



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2023 – 12:49 PM EDT

PDB ID : 2OV2  
Title : The crystal structure of the human RAC3 in complex with the CRIB domain of human p21-activated kinase 4 (PAK4)  
Authors : Ugochukwu, E.; Yang, X.; Elkins, J.M.; Burgess-Brown, N.; Bunkoczi, G.; Debreczeni, J.E.D.; Sundstrom, M.; Arrowsmith, C.H.; Weigelt, J.; Edwards, A.; von Delft, F.; Knapp, S.; Doyle, D.A.; Structural Genomics Consortium (SGC)  
Deposited on : 2007-02-12  
Resolution : 2.10 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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)

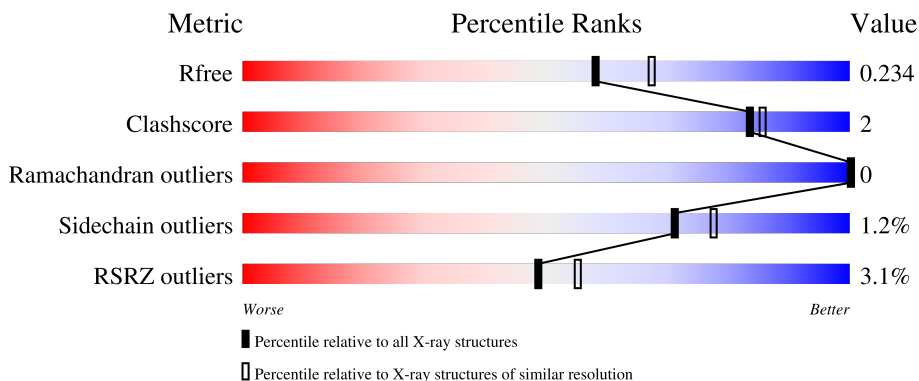
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


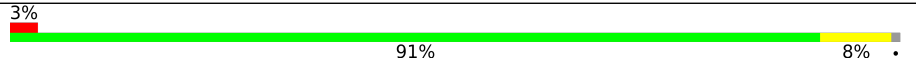
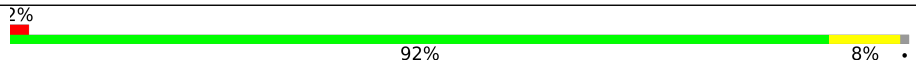
The reported resolution of this entry is 2.10 Å.

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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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  $\geq 3$ , 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	179	 3% 89% 9% ..
1	B	179	 3% 91% 8% .
1	C	179	 2% 92% 8% .

*Continued on next page...*

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
 Validation Pipeline (wwPDB-VP) : 2.35

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	179	<p>2% 95% 5%</p>
1	E	179	<p>1% 91% 8%</p>
1	F	179	<p>3% 93% 7%</p>
1	G	179	<p>3% 91% 9%</p>
1	H	179	<p>4% 96% 1%</p>
2	I	35	<p>6% 83% 14%</p>
2	J	35	<p>9% 94% 6%</p>
2	K	35	<p>3% 89% 9%</p>
2	L	35	<p>17% 97% 1%</p>
2	M	35	<p>3% 94% 3%</p>
2	N	35	<p>0% 91% 9%</p>
2	O	35	<p>6% 97% 1%</p>
2	P	35	<p>0% 94% 6%</p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 14479 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 Ras-related C3 botulinum toxin substrate 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	178	1368	879	224	257	8	0	3	0
1	F	179	1388	891	231	258	8	0	5	0
1	B	178	1384	888	231	258	7	0	2	0
1	C	178	1358	877	220	253	8	0	3	0
1	G	178	1369	878	228	255	8	0	2	0
1	D	179	1365	878	229	251	7	0	1	0
1	E	177	1355	872	225	251	7	0	1	0
1	H	178	1355	871	224	252	8	0	1	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	cloning artifact	UNP P60763
A	178	GLY	-	cloning artifact	UNP P60763
B	0	SER	-	cloning artifact	UNP P60763
B	178	GLY	-	cloning artifact	UNP P60763
C	0	SER	-	cloning artifact	UNP P60763
C	178	GLY	-	cloning artifact	UNP P60763
D	0	SER	-	cloning artifact	UNP P60763
D	178	GLY	-	cloning artifact	UNP P60763
E	0	SER	-	cloning artifact	UNP P60763
E	178	GLY	-	cloning artifact	UNP P60763
F	0	SER	-	cloning artifact	UNP P60763
F	178	GLY	-	cloning artifact	UNP P60763
G	0	SER	-	cloning artifact	UNP P60763

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	178	GLY	-	cloning artifact	UNP P60763
H	0	SER	-	cloning artifact	UNP P60763
H	178	GLY	-	cloning artifact	UNP P60763

- Molecule 2 is a protein called Serine/threonine-protein kinase PAK 4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	34	Total	C	N	O	0	0	0
			266	172	46	48			
2	J	35	Total	C	N	O	0	0	0
			264	171	46	47			
2	K	34	Total	C	N	O	0	1	0
			264	173	45	46			
2	L	35	Total	C	N	O	0	0	0
			267	171	47	49			
2	M	34	Total	C	N	O	0	0	0
			266	173	46	47			
2	N	35	Total	C	N	O	0	0	0
			272	174	47	51			
2	O	34	Total	C	N	O	0	0	0
			264	171	46	47			
2	P	33	Total	C	N	O	0	0	0
			254	164	44	46			

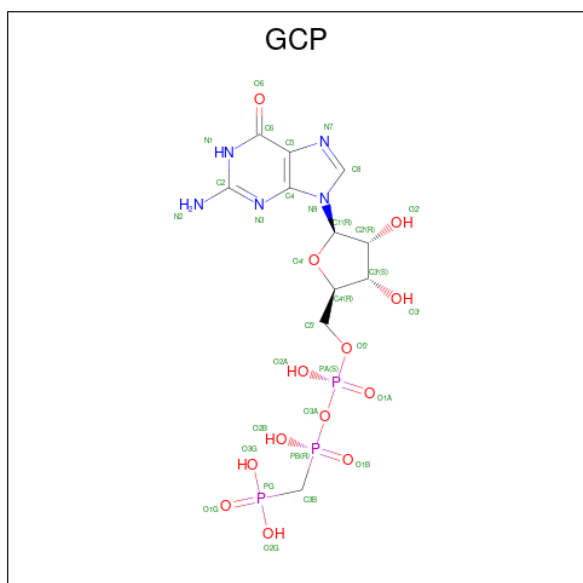
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		
3	F	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		
3	G	1	Total	Mg	0	0
			1	1		
3	D	1	Total	Mg	0	0
			1	1		
3	E	1	Total	Mg	0	0
			1	1		
3	H	1	Total	Mg	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Cl	0	0
			2	2		
4	B	1	Total	Cl	0	0
			1	1		
4	C	1	Total	Cl	0	0
			1	1		
4	E	1	Total	Cl	0	0
			1	1		
4	N	1	Total	Cl	0	0
			1	1		
4	O	1	Total	Cl	0	0
			1	1		
4	P	1	Total	Cl	0	0
			1	1		

- Molecule 5 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: GCP) (formula: C<sub>11</sub>H<sub>18</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



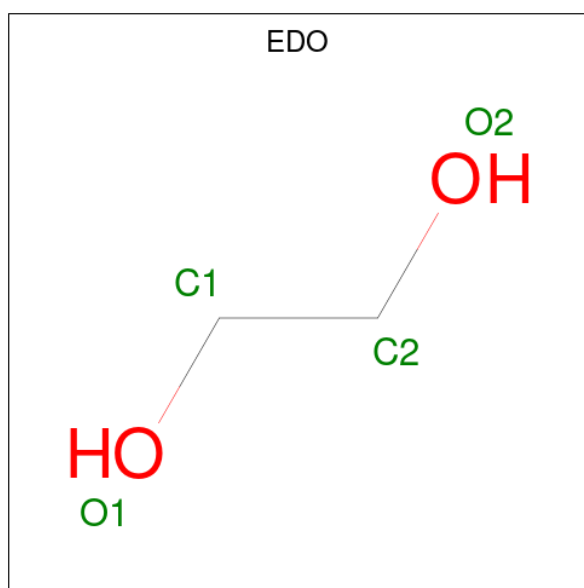
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
5	A	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	F	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	B	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	G	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	D	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	E	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
5	H	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
6	F	1	Total	C	O	0	0
			4	2	2		
6	C	1	Total	C	O	0	0
			4	2	2		
6	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	129	Total	O	0	1
			130	130		

Continued on next page...

*Continued from previous page...*

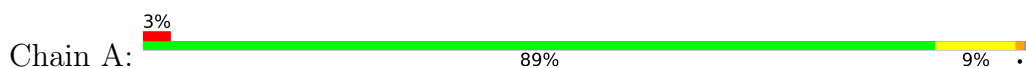
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	95	Total 95	O 95	0	0
7	B	143	Total 143	O 143	0	0
7	C	116	Total 116	O 116	0	0
7	G	122	Total 123	O 123	0	1
7	D	114	Total 114	O 114	0	0
7	E	134	Total 134	O 134	0	0
7	H	105	Total 106	O 106	0	1
7	I	24	Total 24	O 24	0	0
7	J	24	Total 24	O 24	0	0
7	K	16	Total 16	O 16	0	0
7	L	14	Total 14	O 14	0	0
7	M	28	Total 28	O 28	0	0
7	N	21	Total 22	O 22	0	1
7	O	21	Total 21	O 21	0	0
7	P	26	Total 26	O 26	0	0



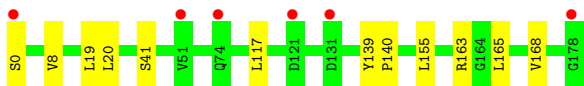
### 3 Residue-property plots [i](#)

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.

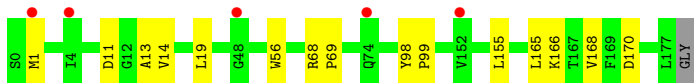
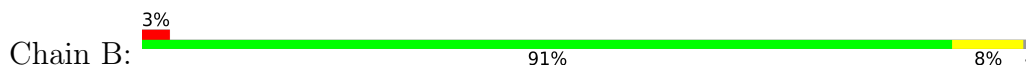
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



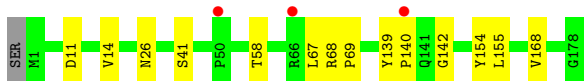
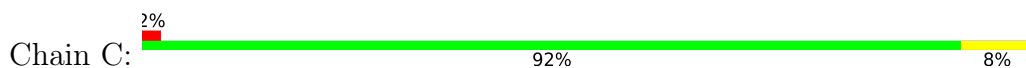
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



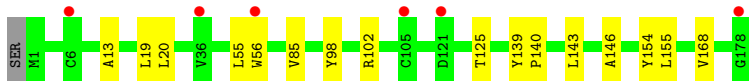
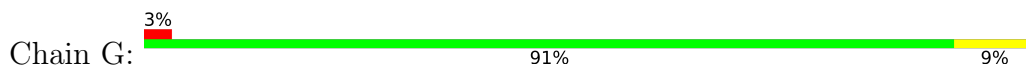
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



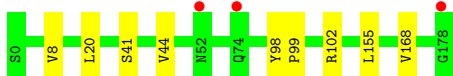
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



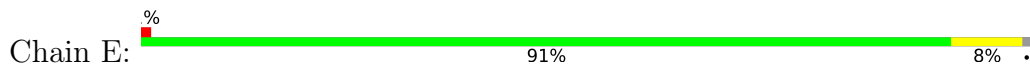
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



- Molecule 1: Ras-related C3 botulinum toxin substrate 3



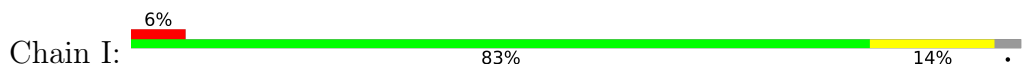
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



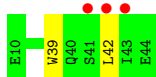
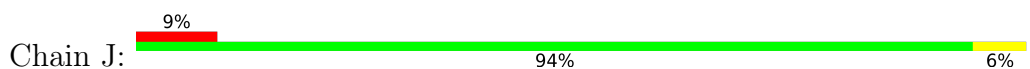
- Molecule 1: Ras-related C3 botulinum toxin substrate 3



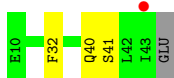
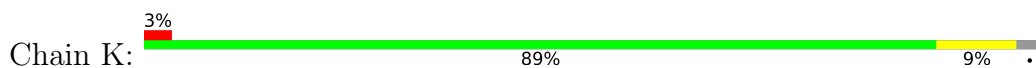
- Molecule 2: Serine/threonine-protein kinase PAK 4



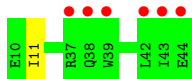
- Molecule 2: Serine/threonine-protein kinase PAK 4



- Molecule 2: Serine/threonine-protein kinase PAK 4



- Molecule 2: Serine/threonine-protein kinase PAK 4




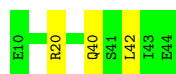
- Molecule 2: Serine/threonine-protein kinase PAK 4

Chain M:  94% 3%



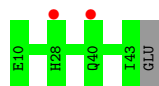
- Molecule 2: Serine/threonine-protein kinase PAK 4

Chain N:  91% 9%



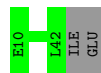
- Molecule 2: Serine/threonine-protein kinase PAK 4

Chain O:  97% 6%



- Molecule 2: Serine/threonine-protein kinase PAK 4

Chain P:  94% 6%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.84Å 73.86Å 133.84Å 88.58° 87.76° 70.64°	Depositor
Resolution (Å)	62.11 – 2.10 62.11 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (62.11-2.10) 98.9 (62.11-2.10)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.172 , 0.225 0.183 , 0.234	Depositor DCC
$R_{free}$ test set	5292 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.5	Xtrriage
Anisotropy	0.089	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 47.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.016 for h,h-k,-l 0.010 for -h,-k,l 0.006 for -h,-h+k,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14479	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage'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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: GCP, EDO, CL, MG

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	0.57	0/1408	0.66	0/1921
1	B	0.54	0/1420	0.67	0/1936
1	C	0.53	0/1398	0.62	0/1909
1	D	0.51	0/1399	0.67	0/1908
1	E	0.55	0/1389	0.69	1/1895 (0.1%)
1	F	0.53	0/1430	0.62	0/1947
1	G	0.53	0/1406	0.65	0/1918
1	H	0.52	0/1388	0.63	0/1893
2	I	0.51	0/275	0.55	0/374
2	J	0.52	0/273	0.56	0/371
2	K	0.47	0/276	0.63	0/375
2	L	0.45	0/276	0.51	0/375
2	M	0.47	0/275	0.55	0/373
2	N	0.49	0/281	0.55	0/381
2	O	0.45	0/273	0.51	0/371
2	P	0.44	0/263	0.55	0/358
All	All	0.53	0/13430	0.64	1/18305 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	128	ARG	NE-CZ-NH2	-5.74	117.43	120.30

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	1368	0	1345	11	0
1	B	1384	0	1372	9	0
1	C	1358	0	1338	7	0
1	D	1365	0	1347	5	0
1	E	1355	0	1330	9	0
1	F	1388	0	1374	6	0
1	G	1369	0	1349	8	0
1	H	1355	0	1329	5	0
2	I	266	0	232	3	0
2	J	264	0	224	2	0
2	K	264	0	233	1	0
2	L	267	0	229	2	0
2	M	266	0	237	0	0
2	N	272	0	232	1	0
2	O	264	0	233	0	0
2	P	254	0	215	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	N	1	0	0	0	0
4	O	1	0	0	0	0
4	P	1	0	0	0	0
5	A	32	0	14	1	0
5	B	32	0	14	1	0
5	C	32	0	14	0	0
5	D	32	0	14	0	0
5	E	32	0	14	1	0
5	F	32	0	14	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	32	0	14	1	0
5	H	32	0	14	1	0
6	C	4	0	6	0	0
6	D	4	0	6	0	0
6	F	4	0	6	0	0
7	A	130	0	0	0	0
7	B	143	0	0	1	0
7	C	116	0	0	0	0
7	D	114	0	0	0	0
7	E	134	0	0	0	0
7	F	95	0	0	0	0
7	G	123	0	0	0	0
7	H	106	0	0	0	1
7	I	24	0	0	1	0
7	J	24	0	0	0	0
7	K	16	0	0	0	0
7	L	14	0	0	0	0
7	M	28	0	0	0	0
7	N	22	0	0	0	0
7	O	21	0	0	0	1
7	P	26	0	0	0	0
All	All	14479	0	12749	63	1

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.

The worst 5 of 63 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:11:ASP:O	1:B:14:VAL:HG22	1.94	0.67
2:J:39:TRP:HA	2:J:42:LEU:HD12	1.76	0.66
1:C:155:LEU:HD13	1:C:168:VAL:HA	1.80	0.63
1:B:1:MET:CE	1:B:1:MET:CG	2.79	0.60
1:C:11:ASP:O	1:C:14:VAL:HG13	2.02	0.59

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:H:297:HOH:O	7:O:175:HOH:O[1_554]	1.97	0.23

## 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	179/179 (100%)	176 (98%)	3 (2%)	0	100	100
1	B	178/179 (99%)	175 (98%)	3 (2%)	0	100	100
1	C	179/179 (100%)	176 (98%)	3 (2%)	0	100	100
1	D	178/179 (99%)	175 (98%)	3 (2%)	0	100	100
1	E	176/179 (98%)	174 (99%)	2 (1%)	0	100	100
1	F	181/179 (101%)	177 (98%)	4 (2%)	0	100	100
1	G	178/179 (99%)	176 (99%)	2 (1%)	0	100	100
1	H	175/179 (98%)	173 (99%)	2 (1%)	0	100	100
2	I	32/35 (91%)	32 (100%)	0	0	100	100
2	J	33/35 (94%)	33 (100%)	0	0	100	100
2	K	33/35 (94%)	33 (100%)	0	0	100	100
2	L	33/35 (94%)	33 (100%)	0	0	100	100
2	M	32/35 (91%)	32 (100%)	0	0	100	100
2	N	33/35 (94%)	32 (97%)	1 (3%)	0	100	100
2	O	32/35 (91%)	32 (100%)	0	0	100	100
2	P	31/35 (89%)	31 (100%)	0	0	100	100
All	All	1683/1712 (98%)	1660 (99%)	23 (1%)	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	146/154 (95%)	144 (99%)	2 (1%)	67	73
1	B	149/154 (97%)	149 (100%)	0	100	100
1	C	144/154 (94%)	141 (98%)	3 (2%)	53	59
1	D	143/154 (93%)	141 (99%)	2 (1%)	67	73
1	E	142/154 (92%)	142 (100%)	0	100	100
1	F	148/154 (96%)	147 (99%)	1 (1%)	84	88
1	G	146/154 (95%)	144 (99%)	2 (1%)	67	73
1	H	143/154 (93%)	143 (100%)	0	100	100
2	I	26/32 (81%)	24 (92%)	2 (8%)	13	9
2	J	24/32 (75%)	24 (100%)	0	100	100
2	K	25/32 (78%)	22 (88%)	3 (12%)	5	2
2	L	26/32 (81%)	26 (100%)	0	100	100
2	M	26/32 (81%)	25 (96%)	1 (4%)	33	34
2	N	26/32 (81%)	24 (92%)	2 (8%)	13	9
2	O	26/32 (81%)	26 (100%)	0	100	100
2	P	24/32 (75%)	24 (100%)	0	100	100
All	All	1364/1488 (92%)	1346 (99%)	18 (1%)	71	75

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	K	41[B]	SER
2	N	42	LEU
2	N	40	GLN
1	D	41	SER
2	K	41[A]	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	162	GLN
1	H	141	GLN
2	L	16	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 16 are monoatomic - leaving 11 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
6	EDO	F	1301	-	3,3,3	0.42	0	2,2,2	0.38	0
5	GCP	E	200	3	27,34,34	1.32	4 (14%)	34,54,54	2.14	9 (26%)
6	EDO	C	1302	-	3,3,3	0.48	0	2,2,2	0.27	0
5	GCP	F	200	3	27,34,34	1.42	3 (11%)	34,54,54	2.17	11 (32%)
5	GCP	G	200	3	27,34,34	1.47	7 (25%)	34,54,54	1.87	10 (29%)
5	GCP	D	200	3	27,34,34	1.21	3 (11%)	34,54,54	1.99	8 (23%)
6	EDO	D	1303	-	3,3,3	0.44	0	2,2,2	0.21	0
5	GCP	C	200	3	27,34,34	1.26	3 (11%)	34,54,54	2.00	9 (26%)
5	GCP	A	200	3	27,34,34	1.19	2 (7%)	34,54,54	2.08	10 (29%)
5	GCP	H	200	3	27,34,34	1.53	5 (18%)	34,54,54	2.17	11 (32%)
5	GCP	B	200	3	27,34,34	1.45	4 (14%)	34,54,54	1.99	11 (32%)

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
6	EDO	F	1301	-	-	0/1/1/1	-
5	GCP	E	200	3	-	2/15/38/38	0/3/3/3
6	EDO	C	1302	-	-	0/1/1/1	-
5	GCP	F	200	3	-	2/15/38/38	0/3/3/3
5	GCP	G	200	3	-	0/15/38/38	0/3/3/3
5	GCP	D	200	3	-	0/15/38/38	0/3/3/3
6	EDO	D	1303	-	-	0/1/1/1	-
5	GCP	C	200	3	-	0/15/38/38	0/3/3/3
5	GCP	A	200	3	-	0/15/38/38	0/3/3/3
5	GCP	H	200	3	-	2/15/38/38	0/3/3/3
5	GCP	B	200	3	-	0/15/38/38	0/3/3/3

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	200	GCP	C5-C6	4.83	1.49	1.41
5	F	200	GCP	C5-C6	4.47	1.49	1.41
5	B	200	GCP	C5-C6	4.35	1.48	1.41
5	C	200	GCP	C5-C6	4.19	1.48	1.41
5	E	200	GCP	C5-C6	3.73	1.47	1.41

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	200	GCP	C4-C5-C6	-5.47	115.58	120.80
5	F	200	GCP	C2-N1-C6	5.22	124.22	115.93
5	A	200	GCP	O1G-PG-C3B	-5.17	100.10	111.24
5	E	200	GCP	C2-N1-C6	5.08	124.00	115.93
5	H	200	GCP	C4-C5-C6	-5.04	115.99	120.80

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

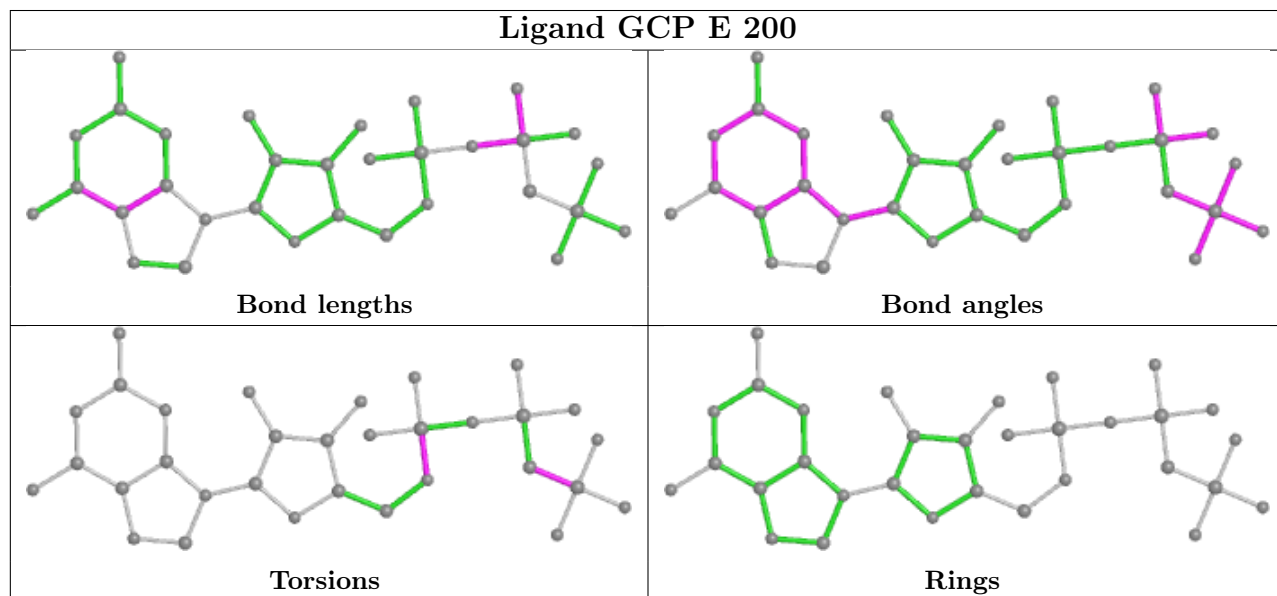
Mol	Chain	Res	Type	Atoms
5	H	200	GCP	C5'-O5'-PA-O1A
5	H	200	GCP	C5'-O5'-PA-O3A
5	F	200	GCP	PB-C3B-PG-O1G
5	E	200	GCP	PB-C3B-PG-O1G
5	F	200	GCP	C5'-O5'-PA-O1A

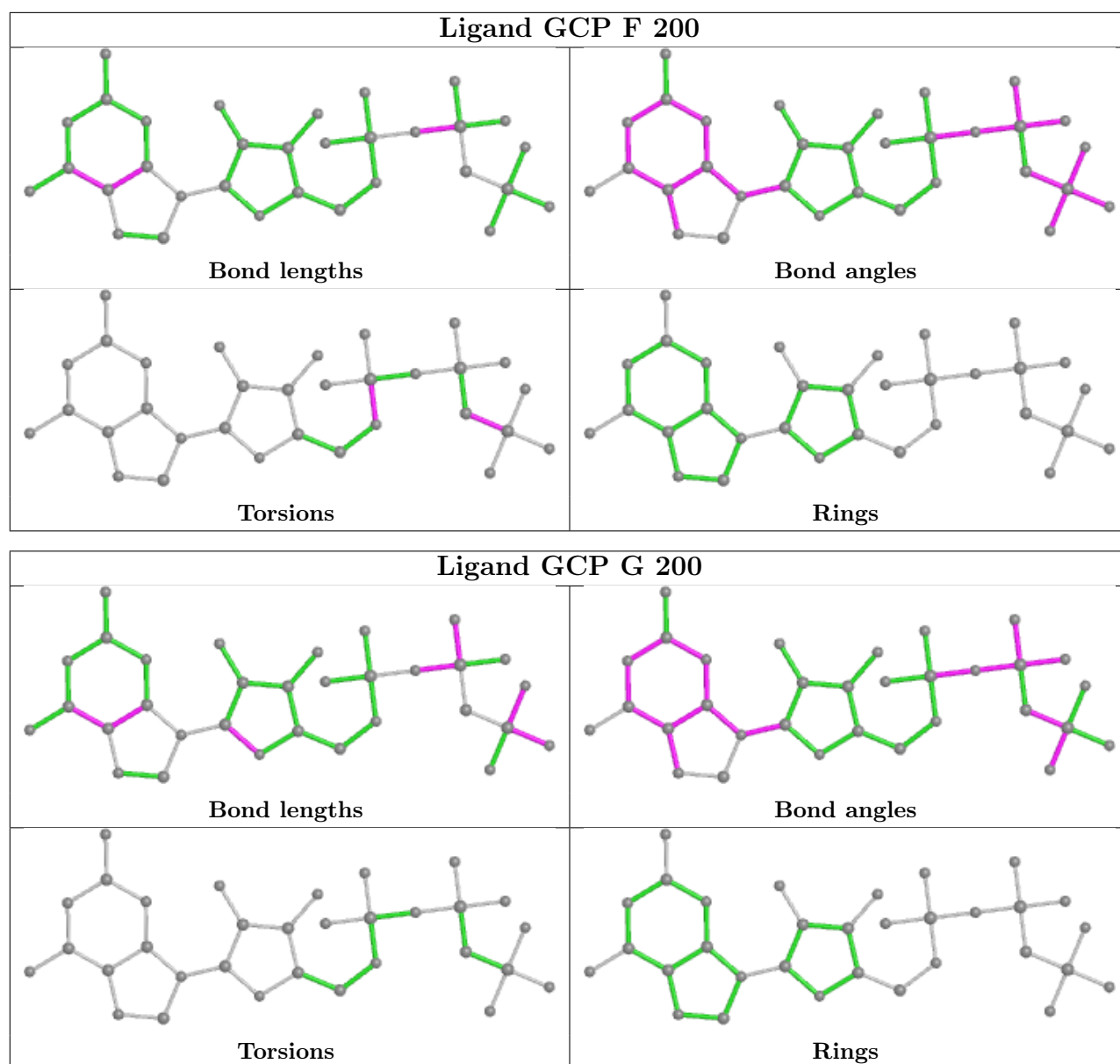
There are no ring outliers.

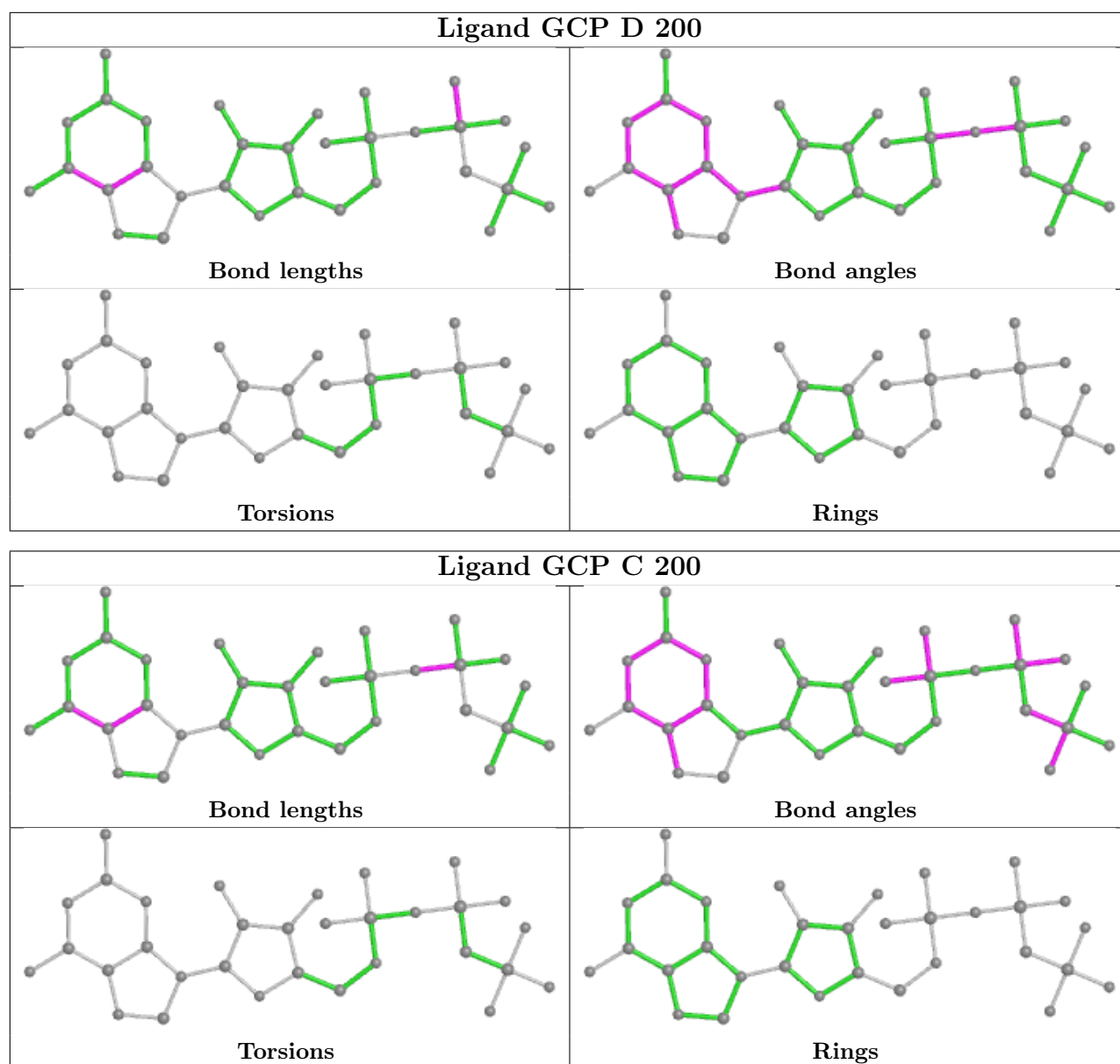
5 monomers are involved in 5 short contacts:

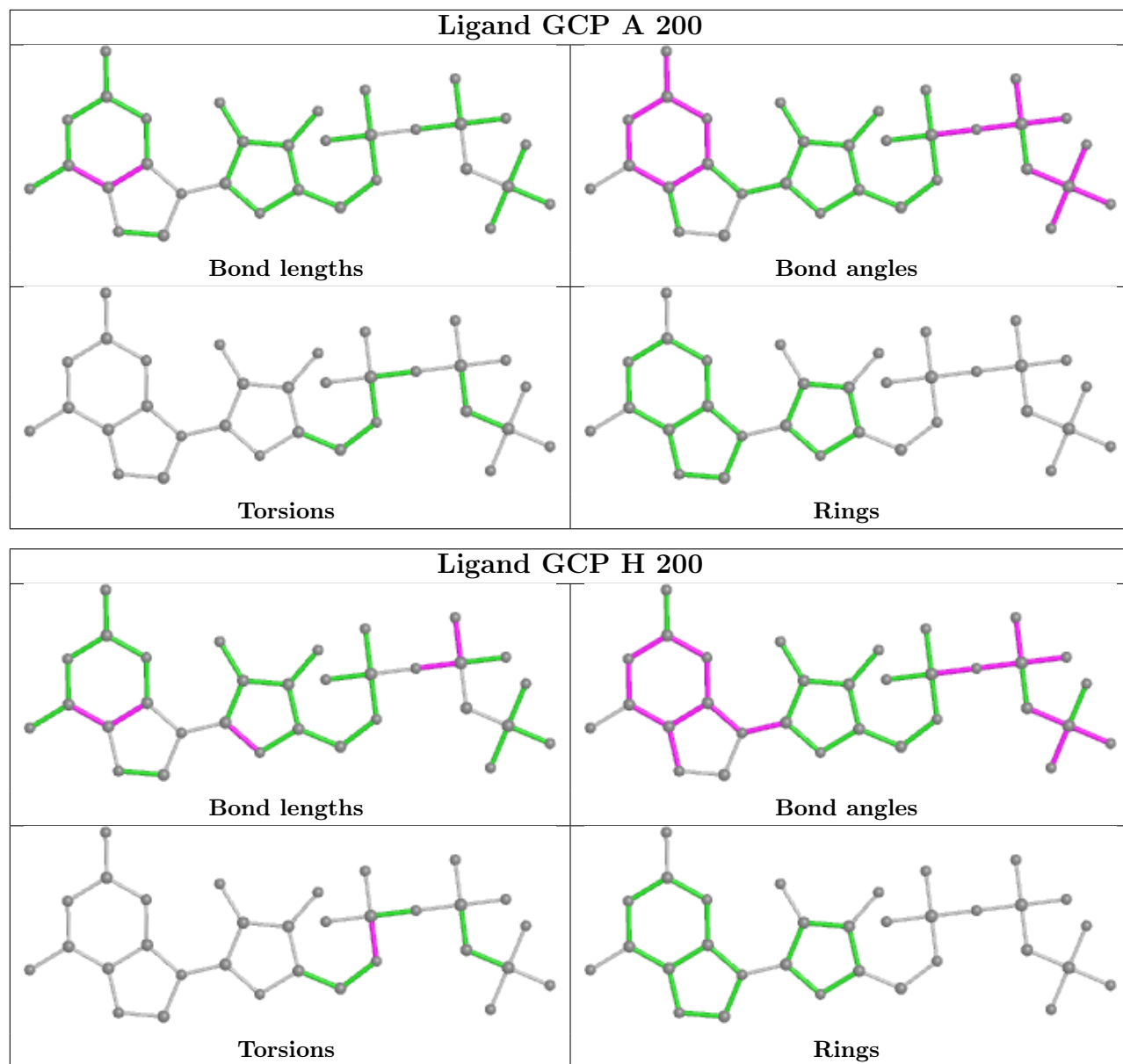
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	200	GCP	1	0
5	G	200	GCP	1	0
5	A	200	GCP	1	0
5	H	200	GCP	1	0
5	B	200	GCP	1	0

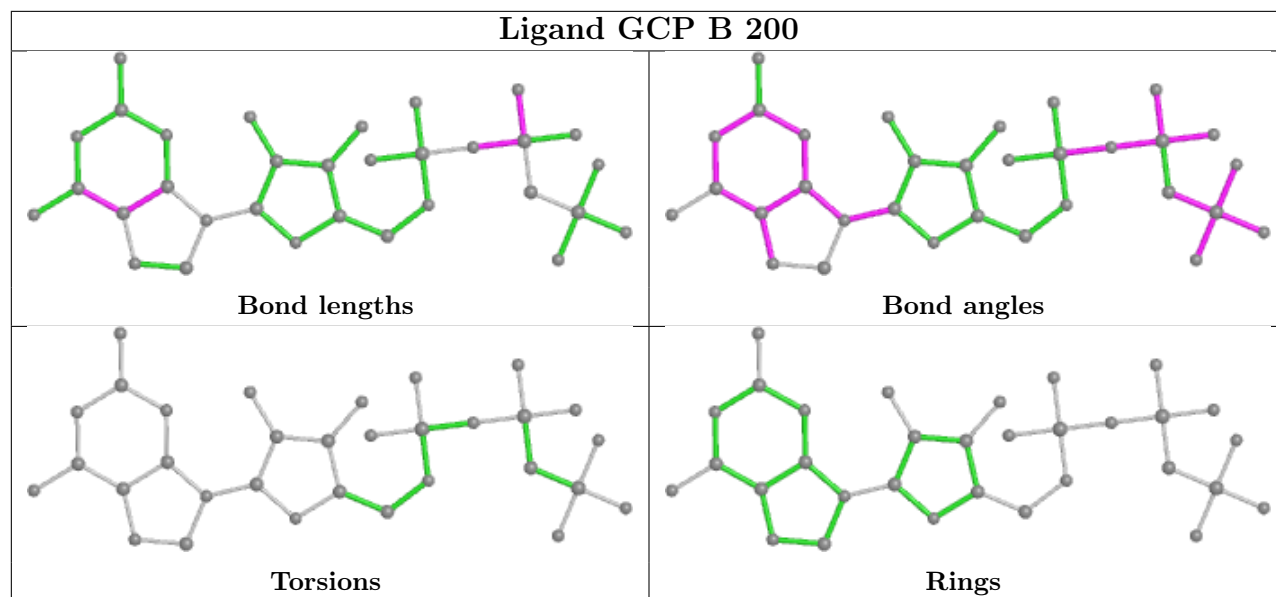
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	178/179 (99%)	0.67	5 (2%) 53 59	34, 38, 45, 48	0
1	B	178/179 (99%)	0.63	5 (2%) 53 59	34, 38, 45, 47	0
1	C	178/179 (99%)	0.67	3 (1%) 70 74	34, 38, 44, 47	0
1	D	179/179 (100%)	0.68	3 (1%) 70 74	34, 38, 46, 48	0
1	E	177/179 (98%)	0.59	1 (0%) 89 91	35, 38, 44, 47	0
1	F	179/179 (100%)	0.72	6 (3%) 45 51	35, 38, 45, 48	0
1	G	178/179 (99%)	0.68	6 (3%) 45 51	34, 38, 45, 49	0
1	H	178/179 (99%)	0.80	8 (4%) 33 38	35, 39, 45, 48	0
2	I	34/35 (97%)	0.72	2 (5%) 22 27	33, 39, 49, 50	0
2	J	35/35 (100%)	0.76	3 (8%) 10 13	32, 40, 51, 52	0
2	K	34/35 (97%)	0.52	1 (2%) 51 57	34, 40, 49, 50	0
2	L	35/35 (100%)	1.10	6 (17%) 1 1	31, 40, 48, 49	0
2	M	34/35 (97%)	0.74	1 (2%) 51 57	33, 40, 49, 49	0
2	N	35/35 (100%)	0.66	0 100 100	34, 41, 50, 50	0
2	O	34/35 (97%)	0.51	2 (5%) 22 27	34, 40, 49, 50	0
2	P	33/35 (94%)	0.83	0 100 100	33, 40, 49, 50	0
All	All	1699/1712 (99%)	0.69	52 (3%) 49 55	31, 39, 46, 52	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	48	GLY	4.1
2	I	43	ILE	4.1
1	F	178[A]	GLY	3.8
1	C	140	PRO	3.4
1	G	178	GLY	3.4

## 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](#)

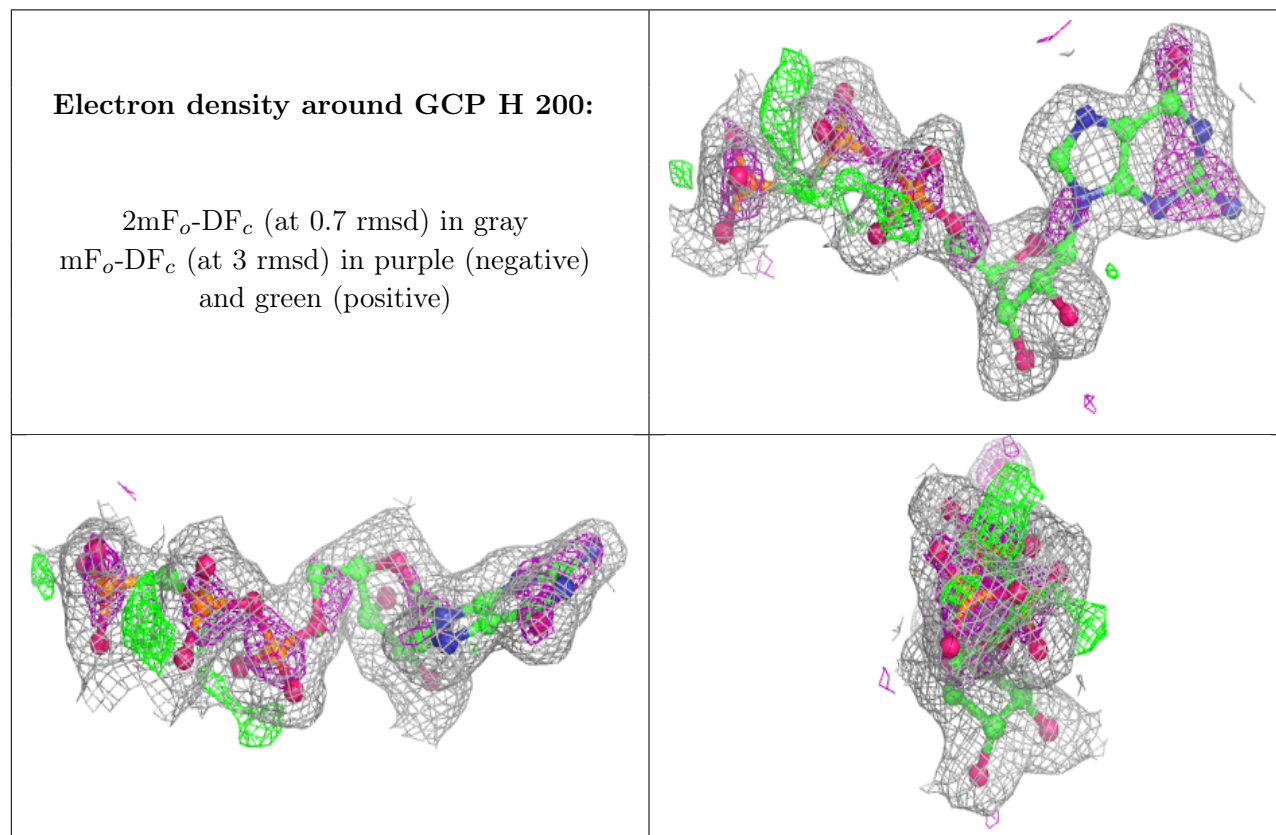
There are no monosaccharides in this entry.

## 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, 95<sup>th</sup> 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.

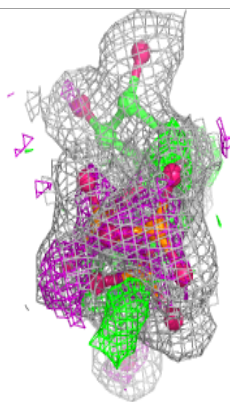
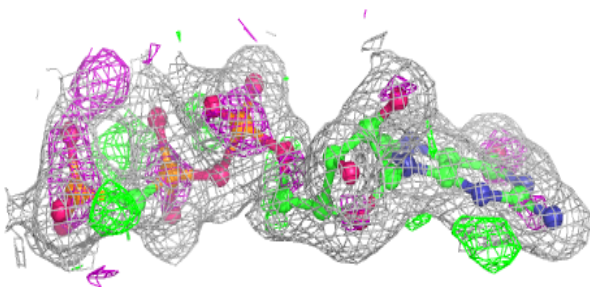
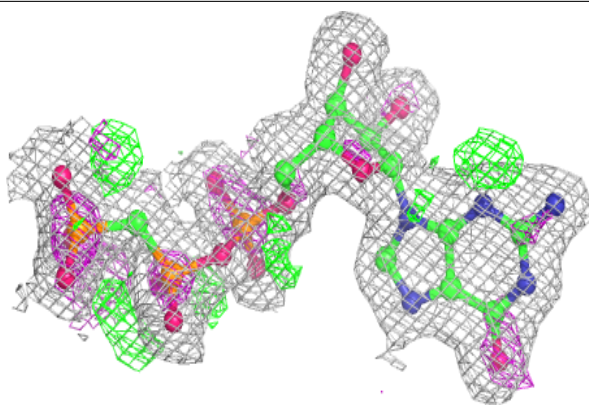
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CL	E	1208	1/1	0.67	0.12	69,69,69,69	0
4	CL	N	1206	1/1	0.79	0.11	84,84,84,84	0
4	CL	A	1207	1/1	0.83	0.12	69,69,69,69	0
6	EDO	F	1301	4/4	0.88	0.14	32,33,33,36	0
4	CL	A	1201	1/1	0.90	0.09	65,65,65,65	0
4	CL	B	1204	1/1	0.91	0.12	65,65,65,65	0
3	MG	A	202	1/1	0.91	0.06	27,27,27,27	0
4	CL	O	1209	1/1	0.92	0.06	60,60,60,60	0
3	MG	F	202	1/1	0.92	0.07	28,28,28,28	0
4	CL	C	1203	1/1	0.94	0.07	59,59,59,59	0
6	EDO	D	1303	4/4	0.94	0.12	35,36,36,36	0
4	CL	P	1205	1/1	0.96	0.07	60,60,60,60	0
5	GCP	H	200	32/32	0.96	0.16	25,33,35,35	0
3	MG	H	202	1/1	0.96	0.07	27,27,27,27	0
6	EDO	C	1302	4/4	0.96	0.10	29,29,31,33	0
3	MG	B	202	1/1	0.96	0.17	23,23,23,23	0
5	GCP	B	200	32/32	0.97	0.10	22,26,27,27	0
5	GCP	C	200	32/32	0.97	0.11	22,25,27,30	0
5	GCP	G	200	32/32	0.97	0.12	19,27,28,29	0
5	GCP	D	200	32/32	0.97	0.11	24,27,28,29	0
5	GCP	E	200	32/32	0.97	0.13	20,28,29,30	0
3	MG	E	202	1/1	0.97	0.09	28,28,28,28	0
3	MG	C	202	1/1	0.97	0.18	24,24,24,24	0
5	GCP	A	200	32/32	0.97	0.11	20,23,25,29	0
5	GCP	F	200	32/32	0.97	0.15	23,27,31,31	0
3	MG	D	202	1/1	0.98	0.15	28,28,28,28	0
3	MG	G	202	1/1	0.98	0.13	25,25,25,25	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.

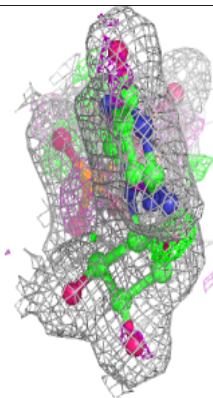
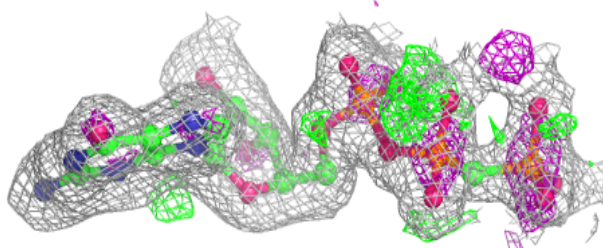
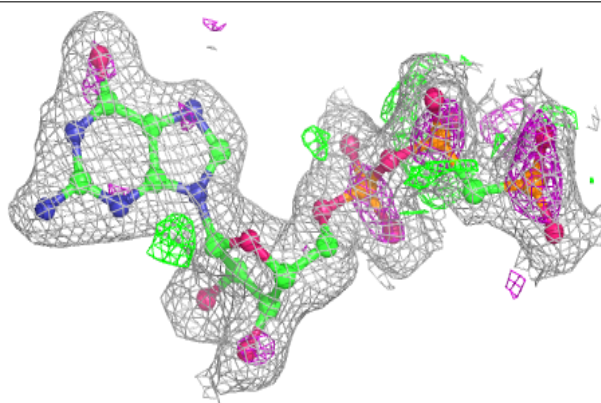


**Electron density around GCP B 200:**

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

**Electron density around GCP C 200:**

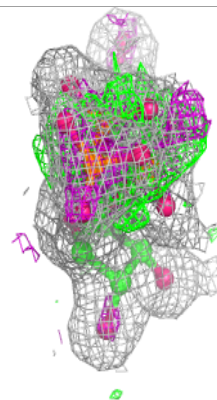
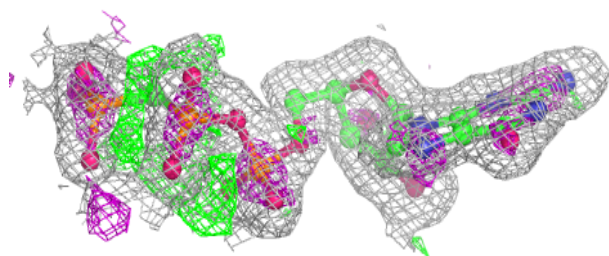
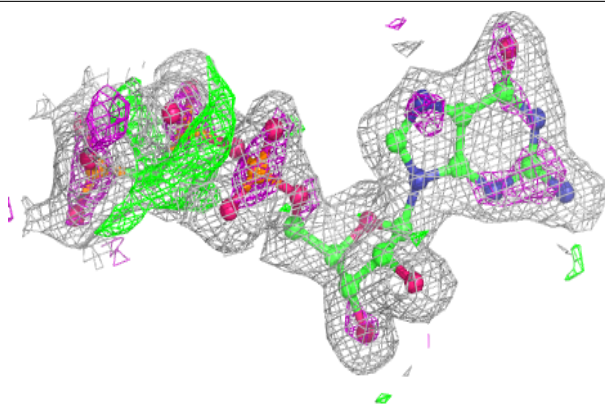
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



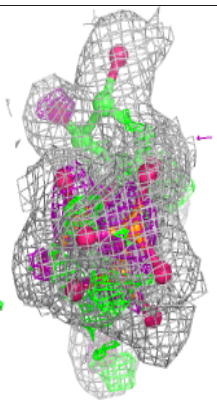
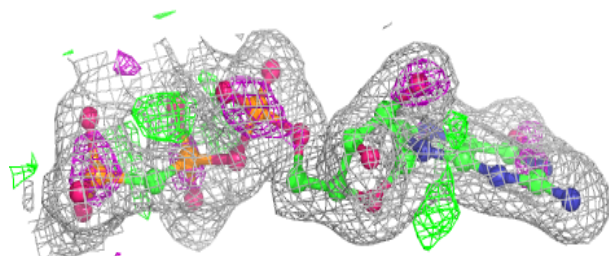
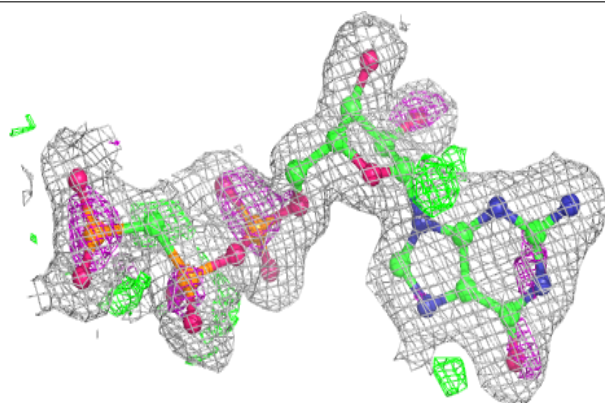


**Electron density around GCP G 200:**

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

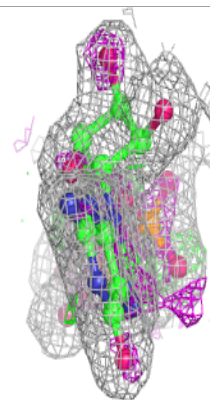
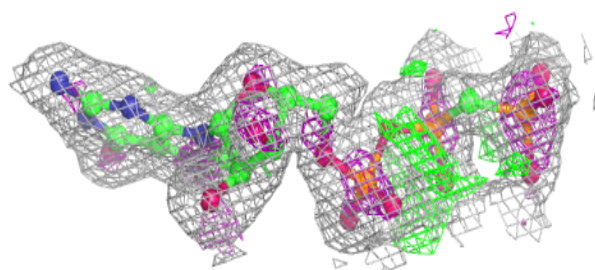
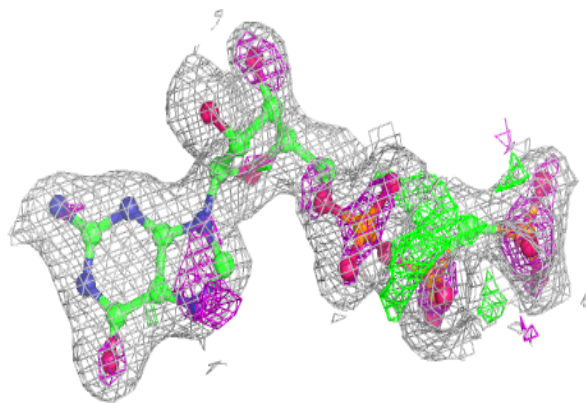
**Electron density around GCP D 200:**

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

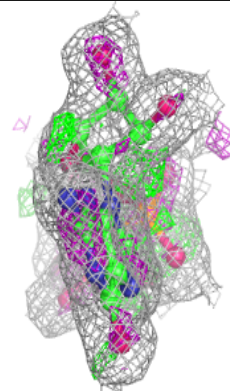
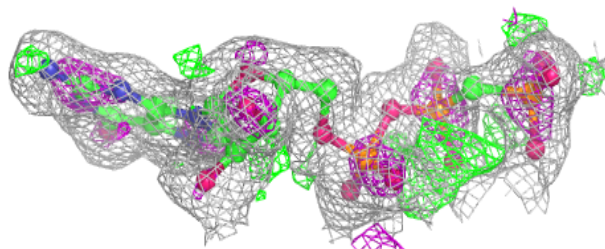
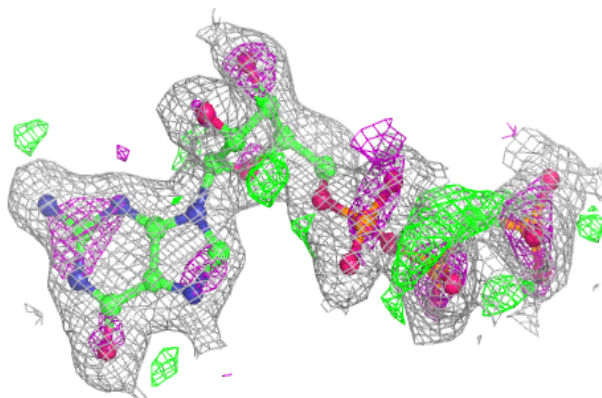


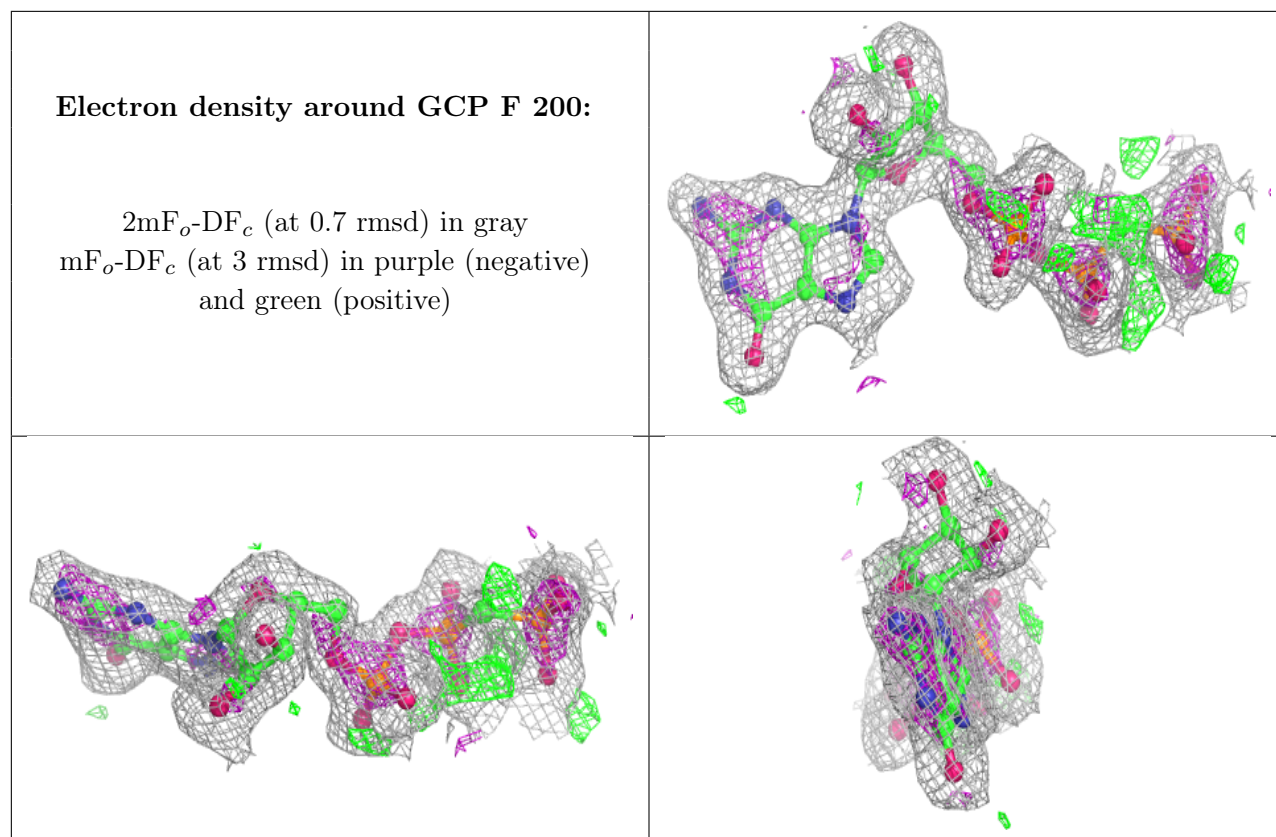
**Electron density around GCP E 200:**

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

**Electron density around GCP A 200:**

$2mF_o-DF_c$  (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.