Full wwPDB X-ray Structure Validation Report (i)

May 24, 2020 – 03:58 am BST

PDB ID	:	2J41
Title	:	Crystal structure of Staphylococcus aureus guanylate monophosphate kinase
Authors	:	El Omari, K.; Dhaliwal, B.; Lockyer, M.; Charles, I.; Hawkins, A.R.; Stammers, D.K.
Deposited on	:	2006-08-24
Resolution	:	1.90 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 following versions of software and data (see references (i)) were used in the production of this report:

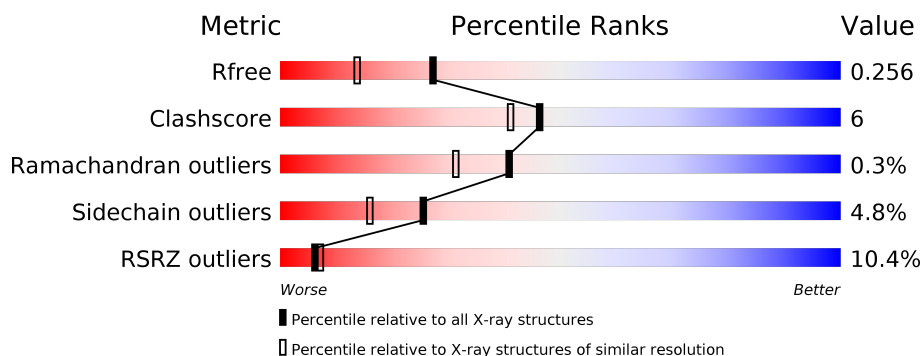|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.90 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	Percentile Ranks	Value
Rfree		0.256
Clashscore		6
Ramachandran outliers		0.3%
Sidechain outliers		4.8%
RSRZ outliers		10.4%

Worse　　　　　　　　　　　　　　　　　　　　　*Better*

▉ Percentile relative to all X-ray structures

▊ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=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	207	7% — 65% / 14% · 19%
1	B	207	6% — 74% / 8% · 16%
1	C	207	9% — 74% / 7% · 16%
1	D	207	13% — 77% / 7% · 14%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6201 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 GUANYLATE KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	167	Total	C	N	O	S	0	1	0
			1363	873	221	263	6			
1	B	173	Total	C	N	O	S	0	0	0
			1406	900	230	269	7			
1	C	174	Total	C	N	O	S	0	0	0
			1412	898	236	272	6			
1	D	178	Total	C	N	O	S	0	0	0
			1449	926	240	276	7			

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

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

Continued on next page...

Continued from previous page...

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

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total K 1 1		0	0
3	D	1	Total K 1 1		0	0
3	C	1	Total K 1 1		0	0

- Molecule 4 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula: $C_{10}H_{14}N_5O_8P$).
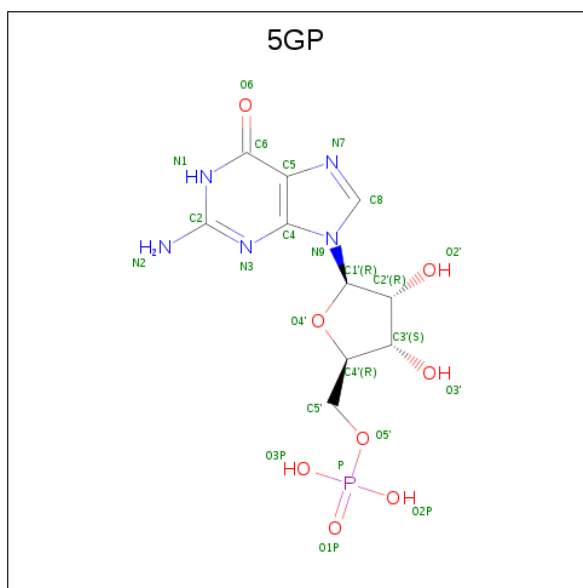
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total C N O P 24 10 5 8 1					0	0
4	C	1	Total C N O P 24 10 5 8 1					0	0
4	D	1	Total C N O P 24 10 5 8 1					0	0

• Molecule 5 is water.

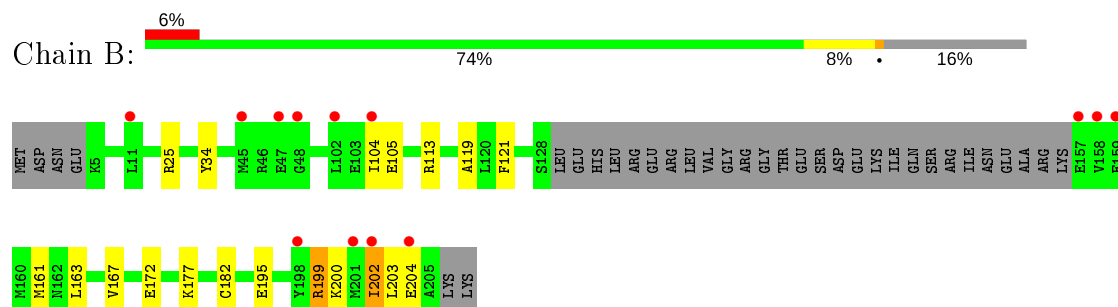| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	95	Total 95	O 95	0	0
5	B	111	Total 111	O 111	0	0
5	C	130	Total 130	O 130	0	0
5	D	135	Total 135	O 135	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: GUANYLATE KINASE
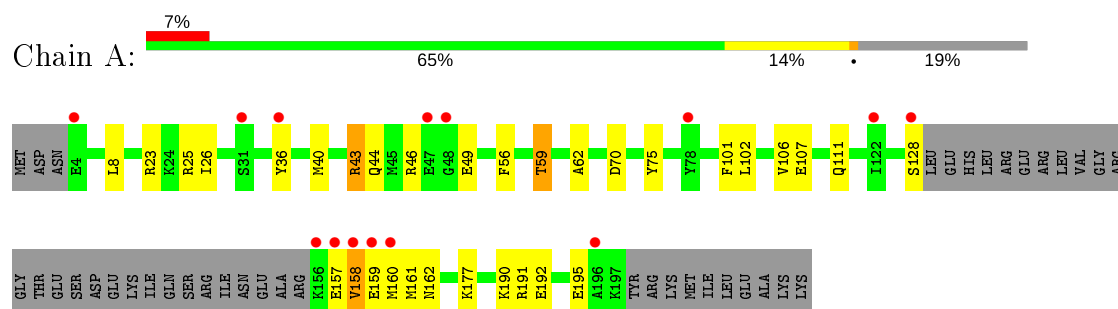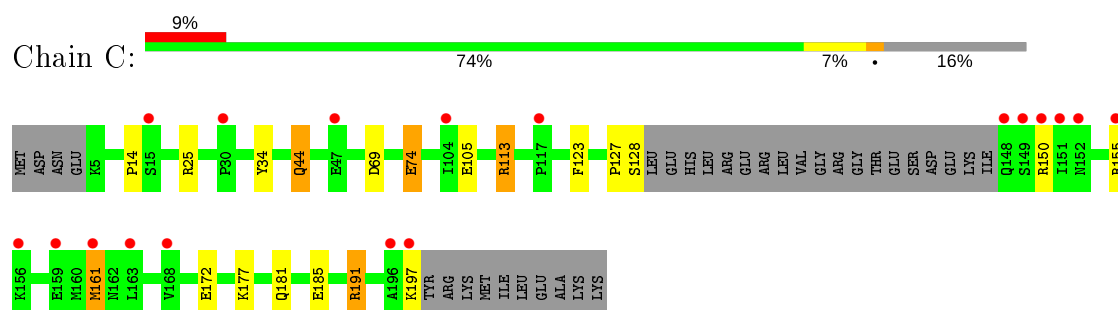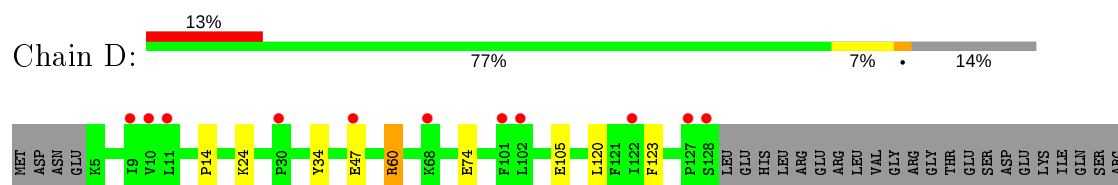

- Molecule 1: GUANYLATE KINASE

- Molecule 1: GUANYLATE KINASE

- Molecule 1: GUANYLATE KINASE

ILE ASN E153 A154 R155 K156 E157 V158 E159 M160 M161 N162 I183 E187 K190 R191 A196 K197 Y198 R199 K200 M201 I202 L203 E204 A205 K206 LYS

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.01Å 93.99Å 83.94Å 90.00° 110.59° 90.00°	Depositor
Resolution (Å)	29.40 – 1.90 29.44 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.1 (29.40-1.90) 99.1 (29.44-1.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.00 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.207 , 0.239 0.228 , 0.256	Depositor DCC
R_{free} test set	3991 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 54.7	EDS
L-test for twinning[2]	$<|L|> = 0.50$, $<L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6201	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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: K, 5GP, 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	0.52	0/1392	0.67	0/1873
1	B	0.61	1/1432 (0.1%)	0.63	0/1926
1	C	0.61	2/1437 (0.1%)	0.68	3/1932 (0.2%)
1	D	0.58	1/1475 (0.1%)	0.65	0/1981
All	All	0.58	4/5736 (0.1%)	0.66	3/7712 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	182	CYS	CB-SG	8.16	1.96	1.82
1	C	161	MET	SD-CE	7.30	2.18	1.77
1	D	156	LYS	CE-NZ	5.83	1.63	1.49
1	C	161	MET	CG-SD	5.79	1.96	1.81

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	161	MET	CG-SD-CE	-8.02	87.36	100.20
1	C	74	GLU	N-CA-CB	-5.26	101.12	110.60
1	C	74	GLU	CA-CB-CG	5.07	124.56	113.40

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	1363	0	1341	26	0
1	B	1406	0	1388	19	0
1	C	1412	0	1391	17	0
1	D	1449	0	1438	12	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	10	0	0	1	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	B	24	0	12	2	0
4	C	24	0	12	0	0
4	D	24	0	12	0	0
5	A	95	0	0	4	0
5	B	111	0	0	1	0
5	C	130	0	0	5	0
5	D	135	0	0	4	0
All	All	6201	0	5594	71	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 6.

All (71) 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:B:161:MET:SD	1:B:161:MET:CE	2.04	1.44
1:C:161:MET:SD	1:C:161:MET:CE	2.18	1.32
1:B:199:ARG:HH21	1:B:199:ARG:HG3	1.15	1.10
1:B:199:ARG:HH21	1:B:199:ARG:CG	1.89	0.84
1:A:59:THR:HG22	1:A:62:ALA:H	1.43	0.82
1:B:199:ARG:NH2	1:B:199:ARG:HG3	1.89	0.81
1:A:23:ARG:HD2	1:A:36[A]:TYR:CE1	2.23	0.74
1:C:181:GLN:HG3	5:C:2114:HOH:O	1.89	0.73
1:A:106:VAL:HG12	1:A:160:MET:HG2	1.72	0.70
1:C:161:MET:HB3	1:C:161:MET:CE	2.22	0.69
2:D:1207:SO4:O2	5:D:2127:HOH:O	2.10	0.69
4:B:1208:5GP:O3'	5:B:2111:HOH:O	2.12	0.68
1:C:161:MET:HB3	1:C:161:MET:HE3	1.75	0.67
1:A:59:THR:CG2	1:A:62:ALA:H	2.07	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:191:ARG:HD3	1:D:161:MET:O	1.95	0.67
1:C:25:ARG:HG3	1:C:177:LYS:HD3	1.78	0.65
1:A:43:ARG:NH2	1:A:49:GLU:OE1	2.29	0.65
1:A:106:VAL:HG23	5:A:2062:HOH:O	1.97	0.63
1:D:60:ARG:HD2	5:D:2049:HOH:O	1.96	0.63
1:A:43:ARG:NH2	1:A:46:ARG:HG3	2.14	0.62
1:A:191:ARG:O	1:A:195:GLU:HB2	1.99	0.62
1:C:113:ARG:HD2	5:C:2104:HOH:O	2.03	0.58
1:B:113:ARG:NH1	1:B:163:LEU:O	2.38	0.56
1:A:40:MET:HG2	1:A:56:PHE:HB2	1.88	0.56
1:C:181:GLN:CG	5:C:2114:HOH:O	2.47	0.56
1:B:25:ARG:HG3	1:B:177:LYS:HD3	1.88	0.56
1:B:161:MET:HE3	1:B:167:VAL:HG21	1.87	0.56
1:D:14:PRO:HD3	1:D:123:PHE:CZ	2.42	0.55
1:C:161:MET:HE3	1:D:191:ARG:HG2	1.89	0.55
1:C:161:MET:CB	1:C:161:MET:CE	2.85	0.54
1:D:153:GLU:HB3	1:D:155:ARG:HH21	1.72	0.54
1:A:43:ARG:HH22	1:A:49:GLU:CD	2.11	0.54
1:B:172:GLU:OE1	1:C:25:ARG:NH1	2.39	0.54
1:B:199:ARG:HD3	1:B:203:LEU:HD22	1.89	0.53
1:D:187:GLU:OE2	1:D:190:LYS:HE3	2.09	0.53
1:A:23:ARG:HD2	1:A:36[A]:TYR:HE1	1.74	0.53
1:B:161:MET:CE	1:B:167:VAL:HG21	2.39	0.53
1:C:161:MET:CG	1:C:161:MET:CE	2.86	0.53
1:C:44:GLN:H	1:C:44:GLN:CD	2.11	0.52
1:A:128:SER:HA	5:A:2073:HOH:O	2.10	0.52
1:A:106:VAL:CG1	1:A:160:MET:HG2	2.40	0.52
1:A:162:ASN:HB3	1:B:195:GLU:OE2	2.10	0.51
1:B:161:MET:HB3	1:B:161:MET:CE	2.40	0.51
1:D:47:GLU:HG3	5:D:2039:HOH:O	2.11	0.50
1:C:128:SER:HB3	5:C:2101:HOH:O	2.12	0.50
1:B:161:MET:HB3	1:B:161:MET:HE3	1.93	0.50
1:B:161:MET:CB	1:B:161:MET:CE	2.90	0.49
1:D:199:ARG:O	1:D:202:ILE:HD13	2.13	0.49
1:A:25:ARG:HG3	1:A:177:LYS:HD3	1.93	0.48
1:A:75:TYR:H	1:A:111:GLN:HE22	1.61	0.48
1:A:36[A]:TYR:CD1	5:A:2058:HOH:O	2.56	0.48
1:A:157:GLU:C	1:A:159:GLU:H	2.16	0.47
1:D:198:TYR:O	1:D:202:ILE:HG23	2.17	0.44
1:C:14:PRO:HD3	1:C:123:PHE:CZ	2.52	0.44
1:C:150:ARG:HA	5:C:2102:HOH:O	2.16	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:113:ARG:HD2	1:B:163:LEU:HD13	1.99	0.44
1:A:158:VAL:HG12	1:A:161:MET:HE3	2.00	0.43
1:A:190:LYS:HG3	1:A:192:GLU:OE1	2.18	0.43
1:A:158:VAL:HG12	1:A:158:VAL:O	2.19	0.42
1:D:120:LEU:HD22	1:D:187:GLU:HG3	2.00	0.42
1:B:25:ARG:NH2	1:C:172:GLU:OE2	2.50	0.42
1:A:111:GLN:HG3	5:A:2063:HOH:O	2.19	0.42
1:B:119:ALA:HB3	1:B:121:PHE:CE2	2.53	0.42
1:D:155:ARG:H	1:D:155:ARG:CZ	2.33	0.42
1:A:59:THR:HG22	1:A:62:ALA:CB	2.50	0.41
1:A:43:ARG:HH21	1:A:46:ARG:HG3	1.85	0.41
1:B:199:ARG:O	1:B:202:ILE:HG22	2.20	0.41
1:A:8:LEU:HD11	1:A:102:LEU:HG	2.02	0.41
1:D:187:GLU:HG2	5:D:2104:HOH:O	2.20	0.41
1:A:26:ILE:HD13	1:A:101:PHE:CZ	2.56	0.40
1:B:104:ILE:HG22	4:B:1208:5GP:C6	2.52	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	164/207 (79%)	161 (98%)	2 (1%)	1 (1%)	25	15
1	B	169/207 (82%)	167 (99%)	2 (1%)	0	100	100
1	C	170/207 (82%)	167 (98%)	2 (1%)	1 (1%)	25	15
1	D	174/207 (84%)	172 (99%)	2 (1%)	0	100	100
All	All	677/828 (82%)	667 (98%)	8 (1%)	2 (0%)	41	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	127	PRO
1	A	158	VAL

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	148/183 (81%)	143 (97%)	5 (3%)	37	28
1	B	152/183 (83%)	146 (96%)	6 (4%)	32	23
1	C	153/183 (84%)	143 (94%)	10 (6%)	17	8
1	D	156/183 (85%)	148 (95%)	8 (5%)	24	14
All	All	609/732 (83%)	580 (95%)	29 (5%)	25	16

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	43	ARG
1	A	44	GLN
1	A	59	THR
1	A	70	ASP
1	A	107	GLU
1	B	34	TYR
1	B	105	GLU
1	B	199	ARG
1	B	200	LYS
1	B	202	ILE
1	B	204	GLU
1	C	34	TYR
1	C	44	GLN
1	C	69	ASP
1	C	74	GLU
1	C	105	GLU
1	C	113	ARG
1	C	155	ARG
1	C	185	GLU
1	C	191	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	197	LYS
1	D	24	LYS
1	D	34	TYR
1	D	60	ARG
1	D	74	GLU
1	D	105	GLU
1	D	155	ARG
1	D	156	LYS
1	D	202	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	111	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 3 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	SO4	D	1208	-	4,4,4	0.16	0	6,6,6	0.42	0
4	5GP	D	1210	3	22,26,26	0.57	0	27,40,40	1.90	6 (22%)
2	SO4	D	1207	-	4,4,4	0.24	0	6,6,6	0.73	0
2	SO4	C	1198	-	4,4,4	0.25	0	6,6,6	0.49	0
2	SO4	A	1198	-	4,4,4	0.16	0	6,6,6	0.42	0
4	5GP	C	1200	3	22,26,26	0.79	0	27,40,40	2.03	7 (25%)
2	SO4	B	1206	-	4,4,4	0.23	0	6,6,6	0.39	0
4	5GP	B	1208	3	22,26,26	0.79	0	27,40,40	1.96	6 (22%)

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	5GP	D	1210	3	-	0/6/26/26	0/3/3/3
4	5GP	B	1208	3	-	0/6/26/26	0/3/3/3
4	5GP	C	1200	3	-	0/6/26/26	0/3/3/3

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	1200	5GP	N3-C2-N1	-5.69	119.63	127.22
4	B	1208	5GP	N3-C2-N1	-4.80	120.82	127.22
4	D	1210	5GP	N3-C2-N1	-4.67	120.99	127.22
4	B	1208	5GP	C5-C6-N1	-4.42	117.38	123.43
4	B	1208	5GP	C6-N1-C2	4.28	122.73	115.93
4	C	1200	5GP	C6-N1-C2	4.15	122.53	115.93
4	C	1200	5GP	C2-N3-C4	3.83	119.73	115.36
4	D	1210	5GP	C6-N1-C2	3.63	121.69	115.93
4	D	1210	5GP	C2-N3-C4	3.54	119.39	115.36
4	D	1210	5GP	C5-C6-N1	-3.44	118.72	123.43
4	C	1200	5GP	C5-C6-N1	-3.14	119.14	123.43
4	B	1208	5GP	C2-N3-C4	3.01	118.79	115.36
4	C	1200	5GP	N2-C2-N3	2.79	122.34	117.79
4	C	1200	5GP	C6-C5-C4	-2.68	118.24	120.80
4	D	1210	5GP	C4-C5-N7	-2.33	106.97	109.40
4	C	1200	5GP	O3'-C3'-C4'	-2.32	104.35	111.05
4	B	1208	5GP	C4-C5-N7	-2.26	107.04	109.40
4	D	1210	5GP	O3'-C3'-C4'	-2.26	104.52	111.05

Continued on next page...

Continued from previous page...

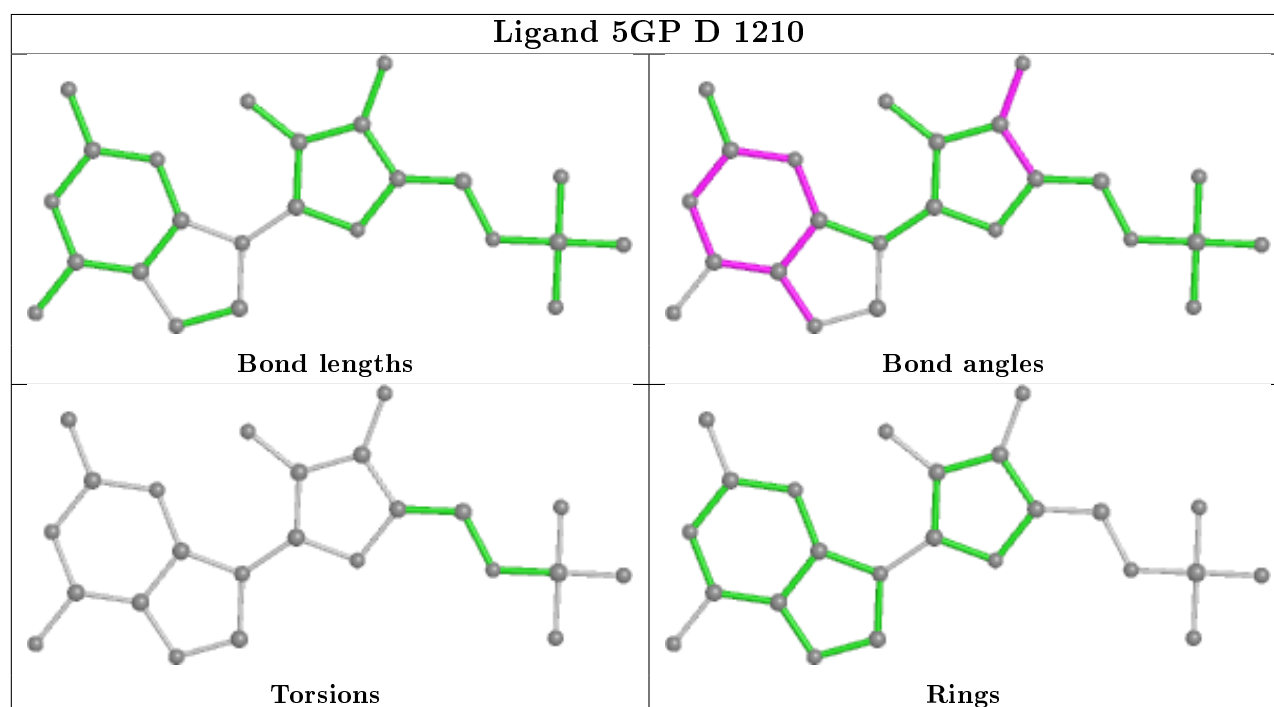| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	1208	5GP	O3'-C3'-C4'	-2.08	105.04	111.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1207	SO4	1	0
4	B	1208	5GP	2	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 5GP D 1210

Bond lengths

Bond angles

Torsions

Rings

Ligand 5GP C 1200

Bond lengths

Bond angles

Torsions

Rings

Ligand 5GP B 1208

Bond lengths

Bond angles

Torsions

Rings
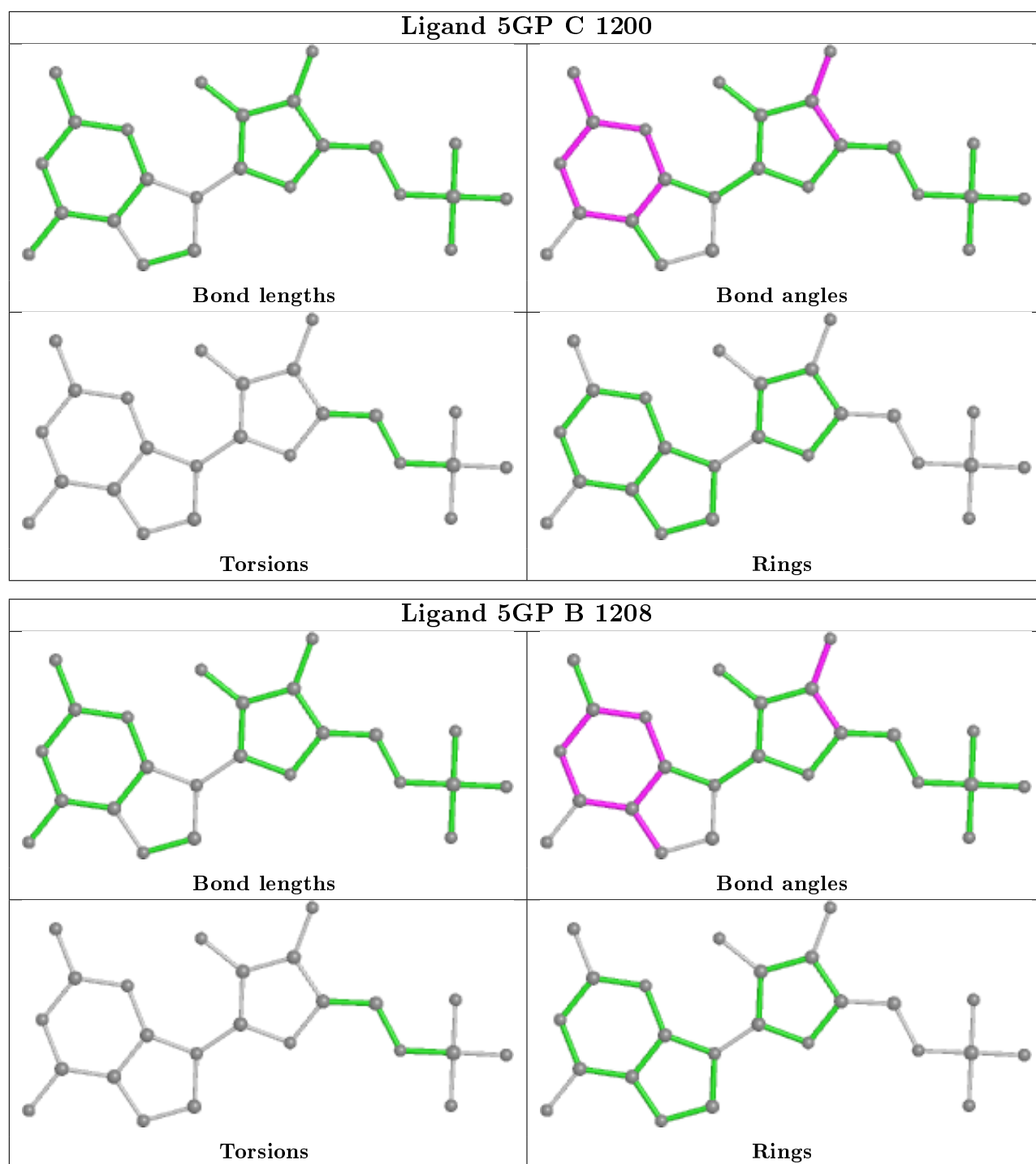
5.7　Other polymers (i)

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 (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	167/207 (80%)	0.50	14 (8%)	11	12	15, 24, 35, 41	0
1	B	173/207 (83%)	0.61	13 (7%)	14	15	14, 22, 44, 49	0
1	C	174/207 (84%)	0.72	18 (10%)	6	7	14, 21, 44, 56	0
1	D	178/207 (85%)	0.72	27 (15%)	2	2	14, 21, 45, 55	0
All	All	692/828 (83%)	0.64	72 (10%)	6	7	14, 22, 41, 56	0

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	148	GLN	7.4
1	D	156	LYS	7.2
1	C	151	ILE	6.5
1	C	149	SER	6.4
1	C	150	ARG	5.9
1	D	155	ARG	5.9
1	D	128	SER	5.8
1	B	201	MET	5.1
1	A	158	VAL	4.7
1	C	152	ASN	4.6
1	D	162	ASN	4.5
1	C	155	ARG	4.4
1	B	47	GLU	4.2
1	D	203	LEU	3.9
1	D	204	GLU	3.9
1	B	202	ILE	3.9
1	D	153	GLU	3.8
1	B	198	TYR	3.8
1	B	159	GLU	3.8
1	D	154	ALA	3.7
1	A	159	GLU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	205	ALA	3.7
1	D	157	GLU	3.6
1	A	78	TYR	3.5
1	C	197	LYS	3.4
1	C	163	LEU	3.3
1	D	206	LYS	3.3
1	C	196	ALA	3.2
1	A	157	GLU	3.0
1	D	102	LEU	2.8
1	D	183	ILE	2.8
1	D	9	ILE	2.8
1	C	30	PRO	2.8
1	D	11	LEU	2.8
1	A	47	GLU	2.8
1	A	160	MET	2.7
1	D	101	PHE	2.7
1	B	104	ILE	2.5
1	D	200	LYS	2.5
1	B	157	GLU	2.5
1	C	47	GLU	2.5
1	A	36[A]	TYR	2.5
1	A	156	LYS	2.4
1	D	201	MET	2.4
1	A	196	ALA	2.4
1	B	204	GLU	2.4
1	A	4	GLU	2.4
1	A	128	SER	2.4
1	A	122	ILE	2.4
1	B	158	VAL	2.4
1	D	10	VAL	2.3
1	B	102	LEU	2.3
1	D	199	ARG	2.3
1	C	159	GLU	2.3
1	B	11	LEU	2.3
1	C	15	SER	2.2
1	D	127	PRO	2.2
1	B	45	MET	2.2
1	C	168	VAL	2.2
1	C	104	ILE	2.1
1	D	47	GLU	2.1
1	C	161	MET	2.1
1	D	159	GLU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	156	LYS	2.1
1	A	48	GLY	2.1
1	D	68	LYS	2.1
1	D	122	ILE	2.1
1	B	48	GLY	2.1
1	C	117	PRO	2.0
1	D	30	PRO	2.0
1	A	31	SER	2.0
1	D	196	ALA	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 carbohydrates 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(Å²)	Q<0.9
3	K	D	1209	1/1	0.95	0.10	31,31,31,31	0
2	SO4	D	1207	5/5	0.95	0.14	31,32,38,39	0
4	5GP	C	1200	24/24	0.95	0.18	16,20,25,28	0
4	5GP	D	1210	24/24	0.96	0.15	18,21,28,29	0
4	5GP	B	1208	24/24	0.96	0.16	18,23,32,32	0
2	SO4	C	1198	5/5	0.97	0.14	30,32,37,37	0
2	SO4	D	1208	5/5	0.98	0.12	31,31,33,34	0
2	SO4	B	1206	5/5	0.98	0.12	32,34,38,39	0
2	SO4	A	1198	5/5	0.98	0.10	39,40,41,42	0
3	K	B	1207	1/1	0.99	0.12	31,31,31,31	0
3	K	C	1199	1/1	0.99	0.10	33,33,33,33	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.
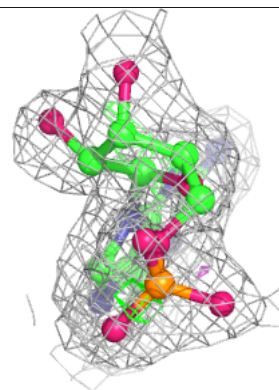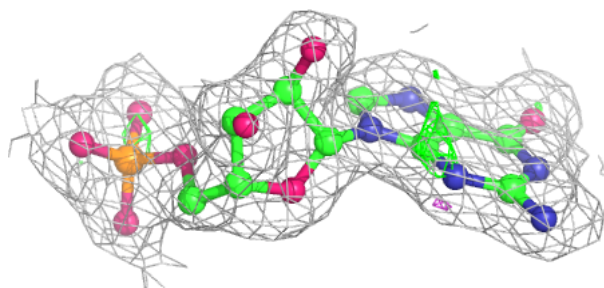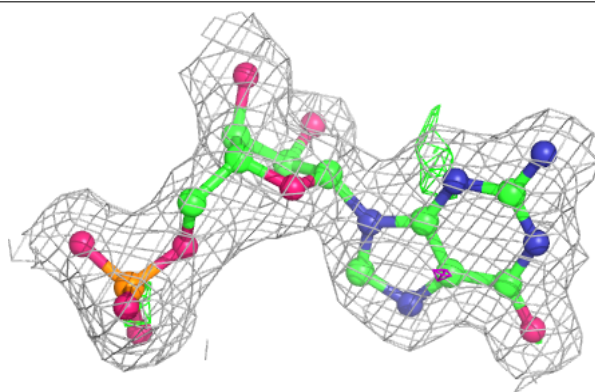
Electron density around 5GP C 1200:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
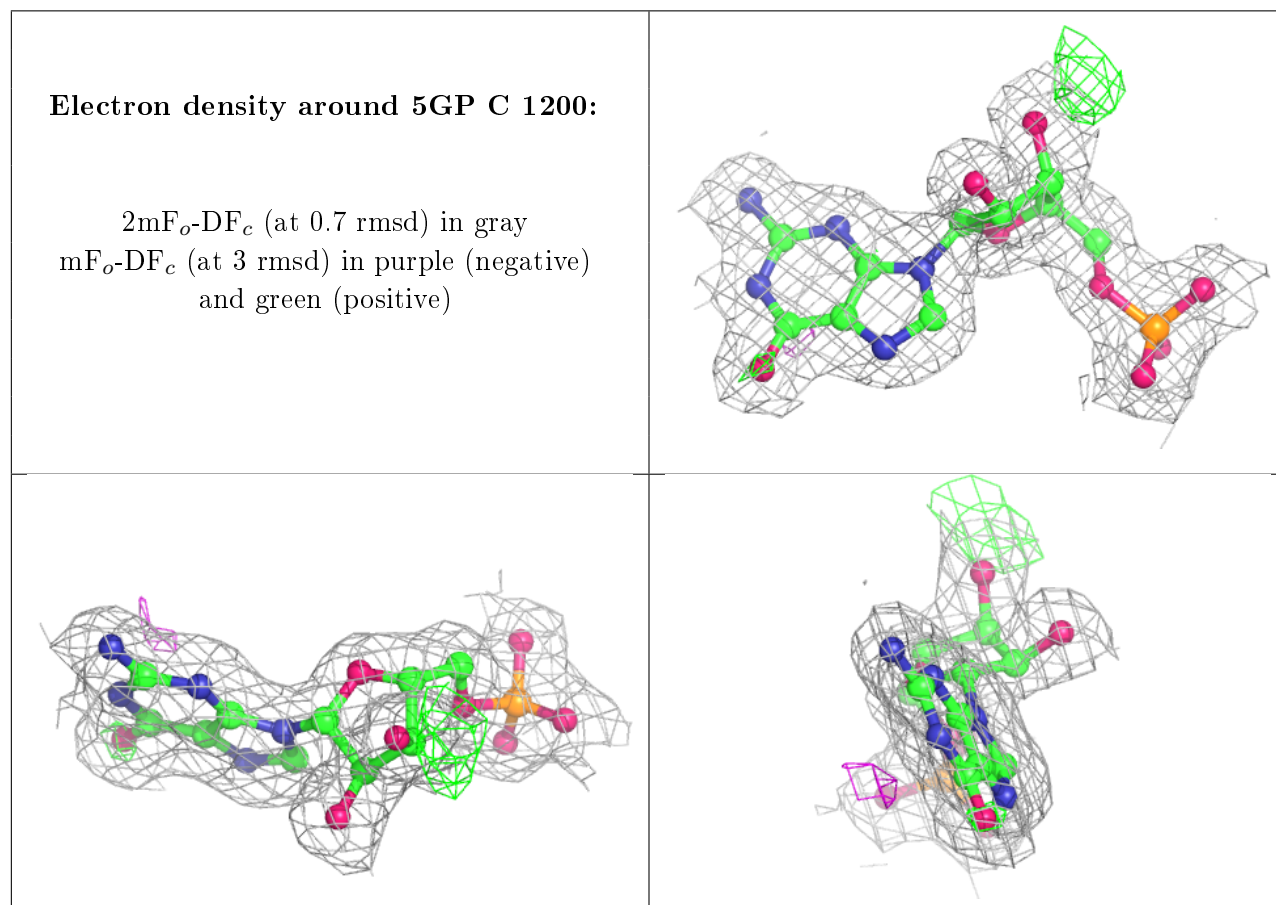
Electron density around 5GP D 1210:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
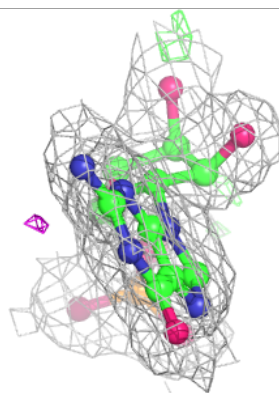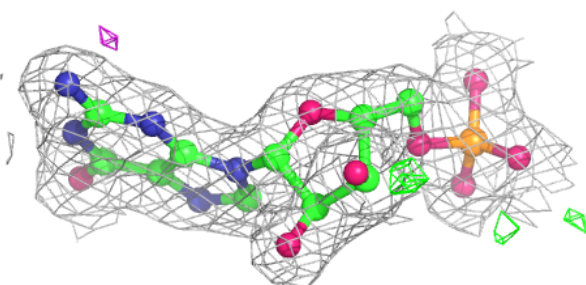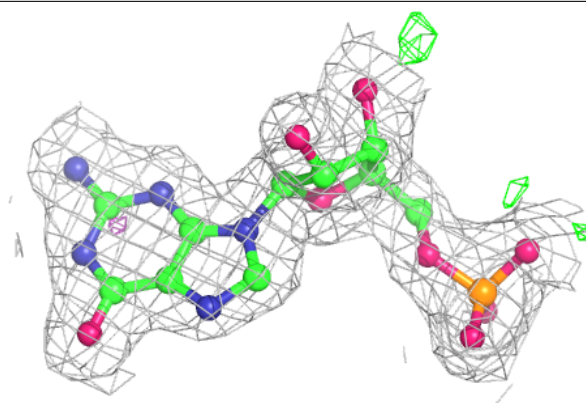
Electron density around 5GP B 1208:

$2mF_o$-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.