Full wwPDB X-ray Structure Validation Report (i)

Aug 19, 2023 – 09:21 PM EDT

PDB ID	:	2GX5
Title	:	N-terminal GAF domain of transcriptional pleiotropic repressor CodY
Authors	:	Wilkinson, A.J.; Levdikov, V.M.; Blagova, E.V.
Deposited on	:	2006-05-08
Resolution	:	1.74 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 (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:

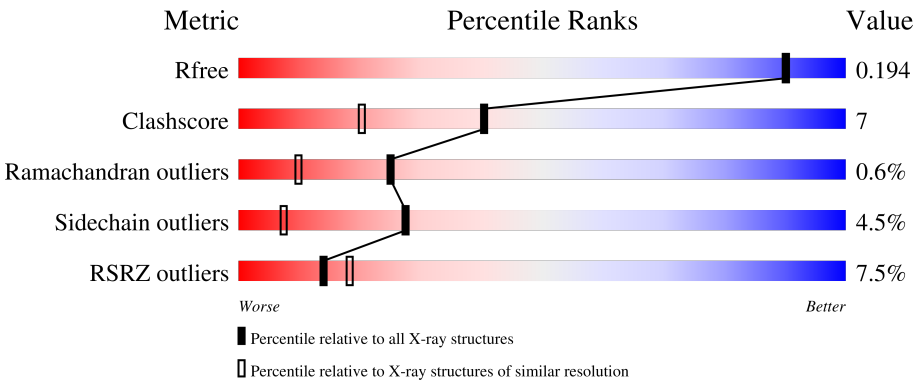
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.74 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	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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	170	7%, 75%, 16%, • 8%
1	B	170	9%, 79%, 14%, • 6%
1	C	170	8%, 75%, 18%, • 5%
1	D	170	5%, 81%, 11%, • 7%

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	PGE	A	515	-	-	-	X
7	PCG	C	514	X	-	-	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 5983 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 GTP-sensing transcriptional pleiotropic repressor codY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	157	1347	848	232	259	8	0	15	0
1	B	160	1368	860	235	265	8	0	15	0
1	C	161	1370	860	236	267	7	0	13	0
1	D	158	1339	843	232	256	8	0	13	0

There are 44 discrepancies between the modelled and reference sequences:

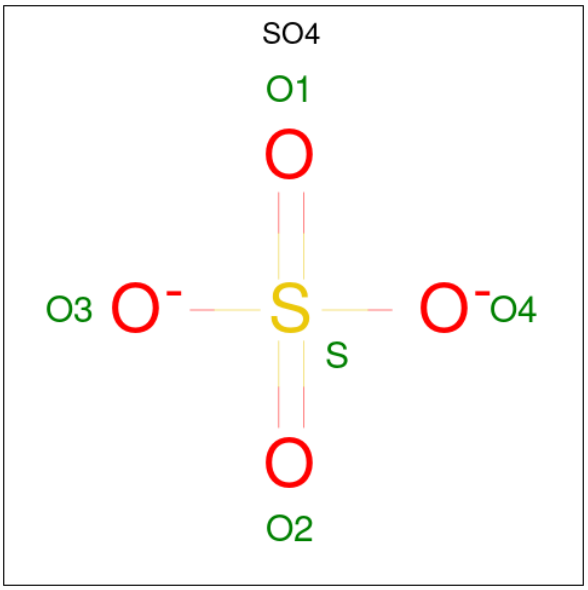
Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	initiating methionine	UNP P39779
A	-8	GLY	-	cloning artifact	UNP P39779
A	-7	SER	-	cloning artifact	UNP P39779
A	-6	SER	-	cloning artifact	UNP P39779
A	-5	HIS	-	expression tag	UNP P39779
A	-4	HIS	-	expression tag	UNP P39779
A	-3	HIS	-	expression tag	UNP P39779
A	-2	HIS	-	expression tag	UNP P39779
A	-1	HIS	-	expression tag	UNP P39779
A	0	HIS	-	expression tag	UNP P39779
A	1	MET	-	cloning artifact	UNP P39779
B	-9	MET	-	initiating methionine	UNP P39779
B	-8	GLY	-	cloning artifact	UNP P39779
B	-7	SER	-	cloning artifact	UNP P39779
B	-6	SER	-	cloning artifact	UNP P39779
B	-5	HIS	-	expression tag	UNP P39779
B	-4	HIS	-	expression tag	UNP P39779
B	-3	HIS	-	expression tag	UNP P39779
B	-2	HIS	-	expression tag	UNP P39779
B	-1	HIS	-	expression tag	UNP P39779
B	0	HIS	-	expression tag	UNP P39779

Continued on next page...

Continued from previous page...

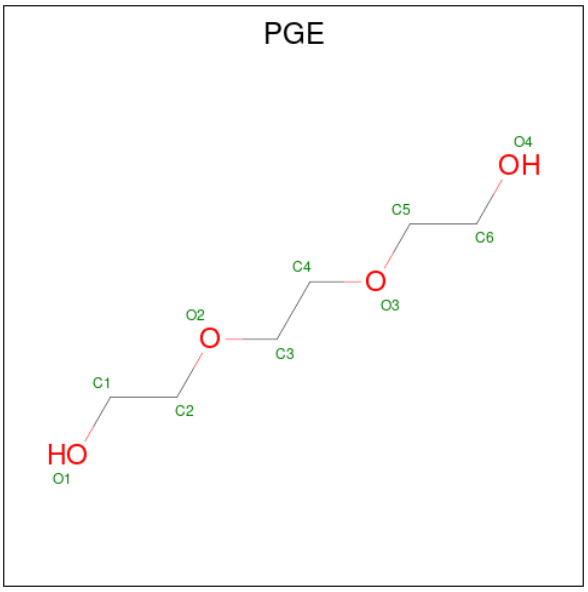
Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	cloning artifact	UNP P39779
C	-9	MET	-	initiating methionine	UNP P39779
C	-8	GLY	-	cloning artifact	UNP P39779
C	-7	SER	-	cloning artifact	UNP P39779
C	-6	SER	-	cloning artifact	UNP P39779
C	-5	HIS	-	expression tag	UNP P39779
C	-4	HIS	-	expression tag	UNP P39779
C	-3	HIS	-	expression tag	UNP P39779
C	-2	HIS	-	expression tag	UNP P39779
C	-1	HIS	-	expression tag	UNP P39779
C	0	HIS	-	expression tag	UNP P39779
C	1	MET	-	cloning artifact	UNP P39779
D	-9	MET	-	initiating methionine	UNP P39779
D	-8	GLY	-	cloning artifact	UNP P39779
D	-7	SER	-	cloning artifact	UNP P39779
D	-6	SER	-	cloning artifact	UNP P39779
D	-5	HIS	-	expression tag	UNP P39779
D	-4	HIS	-	expression tag	UNP P39779
D	-3	HIS	-	expression tag	UNP P39779
D	-2	HIS	-	expression tag	UNP P39779
D	-1	HIS	-	expression tag	UNP P39779
D	0	HIS	-	expression tag	UNP P39779
D	1	MET	-	cloning artifact	UNP P39779

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



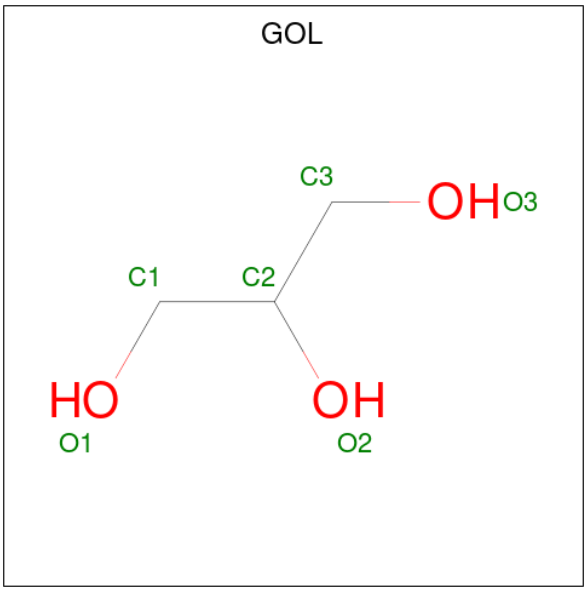
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	D	1	5	4	1	0	0

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



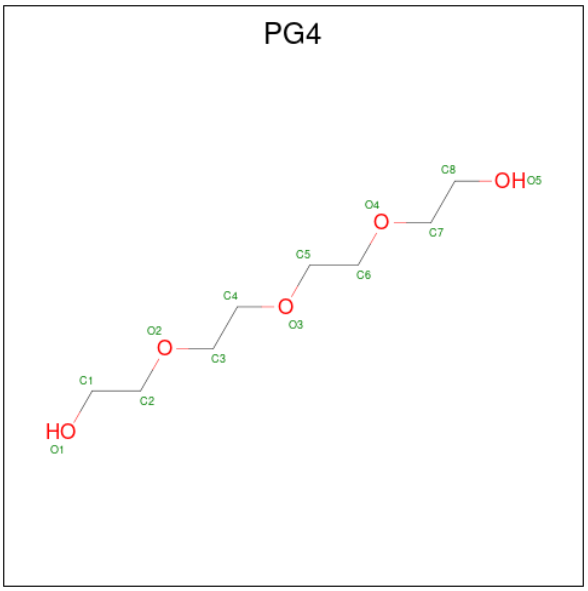
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	10	6	4	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



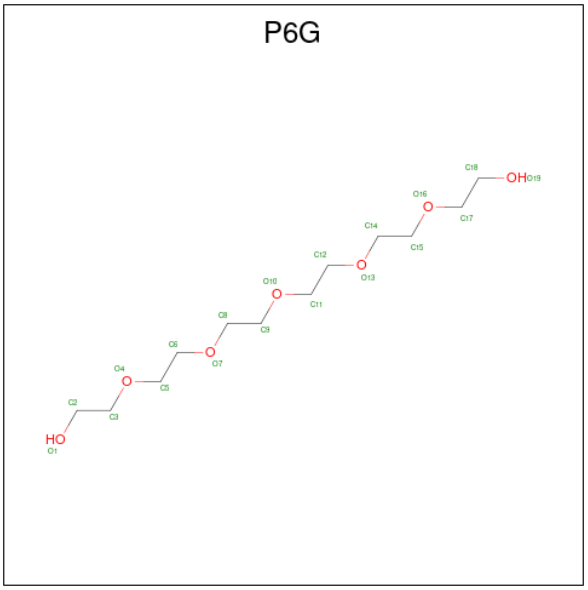
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	6	3	3	0	0
			Total	C	O		
4	A	1	6	3	3	0	0
			Total	C	O		
4	B	1	6	3	3	0	0
			Total	C	O		
4	C	1	6	3	3	0	0
			Total	C	O		
4	C	1	6	3	3	0	0
			Total	C	O		
4	C	1	6	3	3	0	0
			Total	C	O		
4	D	1	6	3	3	0	0
			Total	C	O		
4	D	1	6	3	3	0	0
			Total	C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

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



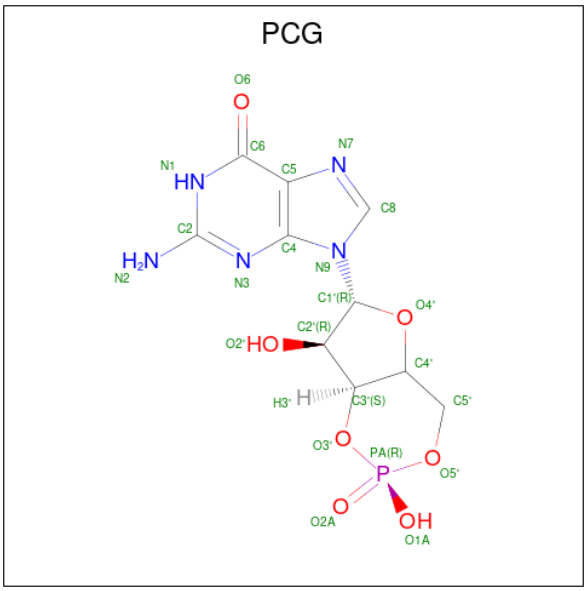
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	C	1	13	8	5	0	0

- Molecule 6 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	C	1	19	12	7	0	0

- Molecule 7 is CYCLIC GUANOSINE MONOPHOSPHATE (three-letter code: PCG) (formula: $C_{10}H_{12}N_5O_7P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
7	C	1	23	10	5	7	1	12	0

- Molecule 8 is water.

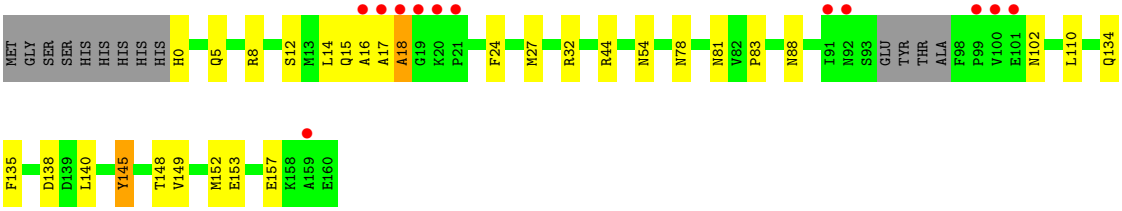
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	98	98	98	0	0
			Total	O		
8	B	80	80	80	0	0
			Total	O		
8	C	101	101	101	0	0
			Total	O		
8	D	109	109	109	0	0

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 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.

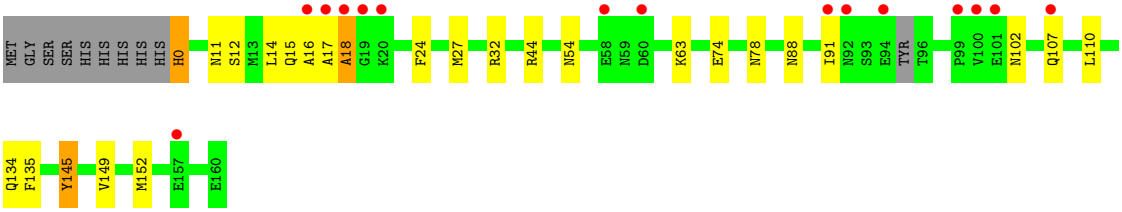
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor codY

Chain A:



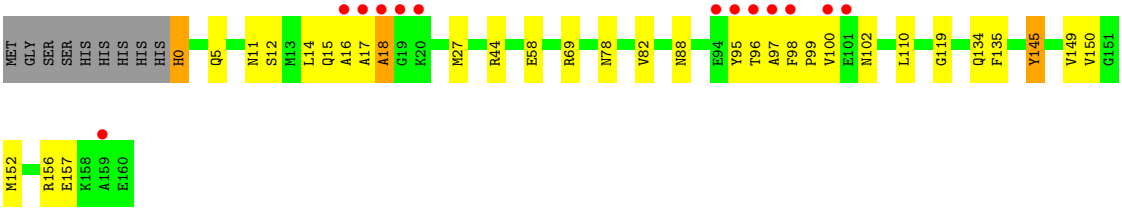
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor codY

Chain B:



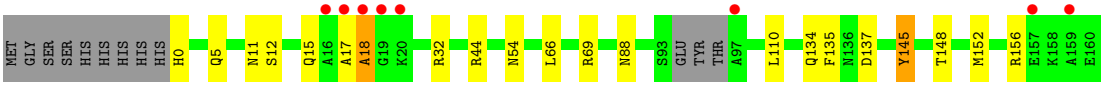
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor codY

Chain C:



• Molecule 1: GTP-sensing transcriptional pleiotropic repressor codY

Chain D:


MET
GLY
SER
SER
HIS
HIS
HIS
HIS
HIS
HO
Q5
N11
S12
Q15
A16
A17
A18
G19
K20
R32
R44
N54
L66
R69
N88
S93
GLU
TYR
THR
A97
L110
Q134
F135
N136
D137
Y145
T148
M152
R156
E157
K158
A159
E160

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	90.21Å 90.21Å 205.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.89 – 1.74 44.66 – 1.58	Depositor EDS
% Data completeness (in resolution range)	97.5 (31.89-1.74) 80.4 (44.66-1.58)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 1.58Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.187 , 0.216 0.194 , 0.194	Depositor DCC
R_{free} test set	4643 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	23.7	Xtriage
Anisotropy	0.200	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 57.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5983	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 51.36 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.6694e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: PCG, GOL, PG4, P6G, PGE, SO4

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	0.37	0/1379	0.54	0/1856
1	B	0.37	0/1400	0.55	0/1885
1	C	0.40	0/1398	0.56	0/1885
1	D	0.39	0/1368	0.56	0/1840
All	All	0.38	0/5545	0.55	0/7466

There are no bond length outliers.

There are no bond angle outliers.

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	1347	0	1367	21	0
1	B	1368	0	1385	24	0
1	C	1370	0	1375	27	0
1	D	1339	0	1363	16	0
2	A	5	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	0	0
2	D	15	0	0	0	0
3	A	10	0	14	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	12	0	16	1	0
4	B	6	0	8	0	0
4	C	18	0	24	1	0
4	D	30	0	40	0	0
5	C	13	0	18	0	0
6	C	19	0	26	0	0
7	C	23	0	11	0	0
8	A	98	0	0	0	0
8	B	80	0	0	0	0
8	C	101	0	0	5	0
8	D	109	0	0	1	0
All	All	5983	0	5647	81	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 7.

All (81) 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:D:5[B]:GLN:NE2	8:D:615:HOH:O	1.67	1.12
1:B:88:ASN:HD21	1:B:135:PHE:H	1.11	0.96
1:A:88:ASN:HD21	1:A:135:PHE:H	1.12	0.95
1:C:88:ASN:HD21	1:C:135:PHE:H	1.13	0.95
1:B:78[B]:ASN:ND2	1:B:102:ASN:HD22	1.66	0.94
1:C:5[B]:GLN:NE2	8:C:568:HOH:O	1.74	0.94
1:A:78[B]:ASN:HD21	1:A:102:ASN:HD22	0.98	0.94
1:D:88:ASN:HD21	1:D:135:PHE:H	1.13	0.91
1:B:78[B]:ASN:HD21	1:B:102:ASN:HD22	0.92	0.87
1:C:78[B]:ASN:HD22	1:C:102:ASN:ND2	1.80	0.80
1:C:97:ALA:HB2	8:C:626:HOH:O	1.79	0.80
1:C:95:TYR:OH	1:C:98:PHE:HB2	1.83	0.79
1:A:78[B]:ASN:ND2	1:A:102:ASN:HD22	1.79	0.78
1:A:12[A]:SER:O	1:A:16[A]:ALA:HB2	1.86	0.75
1:C:100:VAL:HG23	8:C:621:HOH:O	1.86	0.75
1:A:78[B]:ASN:HD21	1:A:102:ASN:ND2	1.81	0.72
1:C:11[A]:ASN:ND2	1:C:145[A]:TYR:OH	2.23	0.72
1:B:78[B]:ASN:HD21	1:B:102:ASN:ND2	1.78	0.72
1:C:82:VAL:HG22	1:C:97:ALA:HB1	1.71	0.71
1:A:17[A]:ALA:O	1:A:18[A]:ALA:HB3	1.90	0.71
1:C:12[A]:SER:HA	1:C:15[A]:GLN:HG2	1.72	0.71
1:B:11[A]:ASN:ND2	1:B:145[A]:TYR:OH	2.25	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:12[A]:SER:HA	1:B:15[A]:GLN:HE21	1.59	0.68
1:B:17[A]:ALA:O	1:B:18[A]:ALA:HB3	1.95	0.66
1:A:24:PHE:CD1	1:A:27[A]:MET:HE3	2.31	0.66
1:A:12[A]:SER:HA	1:A:15[A]:GLN:HG2	1.80	0.63
1:D:88:ASN:HD21	1:D:135:PHE:N	1.92	0.62
1:C:78[B]:ASN:HD22	1:C:102:ASN:HD22	1.46	0.61
1:B:12[A]:SER:HA	1:B:15[A]:GLN:HG2	1.84	0.60
1:D:66:LEU:O	1:D:69:ARG:HD3	2.01	0.60
1:C:97:ALA:HB3	8:C:620:HOH:O	2.01	0.60
1:A:17[A]:ALA:O	1:A:18[A]:ALA:CB	2.53	0.57
1:C:27:MET:HE3	1:C:150:VAL:HG22	1.87	0.57
1:C:17[A]:ALA:O	1:C:18[A]:ALA:HB3	2.05	0.56
1:D:17[A]:ALA:O	1:D:18[A]:ALA:HB3	2.06	0.56
1:C:78[B]:ASN:ND2	1:C:102:ASN:ND2	2.52	0.55
1:A:17[B]:ALA:O	1:A:18[B]:ALA:C	2.45	0.54
1:A:81:ASN:O	1:A:83:PRO:HD3	2.07	0.54
1:B:17[B]:ALA:O	1:B:18[B]:ALA:C	2.45	0.54
1:C:12[A]:SER:HA	1:C:15[A]:GLN:HE21	1.71	0.54
1:C:88:ASN:HD21	1:C:135:PHE:N	1.94	0.54
1:A:14[A]:LEU:HG	1:A:149:VAL:HG22	1.90	0.54
1:B:74:GLU:HG2	1:B:102:ASN:HD21	1.73	0.53
1:D:11[A]:ASN:ND2	1:D:145[A]:TYR:OH	2.42	0.53
1:C:14[A]:LEU:HG	1:C:149:VAL:HG22	1.93	0.51
1:A:145[A]:TYR:CE2	1:B:145[A]:TYR:CE2	2.99	0.51
1:B:12[A]:SER:CA	1:B:15[A]:GLN:HE21	2.22	0.51
1:D:12[A]:SER:HA	1:D:15[A]:GLN:HG2	1.92	0.51
4:C:518:GOL:H31	1:D:15[A]:GLN:NE2	2.27	0.50
1:C:11[A]:ASN:OD1	1:D:148:THR:OG1	2.27	0.49
1:B:24:PHE:HA	1:B:27[A]:MET:HE3	1.94	0.48
1:B:32:ARG:HD2	1:B:54:ASN:HB3	1.95	0.48
1:B:78[B]:ASN:ND2	1:B:102:ASN:ND2	2.49	0.48
1:D:88:ASN:ND2	1:D:135:PHE:H	1.95	0.48
1:A:110[B]:LEU:HD13	1:A:134:GLN:HA	1.95	0.48
1:A:153:GLU:OE1	4:A:520:GOL:H31	2.15	0.47
1:B:88:ASN:ND2	1:B:135:PHE:H	1.95	0.47
1:B:14[A]:LEU:CD2	1:B:149:VAL:HG22	2.46	0.46
1:C:16[A]:ALA:O	1:C:18[A]:ALA:N	2.48	0.46
1:D:17[B]:ALA:O	1:D:18[B]:ALA:C	2.53	0.46
1:B:17[A]:ALA:O	1:B:18[A]:ALA:CB	2.60	0.46
1:D:32:ARG:HD2	1:D:54:ASN:HB3	1.96	0.46
1:A:148:THR:HG1	1:B:11[B]:ASN:ND2	2.14	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15[B]:GLN:C	1:A:17[B]:ALA:N	2.70	0.45
1:A:138:ASP:OD1	1:B:0[A]:HIS:ND1	2.50	0.45
1:A:32:ARG:HD2	1:A:54:ASN:HB3	2.00	0.44
1:A:5[B]:GLN:HE22	1:A:8:ARG:HH11	1.66	0.44
1:B:91:ILE:HG23	1:B:107:GLN:O	2.18	0.43
1:C:12[A]:SER:HA	1:C:15[A]:GLN:NE2	2.33	0.43
1:A:16[A]:ALA:O	1:A:17[A]:ALA:C	2.57	0.43
1:C:0[B]:HIS:ND1	1:C:0[B]:HIS:N	2.61	0.43
1:C:110:LEU:HD13	1:C:134:GLN:HA	2.01	0.42
1:C:99:PRO:HB2	8:C:540:HOH:O	2.18	0.42
1:B:110[A]:LEU:HD23	1:B:134:GLN:HA	2.01	0.42
1:C:11[B]:ASN:HD21	1:D:148:THR:HG1	1.68	0.42
1:B:74:GLU:CG	1:B:102:ASN:HD21	2.32	0.42
1:C:119:GLY:C	1:D:17[A]:ALA:HA	2.39	0.42
1:B:12[A]:SER:O	1:B:16[A]:ALA:HB3	2.20	0.41
1:C:27:MET:CE	1:C:150:VAL:HG22	2.50	0.41
1:C:0[B]:HIS:HB2	1:D:137:ASP:HB2	2.02	0.41
1:D:110:LEU:HD13	1:D:134:GLN:HA	2.04	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	167/170 (98%)	160 (96%)	5 (3%)	2 (1%)	13 3
1	B	170/170 (100%)	163 (96%)	5 (3%)	2 (1%)	13 3
1	C	171/170 (101%)	164 (96%)	5 (3%)	2 (1%)	13 3
1	D	166/170 (98%)	160 (96%)	4 (2%)	2 (1%)	13 3
All	All	674/680 (99%)	647 (96%)	19 (3%)	8 (1%)	25 3

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	18[A]	ALA
1	D	18[B]	ALA
1	A	18[A]	ALA
1	A	18[B]	ALA
1	B	18[A]	ALA
1	B	18[B]	ALA
1	C	18[A]	ALA
1	C	18[B]	ALA

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	150/149 (101%)	142 (95%)	8 (5%)	22 5
1	B	152/149 (102%)	145 (95%)	7 (5%)	27 7
1	C	151/149 (101%)	140 (93%)	11 (7%)	14 2
1	D	148/149 (99%)	141 (95%)	7 (5%)	26 7
All	All	601/596 (101%)	568 (94%)	33 (6%)	27 4

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0[A]	HIS
1	A	0[B]	HIS
1	A	44	ARG
1	A	140	LEU
1	A	145[A]	TYR
1	A	145[B]	TYR
1	A	152	MET
1	A	157	GLU
1	B	0[A]	HIS
1	B	0[B]	HIS
1	B	44	ARG
1	B	63	LYS
1	B	145[A]	TYR
1	B	145[B]	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	152	MET
1	C	0[A]	HIS
1	C	0[B]	HIS
1	C	44	ARG
1	C	58	GLU
1	C	69	ARG
1	C	96	THR
1	C	145[A]	TYR
1	C	145[B]	TYR
1	C	152	MET
1	C	156	ARG
1	C	157	GLU
1	D	0[A]	HIS
1	D	0[B]	HIS
1	D	44	ARG
1	D	145[A]	TYR
1	D	145[B]	TYR
1	D	152	MET
1	D	156	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	ASN
1	A	102	ASN
1	B	88	ASN
1	B	102	ASN
1	C	88	ASN
1	C	102	ASN
1	D	88	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)

23 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	$\#\|Z\| > 2$
5	PG4	C	512	-	12,12,12	0.46	0	11,11,11	0.19	0
2	SO4	C	501	-	4,4,4	0.12	0	6,6,6	0.14	0
2	SO4	B	504	-	4,4,4	0.17	0	6,6,6	0.19	0
4	GOL	A	525	-	5,5,5	0.40	0	5,5,5	0.31	0
2	SO4	A	509	-	4,4,4	0.18	0	6,6,6	0.19	0
6	P6G	C	513	-	18,18,18	0.51	0	17,17,17	0.20	0
4	GOL	D	516	-	5,5,5	0.37	0	5,5,5	0.24	0
4	GOL	D	521	-	5,5,5	0.43	0	5,5,5	0.49	0
7	PCG	C	514	-	22,26,26	1.48	4 (18%)	25,41,41	1.32	4 (16%)
2	SO4	B	510	-	4,4,4	0.13	0	6,6,6	0.11	0
4	GOL	C	518	-	5,5,5	0.44	0	5,5,5	0.29	0
4	GOL	C	517	-	5,5,5	0.40	0	5,5,5	0.50	0
4	GOL	D	524	-	5,5,5	0.37	0	5,5,5	0.20	0
2	SO4	D	505	-	4,4,4	0.15	0	6,6,6	0.06	0
4	GOL	D	519	-	5,5,5	0.47	0	5,5,5	0.37	0
2	SO4	D	502	-	4,4,4	0.15	0	6,6,6	0.09	0
3	PGE	A	515	-	9,9,9	0.49	0	8,8,8	0.22	0
2	SO4	C	511	-	4,4,4	0.14	0	6,6,6	0.04	0
4	GOL	D	522	-	5,5,5	0.43	0	5,5,5	0.18	0
2	SO4	D	506	-	4,4,4	0.16	0	6,6,6	0.17	0
4	GOL	A	520	-	5,5,5	0.33	0	5,5,5	0.38	0
4	GOL	C	526	-	5,5,5	0.27	0	5,5,5	0.51	0
4	GOL	B	523	-	5,5,5	0.41	0	5,5,5	0.23	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
4	GOL	C	517	-	-	2/4/4/4	-
5	PG4	C	512	-	-	5/10/10/10	-
4	GOL	D	524	-	-	2/4/4/4	-
4	GOL	D	519	-	-	4/4/4/4	-
4	GOL	D	522	-	-	2/4/4/4	-
3	PGE	A	515	-	-	4/7/7/7	-
4	GOL	D	516	-	-	2/4/4/4	-
4	GOL	D	521	-	-	2/4/4/4	-
6	P6G	C	513	-	-	8/16/16/16	-
4	GOL	A	520	-	-	0/4/4/4	-
7	PCG	C	514	-	3/3/5/5	0/0/31/31	0/4/4/4
4	GOL	A	525	-	-	2/4/4/4	-
4	GOL	C	518	-	-	0/4/4/4	-
4	GOL	C	526	-	-	2/4/4/4	-
4	GOL	B	523	-	-	0/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	C	514	PCG	PA-O5'	3.83	1.62	1.57
7	C	514	PCG	C6-N1	-2.34	1.34	1.37
7	C	514	PCG	PA-O3'	2.34	1.61	1.57
7	C	514	PCG	O4'-C1'	2.03	1.43	1.41

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	C	514	PCG	O1A-PA-O2A	3.62	120.06	108.73
7	C	514	PCG	C8-N7-C5	2.58	107.91	102.99
7	C	514	PCG	O3'-C3'-C4'	-2.36	108.93	110.71
7	C	514	PCG	C5-C6-N1	2.07	117.61	113.95

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	C	514	PCG	C1'
7	C	514	PCG	C4'
7	C	514	PCG	C3'

All (35) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	516	GOL	O1-C1-C2-C3
4	D	519	GOL	C1-C2-C3-O3
4	D	522	GOL	C1-C2-C3-O3
6	C	513	P6G	O7-C8-C9-O10
6	C	513	P6G	O4-C5-C6-O7
6	C	513	P6G	O13-C14-C15-O16
5	C	512	PG4	O3-C5-C6-O4
4	A	525	GOL	C1-C2-C3-O3
4	C	517	GOL	C1-C2-C3-O3
4	D	519	GOL	O1-C1-C2-C3
4	D	521	GOL	O1-C1-C2-C3
4	D	524	GOL	O1-C1-C2-C3
5	C	512	PG4	O1-C1-C2-O2
4	D	516	GOL	O1-C1-C2-O2
4	D	519	GOL	O1-C1-C2-O2
4	D	522	GOL	O2-C2-C3-O3
3	A	515	PGE	O2-C3-C4-O3
5	C	512	PG4	O4-C7-C8-O5
6	C	513	P6G	O1-C2-C3-O4
6	C	513	P6G	O16-C17-C18-O19
4	D	519	GOL	O2-C2-C3-O3
4	C	517	GOL	O2-C2-C3-O3
4	D	521	GOL	O1-C1-C2-O2
3	A	515	PGE	C4-C3-O2-C2
6	C	513	P6G	C5-C6-O7-C8
5	C	512	PG4	C6-C5-O3-C4
6	C	513	P6G	O10-C11-C12-O13
3	A	515	PGE	C6-C5-O3-C4
5	C	512	PG4	C4-C3-O2-C2
6	C	513	P6G	C15-C14-O13-C12
4	A	525	GOL	O2-C2-C3-O3
4	C	526	GOL	O1-C1-C2-C3
4	C	526	GOL	O1-C1-C2-O2
4	D	524	GOL	O1-C1-C2-O2
3	A	515	PGE	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	518	GOL	1	0

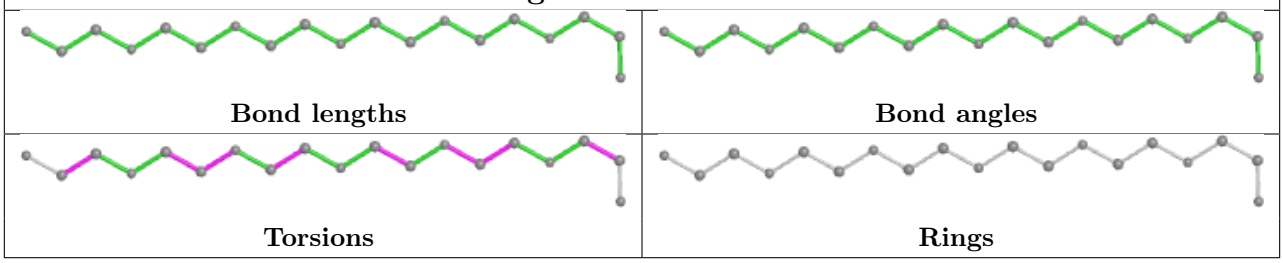
Continued on next page...

Continued from previous page...

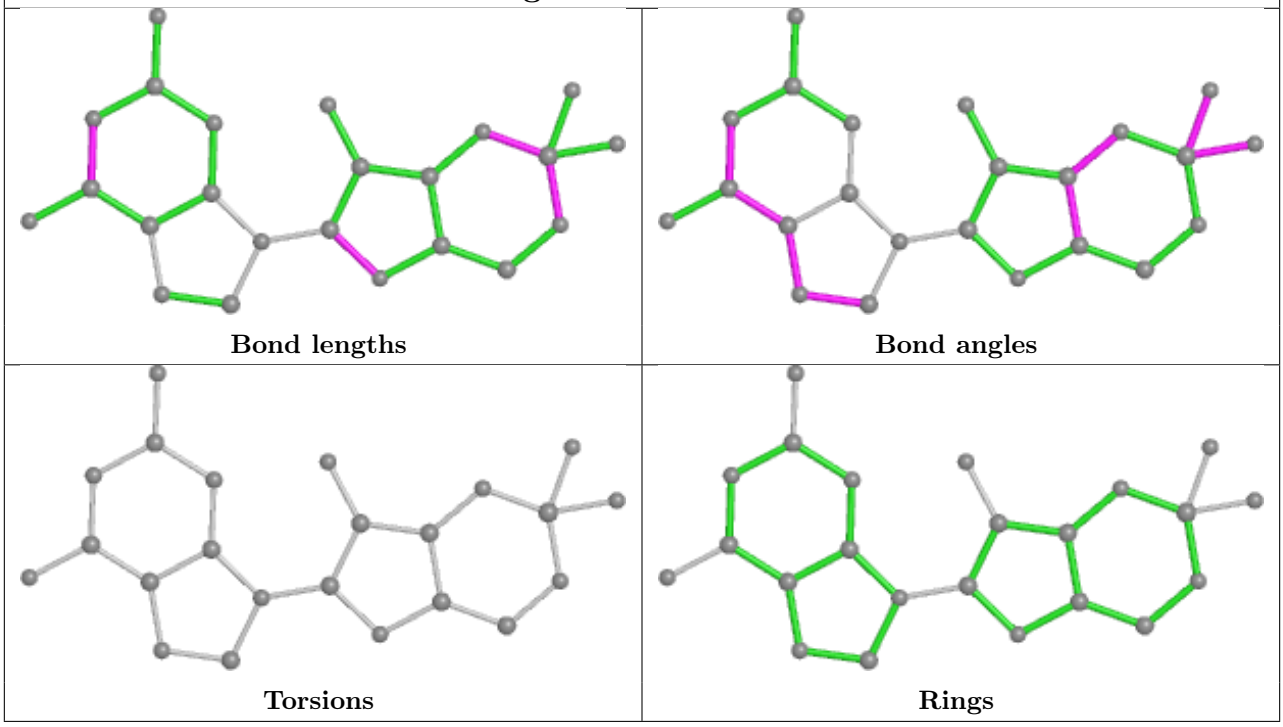
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	520	GOL	1	0

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 P6G C 513



Ligand PCG C 514



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	157/170 (92%)	0.33	12 (7%)	13	18	17, 26, 55, 67	0
1	B	160/170 (94%)	0.53	15 (9%)	8	10	17, 28, 64, 71	0
1	C	161/170 (94%)	0.42	13 (8%)	12	15	14, 22, 56, 66	0
1	D	158/170 (92%)	0.15	8 (5%)	28	33	15, 21, 49, 68	0
All	All	636/680 (93%)	0.36	48 (7%)	14	18	14, 24, 59, 71	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	96	THR	12.3
1	C	97	ALA	9.7
1	B	100	VAL	8.4
1	B	91	ILE	8.3
1	C	19	GLY	8.2
1	B	19	GLY	8.1
1	A	100	VAL	6.6
1	B	16[A]	ALA	6.5
1	D	159	ALA	6.3
1	C	95	TYR	6.3
1	B	17[A]	ALA	5.9
1	C	17[A]	ALA	5.6
1	A	17[A]	ALA	5.3
1	D	97	ALA	4.2
1	D	18[A]	ALA	4.2
1	A	19	GLY	4.1
1	A	91	ILE	4.0
1	D	19	GLY	3.9
1	B	92	ASN	3.8
1	D	17[A]	ALA	3.8
1	B	101[A]	GLU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	100	VAL	3.6
1	C	101[A]	GLU	3.5
1	C	16[A]	ALA	3.5
1	A	16[A]	ALA	3.4
1	A	20	LYS	3.4
1	C	159	ALA	3.2
1	C	18[A]	ALA	3.1
1	C	94	GLU	3.0
1	B	18[A]	ALA	3.0
1	A	92	ASN	2.8
1	D	20	LYS	2.8
1	D	16[A]	ALA	2.8
1	A	159	ALA	2.8
1	C	98	PHE	2.7
1	B	107	GLN	2.7
1	A	99	PRO	2.7
1	B	99	PRO	2.5
1	A	18[A]	ALA	2.5
1	B	58	GLU	2.4
1	B	60	ASP	2.2
1	C	20	LYS	2.2
1	A	101[A]	GLU	2.1
1	B	20	LYS	2.1
1	B	94	GLU	2.1
1	D	157	GLU	2.1
1	A	21	PRO	2.0
1	B	157	GLU	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,

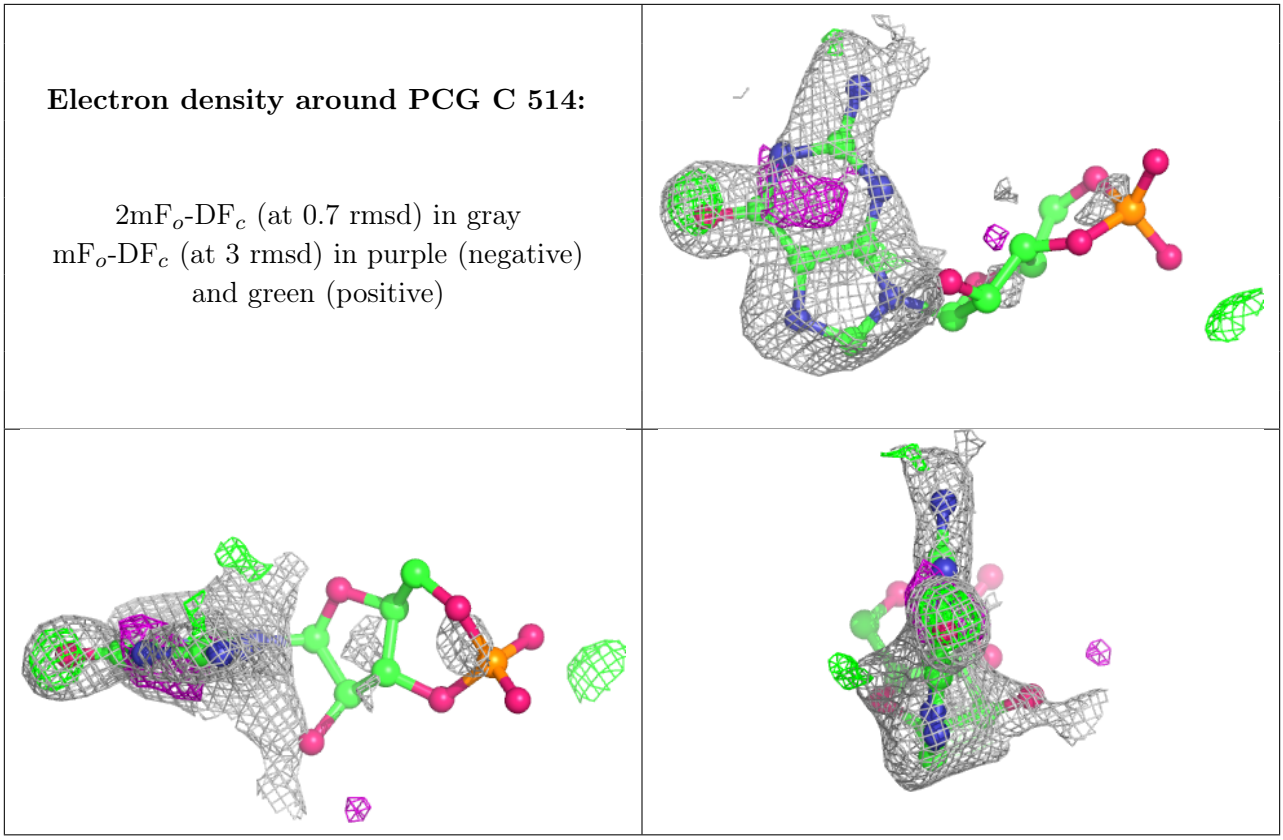
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	PCG	C	514	23/23	0.47	0.24	54,68,69,76	12
2	SO4	D	502	5/5	0.49	0.38	73,75,76,76	0
3	PGE	A	515	10/10	0.52	0.41	65,67,69,69	0
4	GOL	C	526	6/6	0.57	0.15	42,46,48,48	0
6	P6G	C	513	19/19	0.60	0.20	53,58,62,63	0
2	SO4	D	505	5/5	0.74	0.19	46,48,48,48	5
2	SO4	C	511	5/5	0.76	0.14	110,110,110,110	0
4	GOL	C	518	6/6	0.78	0.18	34,41,44,46	0
4	GOL	D	524	6/6	0.79	0.15	43,51,53,55	0
4	GOL	A	520	6/6	0.80	0.14	51,56,58,60	0
5	PG4	C	512	13/13	0.80	0.12	51,53,58,60	0
4	GOL	A	525	6/6	0.81	0.18	42,46,47,48	0
4	GOL	B	523	6/6	0.83	0.15	46,48,49,51	0
4	GOL	D	519	6/6	0.84	0.20	47,52,53,55	0
2	SO4	B	510	5/5	0.85	0.17	46,46,47,47	5
4	GOL	D	516	6/6	0.86	0.13	54,54,55,55	0
2	SO4	A	509	5/5	0.87	0.15	25,29,31,32	5
4	GOL	C	517	6/6	0.87	0.18	40,43,44,47	0
2	SO4	B	504	5/5	0.88	0.20	30,31,33,35	5
4	GOL	D	522	6/6	0.88	0.28	45,45,47,48	0
4	GOL	D	521	6/6	0.90	0.18	34,39,42,43	0
2	SO4	D	506	5/5	0.92	0.15	22,26,29,30	5
2	SO4	C	501	5/5	0.92	0.19	24,29,30,31	5

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.

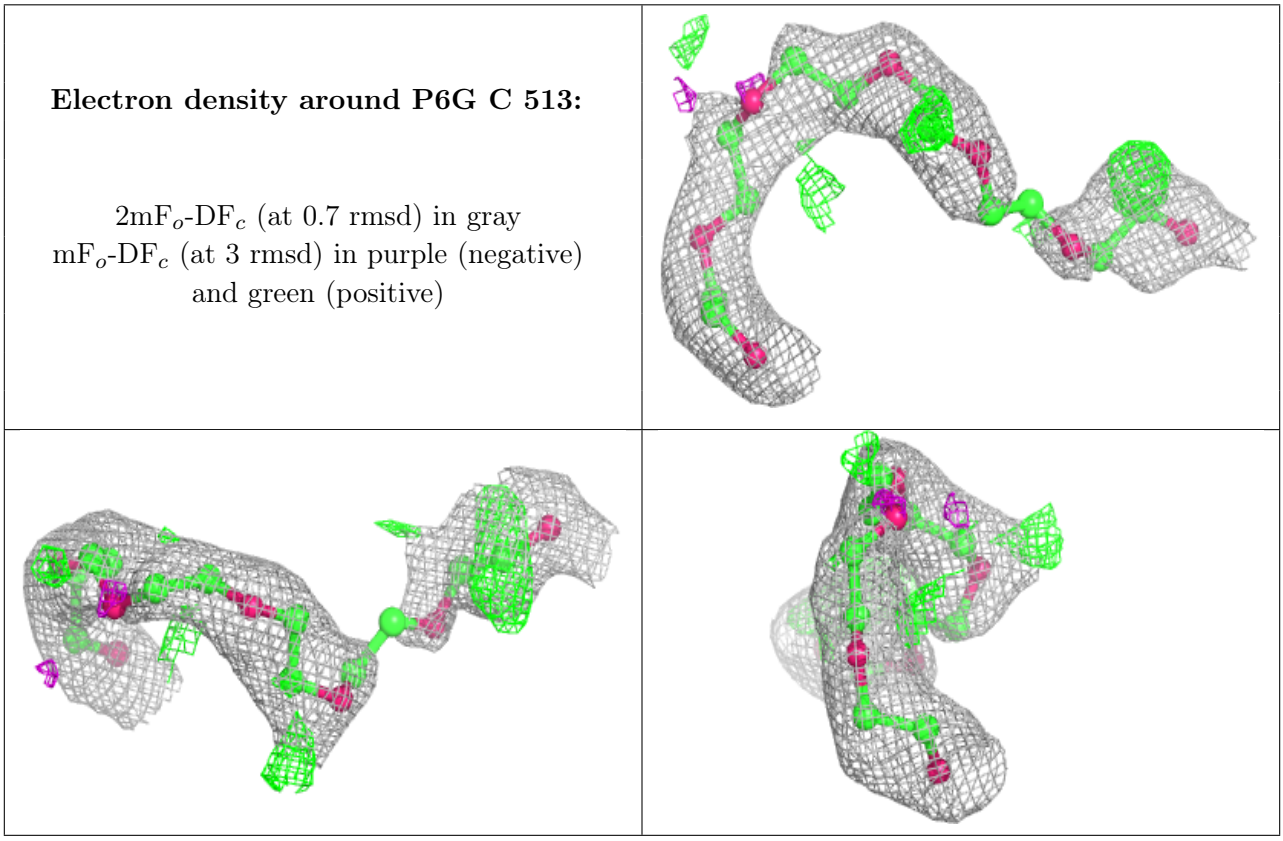
Electron density around PCG C 514:

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



Electron density around P6G C 513:

$2mF_o$-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.