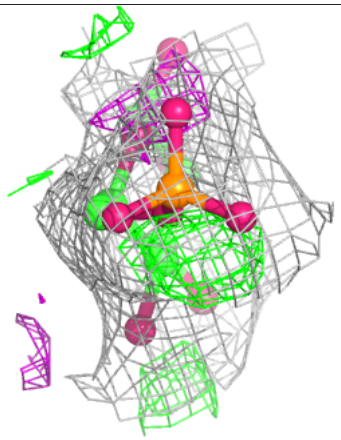


Electron density around G6P B 919:

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

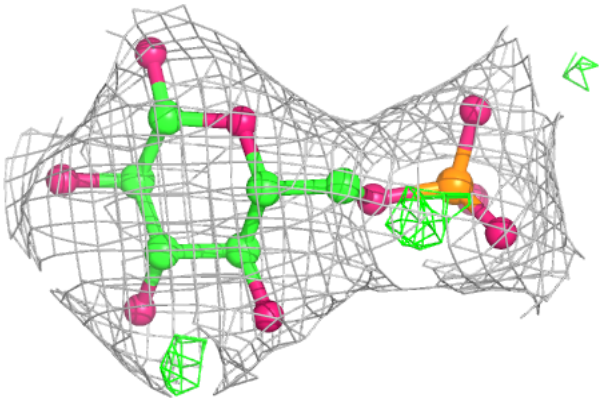


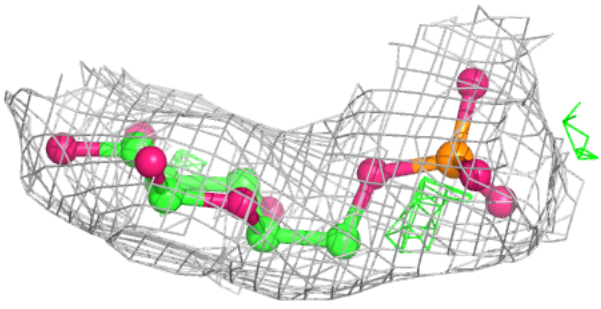


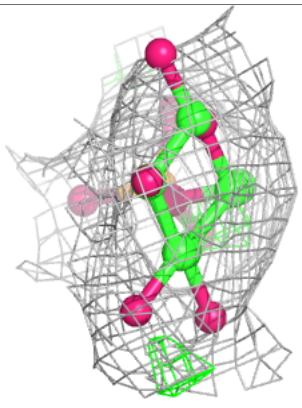


Electron density around G6P B 921:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.