



Full wwPDB X-ray Structure Validation Report ⓘ

Jul 12, 2021 – 11:42 am BST

PDB ID : 7OPG
Title : Crystal structure of CLK1 in complex with compound 2 (CC513)
Authors : Chaikuad, A.; Routier, S.; Bonnet, P.; Knapp, S.; Structural Genomics Consortium (SGC)
Deposited on : 2021-05-31
Resolution : 1.93 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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)
Xtriage (Phenix) : 1.13
EDS : 2.22
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) : 2.22

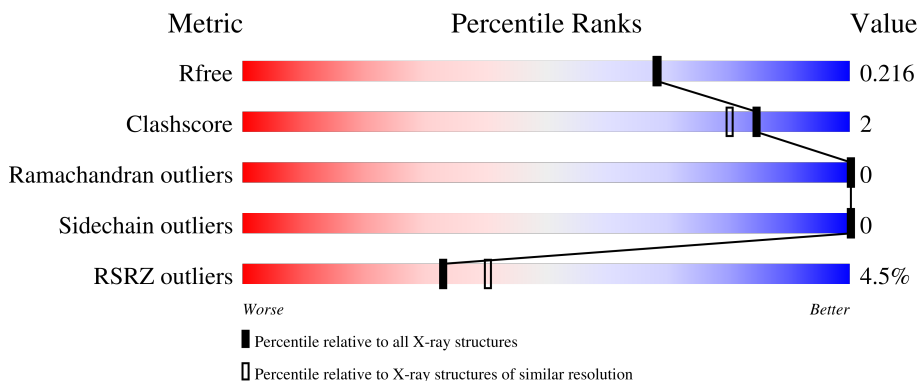
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

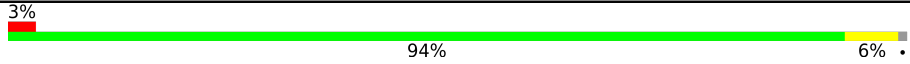
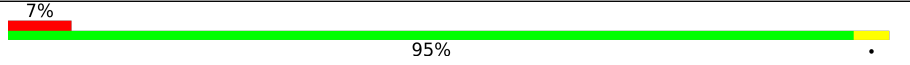
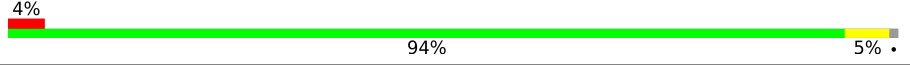
The reported resolution of this entry is 1.93 Å.

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))
R_{free}	130704	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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 $\leq 5\%$. 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	339	
1	B	339	
1	C	339	

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	GOL	A	504	-	-	X	-

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 9093 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.

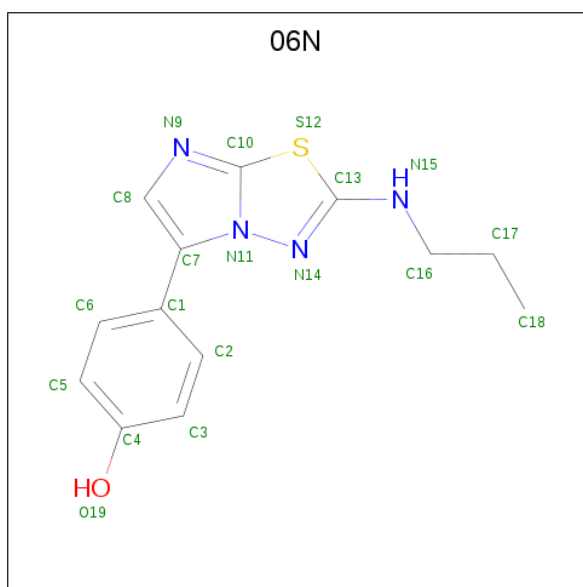
- Molecule 1 is a protein called Dual specificity protein kinase CLK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	337	2787	1780	486	504	17	0	4	0
1	B	338	2771	1769	483	502	17	0	2	0
1	C	337	2774	1772	484	501	17	0	3	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP P49759
A	0	MET	-	expression tag	UNP P49759
A	432	ALA	ARG	conflict	UNP P49759
B	-1	SER	-	expression tag	UNP P49759
B	0	MET	-	expression tag	UNP P49759
B	432	ALA	ARG	conflict	UNP P49759
C	-1	SER	-	expression tag	UNP P49759
C	0	MET	-	expression tag	UNP P49759
C	432	ALA	ARG	conflict	UNP P49759

- Molecule 2 is 4-[2-(propylamino)imidazo[2,1-b][1,3,4]thiadiazol-5-yl]phenol (three-letter code: 06N) (formula: C₁₃H₁₄N₄OS) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	19	13	4	1	1	0	0
2	B	1	19	13	4	1	1	0	0
2	C	1	19	13	4	1	1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



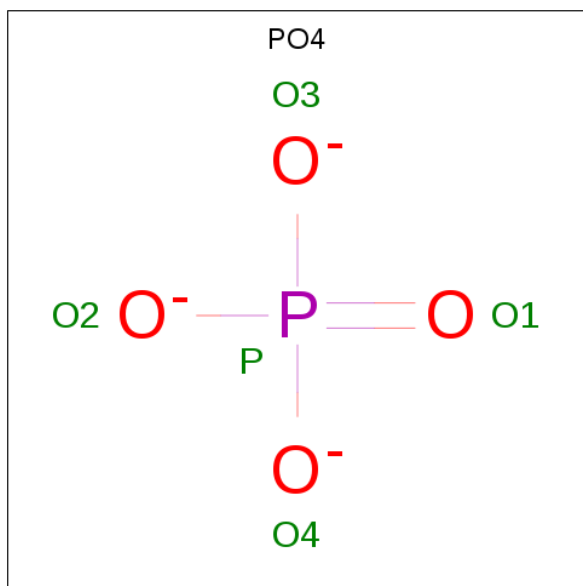
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	6	3	3	0	0

Continued on next page...

Continued from previous page...

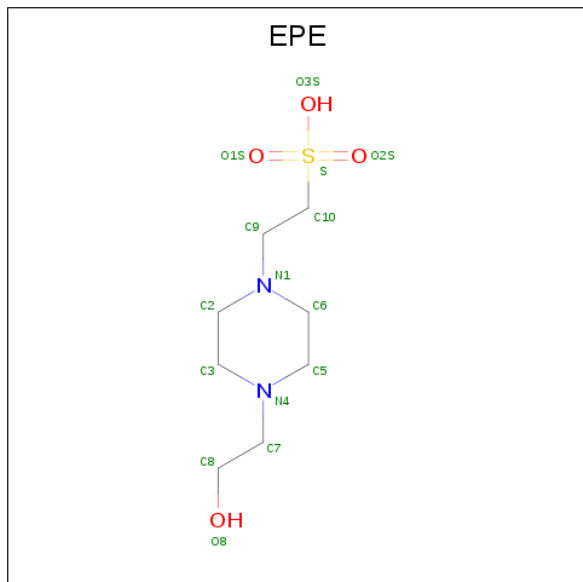
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	P	0	0
			5	4	1		
4	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
5	B	1	15	8	2	4	1	0	0

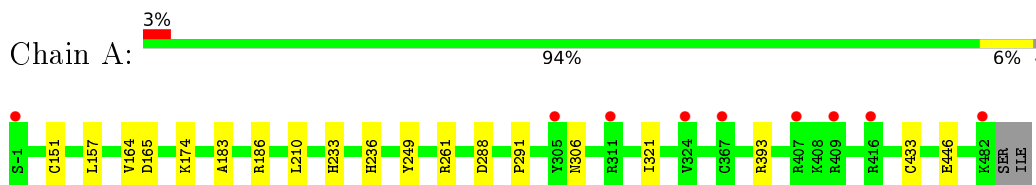
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	192	Total	O	0	0
			192	192		
6	B	222	Total	O	0	0
			222	222		
6	C	205	Total	O	0	0
			205	205		

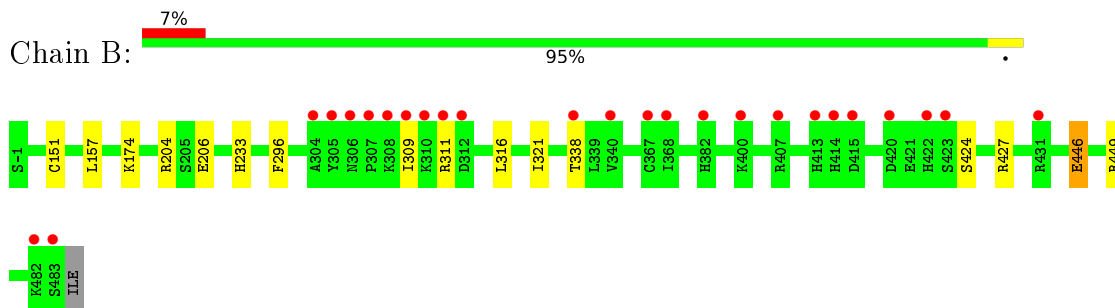
3 Residue-property plots [i](#)

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.

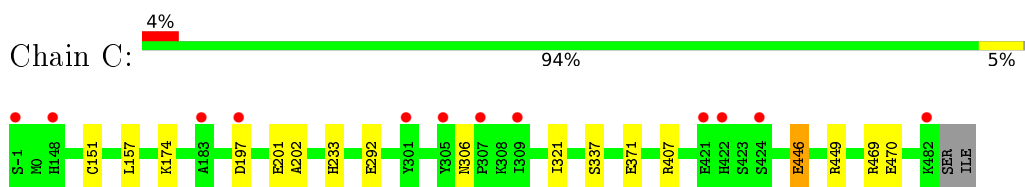
- Molecule 1: Dual specificity protein kinase CLK1



- Molecule 1: Dual specificity protein kinase CLK1



- Molecule 1: Dual specificity protein kinase CLK1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.39Å 116.93Å 91.80Å 90.00° 99.04° 90.00°	Depositor
Resolution (Å)	31.01 – 1.93 29.55 – 1.93	Depositor EDS
% Data completeness (in resolution range)	99.5 (31.01-1.93) 99.6 (29.55-1.93)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.72 (at 1.93Å)	Xtrriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.175 , 0.208 0.184 , 0.216	Depositor DCC
R_{free} test set	4356 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	29.0	Xtrriage
Anisotropy	0.244	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 41.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9093	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.01% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: 06N, PO4, GOL, EPE

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	0.75	1/2867 (0.0%)	0.75	1/3874 (0.0%)
1	B	0.82	1/2846 (0.0%)	0.78	0/3849
1	C	0.83	4/2854 (0.1%)	0.79	2/3858 (0.1%)
All	All	0.80	6/8567 (0.1%)	0.78	3/11581 (0.0%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	446	GLU	CD-OE1	6.32	1.32	1.25
1	A	446	GLU	CD-OE1	5.97	1.32	1.25
1	C	371	GLU	CD-OE2	5.87	1.32	1.25
1	C	337	SER	CB-OG	-5.26	1.35	1.42
1	C	292	GLU	CD-OE1	5.11	1.31	1.25
1	C	446	GLU	CD-OE1	5.07	1.31	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	407	ARG	NE-CZ-NH1	6.36	123.48	120.30
1	C	469	ARG	NE-CZ-NH1	6.14	123.37	120.30
1	A	288	ASP	CB-CG-OD1	5.14	122.93	118.30

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	2787	0	2749	16	0
1	B	2771	0	2706	12	0
1	C	2774	0	2733	8	0
2	A	19	0	0	0	0
2	B	19	0	0	0	0
2	C	19	0	0	0	0
3	A	24	0	32	5	0
3	B	24	0	32	2	0
3	C	12	0	16	1	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
5	B	15	0	18	1	0
6	A	192	0	0	2	0
6	B	222	0	0	6	0
6	C	205	0	0	1	0
All	All	9093	0	8286	38	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 2.

All (38) 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:233:HIS:ND1	6:A:601:HOH:O	2.23	0.71
1:B:296:PHE:CZ	1:B:316:LEU:HD11	2.25	0.71
1:B:424:SER:HB2	6:B:604:HOH:O	1.90	0.69
1:B:424:SER:CB	6:B:604:HOH:O	2.40	0.69
3:B:504:GOL:H32	6:B:605:HOH:O	1.95	0.65
1:A:261:ARG:H	3:A:504:GOL:C1	2.12	0.63
1:C:202:ALA:HB1	3:C:503:GOL:H32	1.80	0.63
1:B:309:ILE:O	1:B:311:ARG:HG2	2.00	0.61
1:A:261:ARG:HB2	3:A:504:GOL:H12	1.84	0.60
1:B:151:CYS:SG	1:B:174:LYS:NZ	2.63	0.59
1:A:261:ARG:H	3:A:504:GOL:H11	1.68	0.58
1:A:151:CYS:SG	1:A:174:LYS:NZ	2.61	0.57
1:C:151:CYS:SG	1:C:174:LYS:NZ	2.62	0.56
1:B:206:GLU:OE1	3:B:502:GOL:O1	2.25	0.54
1:B:338:THR:HG23	6:B:649:HOH:O	2.08	0.52
1:A:164:VAL:O	1:A:165[A]:ASP:HB3	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:157:LEU:HD22	1:A:233:HIS:CD2	2.47	0.49
1:B:204:ARG:NH1	6:B:611:HOH:O	2.46	0.49
1:C:321:ILE:C	1:C:321:ILE:HD12	2.34	0.48
1:A:261:ARG:N	3:A:504:GOL:H11	2.28	0.48
1:B:427:ARG:HB3	6:B:614:HOH:O	2.13	0.47
5:B:507:EPE:O3S	5:B:507:EPE:H62	2.15	0.47
1:C:197:ASP:O	1:C:201:GLU:HG2	2.14	0.47
1:A:236:HIS:HE1	6:A:761:HOH:O	1.98	0.46
1:C:470:GLU:HG2	6:C:608:HOH:O	2.16	0.45
1:B:157:LEU:HD22	1:B:233:HIS:CD2	2.52	0.44
1:C:446:GLU:OE2	1:C:449:ARG:NH1	2.46	0.44
1:A:261:ARG:HB2	3:A:504:GOL:C1	2.46	0.44
1:A:249:TYR:HB2	1:A:291:PRO:HB2	2.00	0.43
1:A:393:ARG:HG2	1:A:433:CYS:HB2	2.00	0.43
1:A:183:ALA:HB1	1:A:186:ARG:HD2	2.00	0.43
1:A:321:ILE:HD12	1:A:321:ILE:C	2.38	0.43
1:B:446:GLU:OE2	1:B:449:ARG:NH1	2.42	0.43
1:C:157:LEU:HD22	1:C:233:HIS:CD2	2.55	0.42
1:C:306:ASN:OD1	1:C:306:ASN:C	2.58	0.41
1:A:306:ASN:OD1	1:A:306:ASN:C	2.60	0.41
1:B:321:ILE:C	1:B:321:ILE:HD12	2.41	0.41
1:A:210:LEU:HD23	1:A:210:LEU:HA	1.95	0.40

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	339/339 (100%)	329 (97%)	10 (3%)	0	100	100
1	B	338/339 (100%)	327 (97%)	11 (3%)	0	100	100
1	C	338/339 (100%)	329 (97%)	9 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1015/1017 (100%)	985 (97%)	30 (3%)	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	309/308 (100%)	309 (100%)	0	100	100
1	B	304/308 (99%)	304 (100%)	0	100	100
1	C	307/308 (100%)	307 (100%)	0	100	100
All	All	920/924 (100%)	920 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

16 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	GOL	A	503	-	5,5,5	0.46	0	5,5,5	0.41	0
3	GOL	B	505	-	5,5,5	0.56	0	5,5,5	1.13	0
3	GOL	C	503	-	5,5,5	0.47	0	5,5,5	1.16	0
3	GOL	A	504	-	5,5,5	0.89	0	5,5,5	1.33	1 (20%)
3	GOL	A	505	-	5,5,5	0.39	0	5,5,5	0.58	0
3	GOL	A	502	-	5,5,5	0.37	0	5,5,5	0.92	0
3	GOL	B	503	-	5,5,5	0.57	0	5,5,5	0.77	0
3	GOL	B	504	-	5,5,5	0.52	0	5,5,5	0.61	0
2	06N	C	501	-	14,21,21	0.83	1 (7%)	16,29,29	0.90	1 (6%)
3	GOL	B	502	-	5,5,5	0.39	0	5,5,5	0.86	0
2	06N	A	501	-	14,21,21	0.63	0	16,29,29	0.88	0
3	GOL	C	502	-	5,5,5	0.64	0	5,5,5	0.87	0
4	PO4	B	506	-	4,4,4	0.88	0	6,6,6	0.62	0
4	PO4	C	504	-	4,4,4	1.15	0	6,6,6	0.30	0
5	EPE	B	507	-	15,15,15	1.95	1 (6%)	18,20,20	1.91	6 (33%)
2	06N	B	501	-	14,21,21	0.68	1 (7%)	16,29,29	0.88	1 (6%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	503	-	-	2/4/4/4	-
3	GOL	B	505	-	-	3/4/4/4	-
3	GOL	C	503	-	-	4/4/4/4	-
3	GOL	A	504	-	-	0/4/4/4	-
3	GOL	A	505	-	-	4/4/4/4	-
3	GOL	A	502	-	-	2/4/4/4	-
3	GOL	B	503	-	-	1/4/4/4	-
3	GOL	B	504	-	-	2/4/4/4	-
2	06N	C	501	-	-	1/6/8/8	0/3/3/3
3	GOL	B	502	-	-	1/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	06N	A	501	-	-	0/6/8/8	0/3/3/3
3	GOL	C	502	-	-	0/4/4/4	-
5	EPE	B	507	-	-	5/9/19/19	0/1/1/1
2	06N	B	501	-	-	1/6/8/8	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	507	EPE	C10-S	-6.97	1.67	1.77
2	C	501	06N	C1-C7	-2.16	1.45	1.48
2	B	501	06N	C10-N9	-2.11	1.30	1.33

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	507	EPE	O3S-S-C10	3.77	111.87	105.77
5	B	507	EPE	C6-C5-N4	-2.81	104.88	110.64
5	B	507	EPE	C2-C3-N4	-2.80	104.90	110.64
5	B	507	EPE	O1S-S-C10	2.74	110.22	106.92
2	B	501	06N	C1-C7-N11	2.59	127.36	123.44
2	C	501	06N	C1-C7-N11	2.58	127.35	123.44
5	B	507	EPE	C6-N1-C2	2.27	113.95	108.83
5	B	507	EPE	O2S-S-O1S	-2.19	106.38	113.95
3	A	504	GOL	O2-C2-C1	-2.19	99.50	109.12

There are no chirality outliers.

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	GOL	O1-C1-C2-C3
3	A	503	GOL	O1-C1-C2-C3
3	B	505	GOL	O1-C1-C2-C3
3	C	503	GOL	O1-C1-C2-C3
5	B	507	EPE	C10-C9-N1-C6
5	B	507	EPE	S-C10-C9-N1
5	B	507	EPE	C9-C10-S-O2S
3	B	505	GOL	O1-C1-C2-O2
3	A	505	GOL	O1-C1-C2-C3
3	A	505	GOL	C1-C2-C3-O3
3	B	502	GOL	C1-C2-C3-O3
3	B	504	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

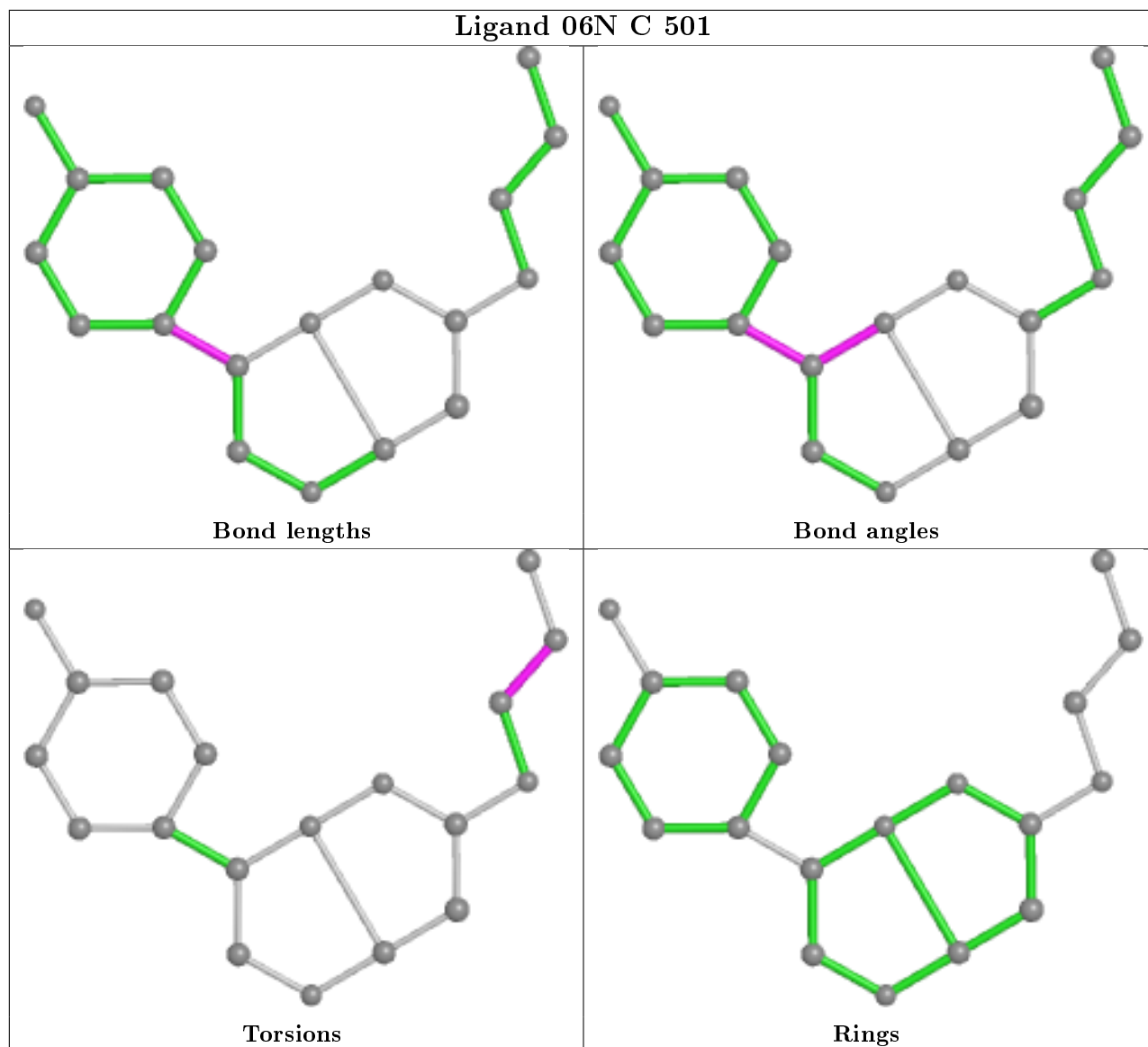
Mol	Chain	Res	Type	Atoms
3	C	503	GOL	C1-C2-C3-O3
3	A	502	GOL	O1-C1-C2-O2
3	A	503	GOL	O1-C1-C2-O2
3	A	505	GOL	O1-C1-C2-O2
3	C	503	GOL	O2-C2-C3-O3
3	B	504	GOL	O1-C1-C2-O2
2	C	501	06N	N15-C16-C17-C18
3	C	503	GOL	O1-C1-C2-O2
3	B	503	GOL	O1-C1-C2-C3
5	B	507	EPE	C9-C10-S-O1S
3	A	505	GOL	O2-C2-C3-O3
3	B	505	GOL	C1-C2-C3-O3
2	B	501	06N	N15-C16-C17-C18
5	B	507	EPE	N4-C7-C8-O8

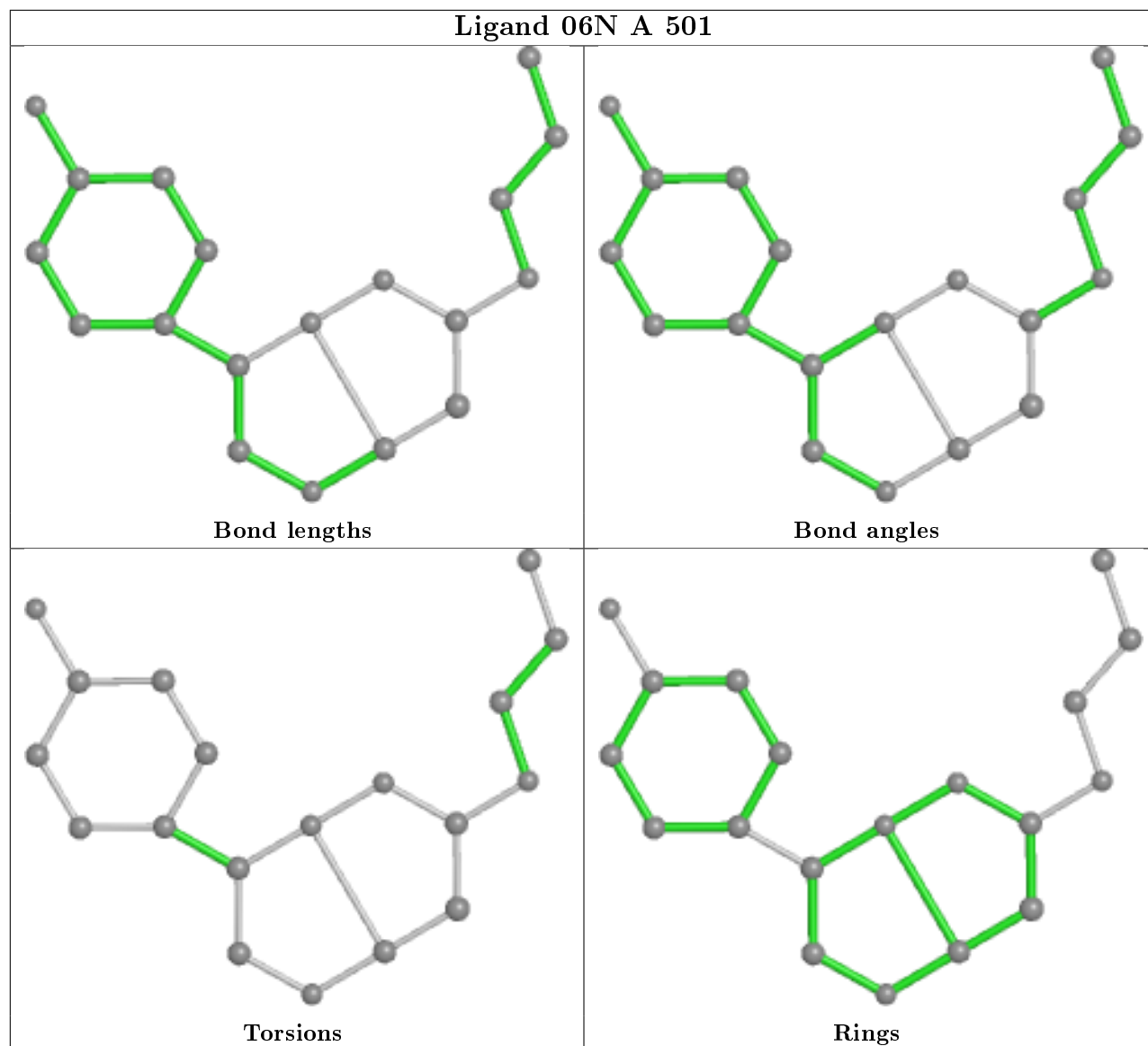
There are no ring outliers.

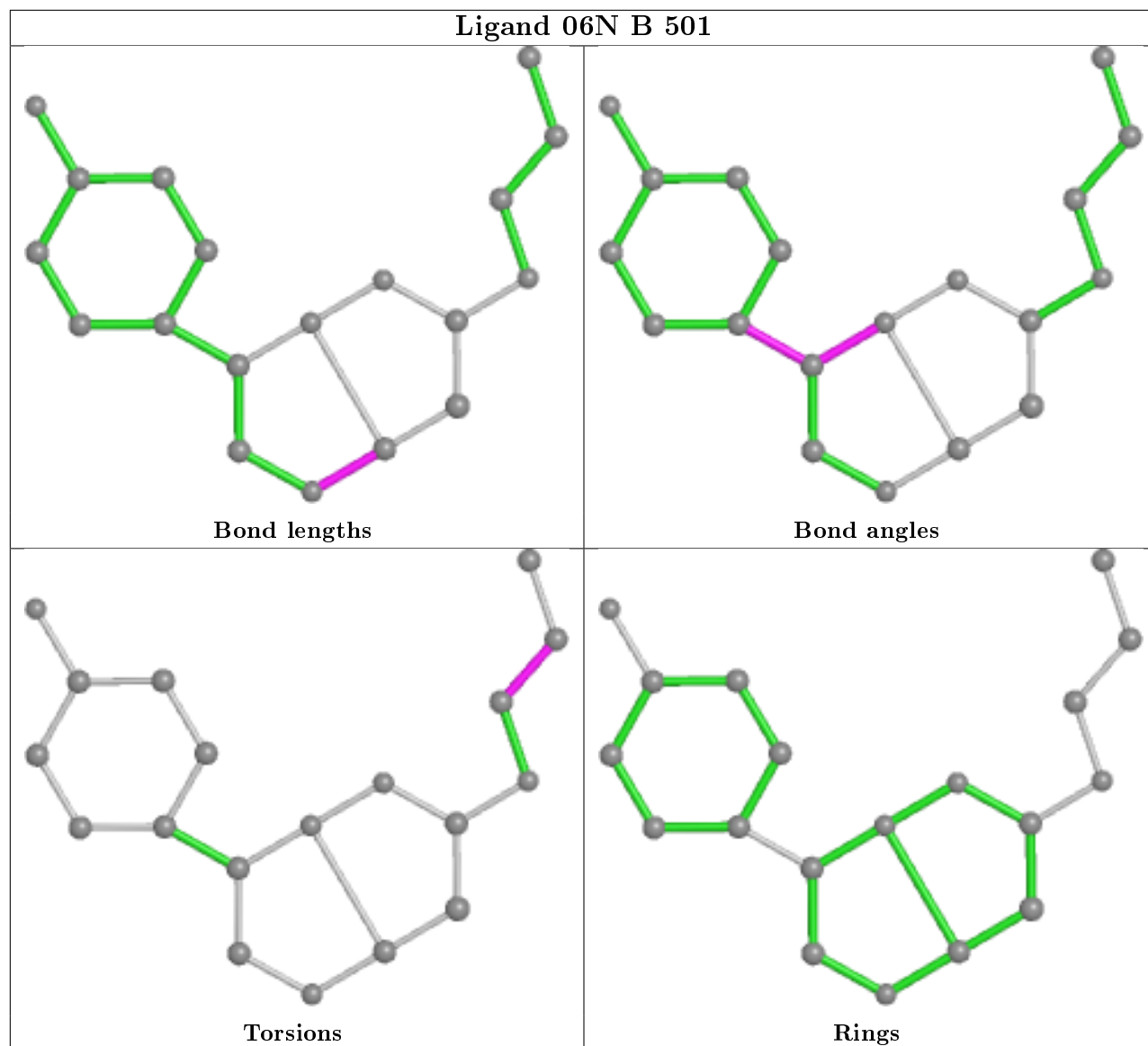
5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	503	GOL	1	0
3	A	504	GOL	5	0
3	B	504	GOL	1	0
3	B	502	GOL	1	0
5	B	507	EPE	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.







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

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	337/339 (99%)	0.02	9 (2%) 54 61	21, 38, 68, 93	0
1	B	338/339 (99%)	0.15	25 (7%) 14 20	19, 35, 86, 107	0
1	C	337/339 (99%)	-0.01	12 (3%) 42 50	21, 37, 73, 98	0
All	All	1012/1017 (99%)	0.06	46 (4%) 33 40	19, 36, 76, 107	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	483	SER	6.7
1	B	305	TYR	6.1
1	B	414	HIS	5.8
1	B	309	ILE	5.3
1	B	306	ASN	5.3
1	C	307	PRO	4.6
1	A	-1	SER	4.5
1	A	482	LYS	4.4
1	B	307	PRO	4.2
1	B	413	HIS	4.1
1	B	304	ALA	4.0
1	B	312	ASP	3.8
1	C	482	LYS	3.7
1	B	310	LYS	3.7
1	B	311	ARG	3.5
1	B	340	VAL	3.5
1	C	301	TYR	3.4
1	C	305	TYR	3.3
1	C	183	ALA	3.2
1	B	308	LYS	3.2
1	C	-1	SER	3.2
1	B	482	LYS	2.9
1	B	338	THR	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	407	ARG	2.8
1	C	309	ILE	2.8
1	A	305	TYR	2.7
1	A	324	VAL	2.6
1	B	420	ASP	2.5
1	C	422	HIS	2.5
1	B	368	ILE	2.5
1	B	422	HIS	2.4
1	C	424	SER	2.4
1	A	416	ARG	2.3
1	B	400	LYS	2.3
1	C	148	HIS	2.3
1	B	382	HIS	2.3
1	B	415	ASP	2.2
1	B	367	CYS	2.2
1	B	407	ARG	2.2
1	C	197	ASP	2.2
1	A	367	CYS	2.1
1	B	423	SER	2.1
1	A	409	ARG	2.1
1	C	421	GLU	2.0
1	B	431	ARG	2.0
1	A	311	ARG	2.0

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.

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	GOL	B	504	6/6	0.74	0.15	52,63,66,66	0
3	GOL	A	505	6/6	0.82	0.26	50,59,64,67	0
5	EPE	B	507	15/15	0.82	0.16	65,79,88,92	0
3	GOL	C	502	6/6	0.83	0.18	39,52,53,57	0
3	GOL	A	504	6/6	0.84	0.17	38,43,49,52	0
3	GOL	C	503	6/6	0.89	0.14	45,49,49,50	0
3	GOL	B	502	6/6	0.89	0.14	47,49,52,55	0
3	GOL	B	505	6/6	0.91	0.17	39,56,61,63	0
3	GOL	A	503	6/6	0.93	0.17	48,51,55,62	0
3	GOL	B	503	6/6	0.95	0.15	26,41,51,54	0
4	PO4	C	504	5/5	0.95	0.11	60,63,77,78	0
3	GOL	A	502	6/6	0.95	0.18	31,52,54,65	0
4	PO4	B	506	5/5	0.96	0.20	52,59,66,76	0
2	06N	B	501	19/19	0.97	0.12	17,22,39,43	0
2	06N	C	501	19/19	0.97	0.11	21,24,58,60	0
2	06N	A	501	19/19	0.97	0.11	22,27,48,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.