W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 03:27 pm BST

PDB ID	:	6F2H
Title	:	Structure of Protease 1 from Pyrococcus horikoshii co-crystallized in presence of 10 mM Tb-Xo4 and potassium iodide.
Authors	:	Engilberge, S.; Riobe, F.; Di Pietro, S.; Franzetti, B.; Girard, E.; Dumont, E.; Maury, O.
Deposited on	:	2017-11-24
Resolution	:	2.19 Å(reported)

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:
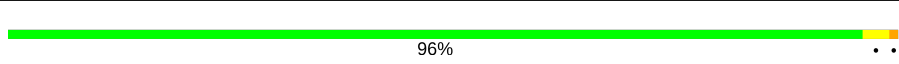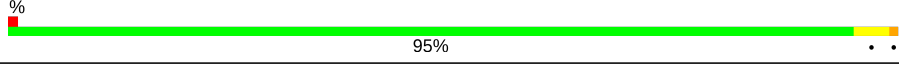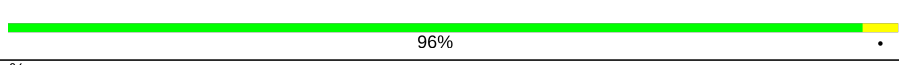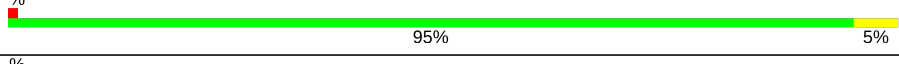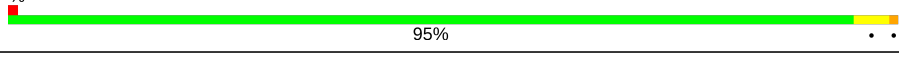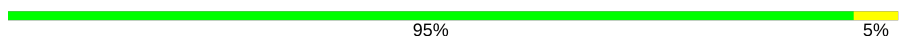*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.19 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		0.196
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		2.7%
RSRZ outliers		0.5%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
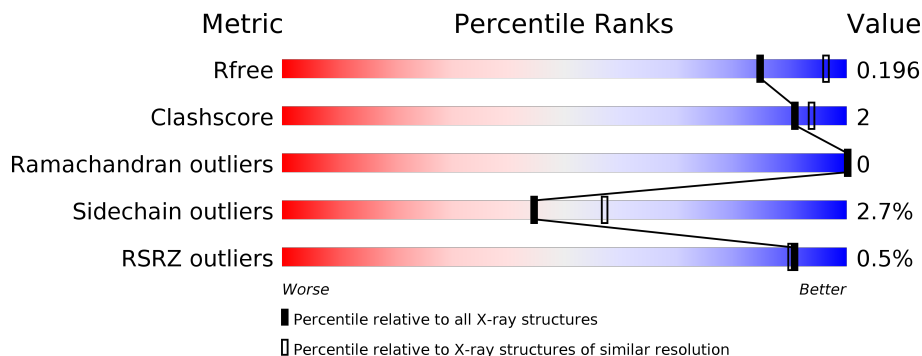
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	166	96%
1	B	166	95%
1	C	166	96%
1	D	166	95% 5%
1	E	166	95%
1	F	166	95% 5%

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	166	94% 5%
1	H	166	93% 7%
1	I	166	92% 7%
1	J	166	92% 8%
1	K	166	90% 9%
1	L	166	95% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

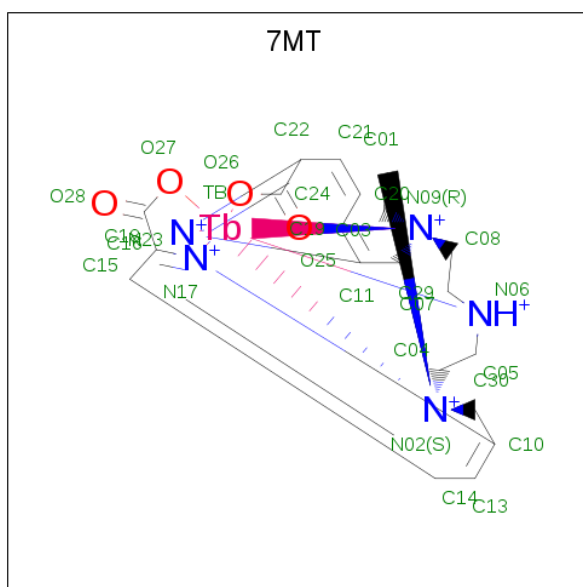| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	IOD	B	202	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 32784 atoms, of which 15986 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 Deglycase PH1704.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	B	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	C	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	D	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	E	166	Total 2628	C 843	H 1314	N 225	O 241	S 5	0	0	0
1	F	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	G	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	H	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	I	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	J	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	K	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0
1	L	166	Total 2629	C 843	H 1315	N 225	O 241	S 5	0	0	0

- Molecule 2 is Tb-Xo4 (three-letter code: 7MT) (formula: $C_{20}H_{23}N_5O_4Tb$).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	B	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	E	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	F	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	G	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	I	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	J	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	K	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		
2	L	1	Total	C	H	N	O	Tb	0	0
			53	20	23	5	4	1		

- Molecule 3 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total	I	0	0
			1	1		
3	J	1	Total	I	0	0
			1	1		
3	K	1	Total	I	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total 1	I 1	0	0
3	H	1	Total 1	I 1	0	0
3	B	1	Total 1	I 1	0	0
3	I	1	Total 1	I 1	0	0
3	C	1	Total 1	I 1	0	0
3	A	1	Total 1	I 1	0	0
3	L	1	Total 1	I 1	0	0
3	F	1	Total 1	I 1	0	0

- Molecule 4 is TERBIUM(III) ION (three-letter code: TB) (formula: Tb).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	J	1	Total 1	Tb 1	0	0
4	D	1	Total 1	Tb 1	0	0
4	K	1	Total 1	Tb 1	0	0
4	E	2	Total 2	Tb 2	0	0
4	H	1	Total 1	Tb 1	0	0
4	C	1	Total 1	Tb 1	0	0
4	F	2	Total 2	Tb 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	73	Total 73	O 73	0	0
5	B	73	Total 73	O 73	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	76	Total 76	O 76	0	0
5	D	75	Total 75	O 75	0	0
5	E	76	Total 76	O 76	0	0
5	F	58	Total 58	O 58	0	0
5	G	82	Total 82	O 82	0	0
5	H	36	Total 36	O 36	0	0
5	I	36	Total 36	O 36	0	0
5	J	50	Total 50	O 50	0	0
5	K	62	Total 62	O 62	0	0
5	L	43	Total 43	O 43	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.
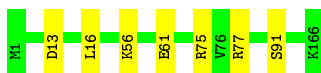
- Molecule 1: Deglycase PH1704

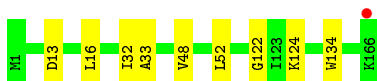Chain A:
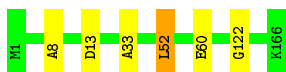
- Molecule 1: Deglycase PH1704

Chain B:
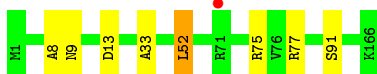
- Molecule 1: Deglycase PH1704

Chain C:

- Molecule 1: Deglycase PH1704

Chain D:

- Molecule 1: Deglycase PH1704

Chain E:

- Molecule 1: Deglycase PH1704

Chain F:

95% 5%

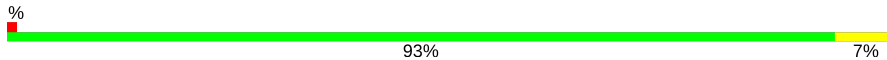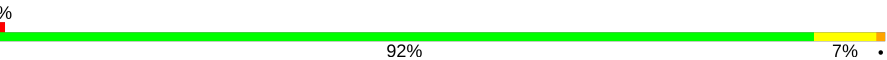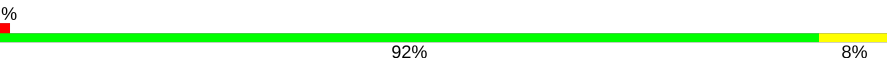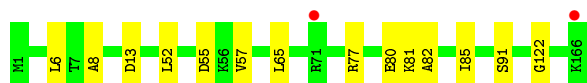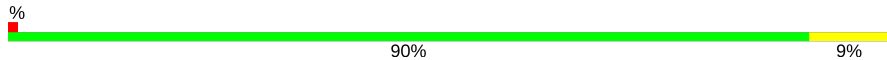M1 L6 T7 A8 D13 A33 L52 L65 S91 D126 K166

• Molecule 1: Deglycase PH1704

Chain G:

94% 5%

M1 K2 L6 T7 A8 D13 L52 E61 L65 R71 R77 E80 K124 K166

• Molecule 1: Deglycase PH1704

Chain H:

93% 7%

M1 L6 D13 E26 I32 A33 V48 L52 L65 R71 L112 D126 V132 K166

• Molecule 1: Deglycase PH1704

Chain I:

92% 7%

M1 A8 E12 D13 I32 A33 V48 L52 D63 R71 R77 S91 L106 R113 L154 K166

• Molecule 1: Deglycase PH1704

Chain J:

92% 8%

M1 L6 T7 A8 D13 L52 D55 K56 V57 L65 R71 R77 E80 K81 A82 I85 S91 G122 K166

• Molecule 1: Deglycase PH1704

Chain K:

90% 9%

M1 A8 D13 A33 L52 T53 R71 R75 V76 R77 R87 A109 R113 G122 G131 V135 E138 V139 V140 K166

• Molecule 1: Deglycase PH1704

Chain L:

95% 5%

M1 A8 D13 A33 L52 E60 E80 R113 G122 I123 K124 K166
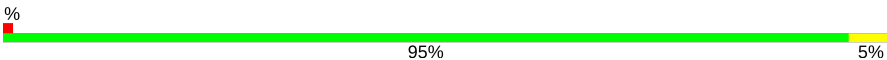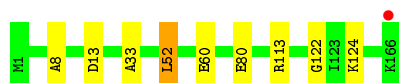
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	82.58Å 84.70Å 144.34Å 90.00° 90.97° 90.00°	Depositor
Resolution (Å)	46.01 – 2.19 46.01 – 2.19	Depositor EDS
% Data completeness (in resolution range)	98.6 (46.01-2.19) 99.1 (46.01-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 2.20Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.150 , 0.193 0.158 , 0.196	Depositor DCC
R_{free} test set	5027 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	32.2	Xtriage
Anisotropy	0.579	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 59.0	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.019 for -k,-h,-l 0.022 for k,h,-l 0.035 for h,-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	32784	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% 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: IOD, TB, 7MT

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	0.55	0/1344	0.58	0/1816
1	B	0.53	0/1344	0.56	0/1816
1	C	0.51	0/1344	0.54	0/1816
1	D	0.50	0/1344	0.55	0/1816
1	E	0.52	0/1344	0.59	1/1816 (0.1%)
1	F	0.49	0/1344	0.55	0/1816
1	G	0.52	0/1344	0.56	0/1816
1	H	0.45	0/1344	0.53	0/1816
1	I	0.46	0/1344	0.54	0/1816
1	J	0.47	0/1344	0.54	0/1816
1	K	0.50	0/1344	0.58	1/1816 (0.1%)
1	L	0.51	0/1344	0.58	0/1816
All	All	0.50	0/16128	0.56	2/21792 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	77	ARG	CB-CG-CD	-6.62	94.38	111.60
1	K	75	ARG	CG-CD-NE	5.38	123.11	111.80

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	1314	1315	1319	3	0
1	B	1314	1315	1319	5	0
1	C	1314	1315	1319	3	0
1	D	1314	1315	1319	6	0
1	E	1314	1314	1319	5	0
1	F	1314	1315	1319	4	0
1	G	1314	1315	1319	5	0
1	H	1314	1315	1319	7	0
1	I	1314	1315	1319	5	0
1	J	1314	1315	1319	6	0
1	K	1314	1315	1319	8	0
1	L	1314	1315	1319	3	0
2	A	30	23	0	1	0
2	B	30	23	0	1	0
2	E	30	23	0	2	0
2	F	30	23	0	1	0
2	G	30	23	0	2	0
2	I	30	23	0	1	0
2	J	30	23	0	1	0
2	K	30	23	0	1	0
2	L	30	23	0	1	0
3	A	1	0	0	0	0
3	B	1	0	0	2	0
3	C	1	0	0	1	0
3	E	1	0	0	1	0
3	F	1	0	0	0	0
3	G	1	0	0	1	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	2	0	0	0	0
4	F	2	0	0	0	0
4	H	1	0	0	0	0
4	J	1	0	0	0	0
4	K	1	0	0	0	0
5	A	73	0	0	0	0
5	B	73	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	76	0	0	0	0
5	D	75	0	0	0	0
5	E	76	0	0	0	0
5	F	58	0	0	0	0
5	G	82	0	0	0	0
5	H	36	0	0	0	0
5	I	36	0	0	0	0
5	J	50	0	0	0	0
5	K	62	0	0	0	0
5	L	43	0	0	0	0
All	All	16798	15986	15828	55	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

The worst 5 of 55 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:E:8:ALA:HB1	2:E:203:7MT:O25	1.77	0.85
1:G:8:ALA:HB1	2:G:201:7MT:O25	1.92	0.70
1:K:135:VAL:HG23	1:K:140:VAL:HG21	1.74	0.69
1:F:8:ALA:HB1	2:F:203:7MT:O25	1.96	0.65
1:E:77:ARG:HD2	1:E:108:SER:OG	1.96	0.64

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	164/166 (99%)	161 (98%)	3 (2%)	0	100	100
1	B	164/166 (99%)	162 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	D	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	E	164/166 (99%)	162 (99%)	2 (1%)	0	100	100
1	F	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	G	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	H	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	I	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	J	164/166 (99%)	163 (99%)	1 (1%)	0	100	100
1	K	164/166 (99%)	162 (99%)	2 (1%)	0	100	100
1	L	164/166 (99%)	162 (99%)	2 (1%)	0	100	100
All	All	1968/1992 (99%)	1950 (99%)	18 (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	140/140 (100%)	137 (98%)	3 (2%)	53	67
1	B	140/140 (100%)	137 (98%)	3 (2%)	53	67
1	C	140/140 (100%)	136 (97%)	4 (3%)	42	54
1	D	140/140 (100%)	139 (99%)	1 (1%)	84	91
1	E	140/140 (100%)	138 (99%)	2 (1%)	67	80
1	F	140/140 (100%)	138 (99%)	2 (1%)	67	80
1	G	140/140 (100%)	136 (97%)	4 (3%)	42	54
1	H	140/140 (100%)	138 (99%)	2 (1%)	67	80
1	I	140/140 (100%)	132 (94%)	8 (6%)	20	24
1	J	140/140 (100%)	135 (96%)	5 (4%)	35	45
1	K	140/140 (100%)	135 (96%)	5 (4%)	35	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	140/140 (100%)	134 (96%)	6 (4%)	29	36
All	All	1680/1680 (100%)	1635 (97%)	45 (3%)	44	57

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	52	LEU
1	I	91	SER
1	L	60	GLU
1	I	12	GLU
1	I	106	LEU

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

5.3.3   RNA (i)

There are no RNA molecules in this entry.

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

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

5.5   Carbohydrates (i)

There are no carbohydrates in this entry.

5.6   Ligand geometry (i)

Of 29 ligands modelled in this entry, 20 are monoatomic - leaving 9 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	7MT	F	203	-	21,38,38	1.47	3 (14%)	20,76,76	2.18	5 (25%)
2	7MT	G	201	-	21,38,38	1.34	3 (14%)	20,76,76	2.01	5 (25%)
2	7MT	L	201	-	21,38,38	1.32	2 (9%)	20,76,76	2.08	5 (25%)
2	7MT	J	202	-	21,38,38	1.31	2 (9%)	20,76,76	1.51	3 (15%)
2	7MT	B	201	-	21,38,38	1.44	3 (14%)	20,76,76	2.24	7 (35%)
2	7MT	E	203	-	21,38,38	1.26	2 (9%)	20,76,76	1.88	5 (25%)
2	7MT	I	201	-	21,38,38	1.61	3 (14%)	20,76,76	1.95	5 (25%)
2	7MT	A	201	-	21,38,38	1.26	2 (9%)	20,76,76	2.07	5 (25%)
2	7MT	K	202	-	21,38,38	1.45	2 (9%)	20,76,76	2.75	8 (40%)

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	201	7MT	C30-C10	4.73	1.56	1.51
2	K	202	7MT	C30-C10	4.40	1.56	1.51
2	F	203	7MT	C30-C10	4.34	1.56	1.51
2	B	201	7MT	C30-C10	4.03	1.55	1.51
2	E	203	7MT	C30-C10	3.68	1.55	1.51

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	202	7MT	O26-C24-C22	8.20	127.75	115.78
2	F	203	7MT	O26-C24-C22	6.55	125.33	115.78
2	B	201	7MT	O26-C24-C22	6.02	124.56	115.78
2	I	201	7MT	O26-C24-C22	5.47	123.76	115.78
2	G	201	7MT	O26-C24-C22	5.29	123.50	115.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	203	7MT	1	0
2	G	201	7MT	2	0
2	L	201	7MT	1	0
2	J	202	7MT	1	0

Continued from previous page...

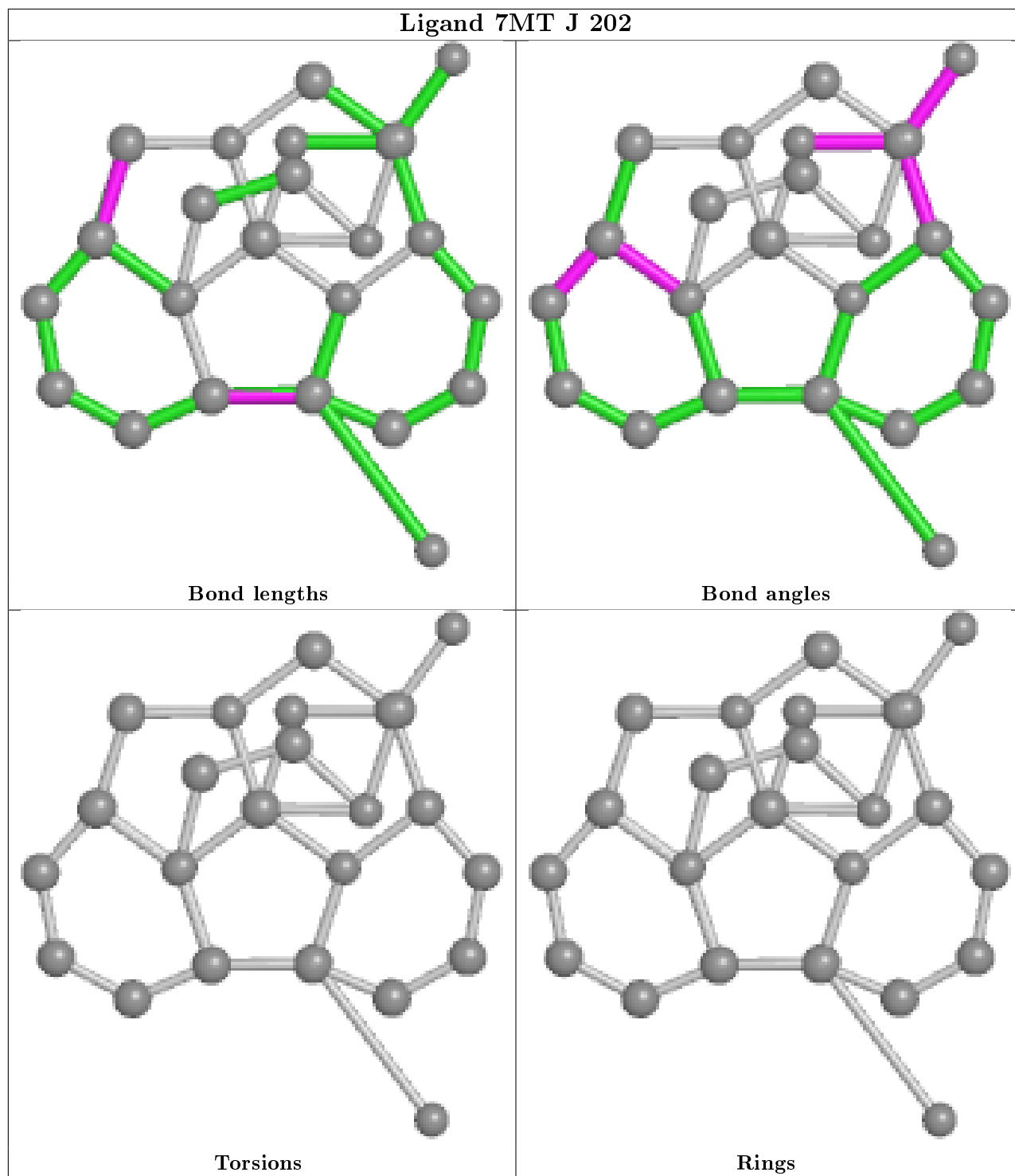| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	201	7MT	1	0
2	E	203	7MT	2	0
2	I	201	7MT	1	0
2	A	201	7MT	1	0
2	K	202	7MT	1	0

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.
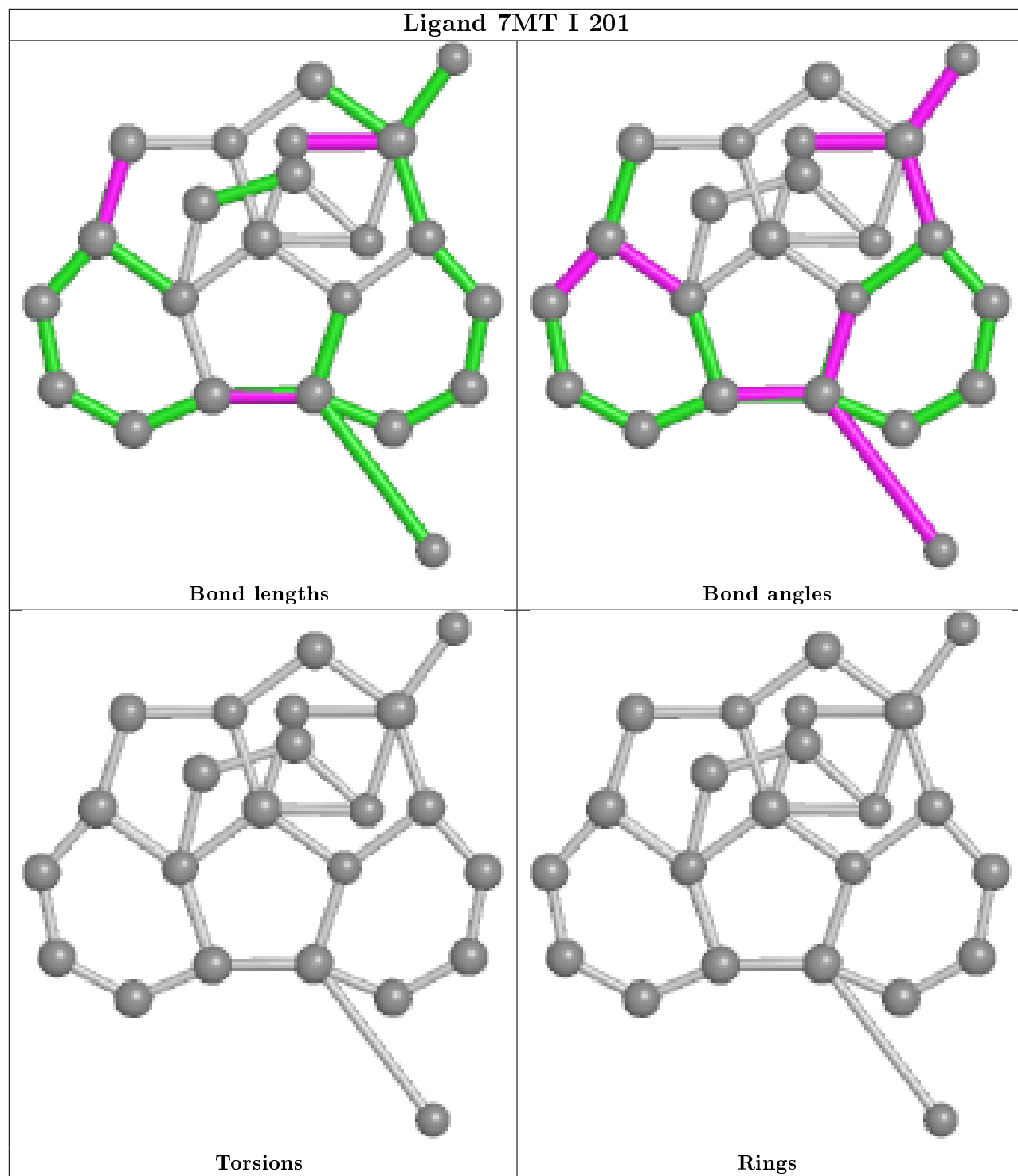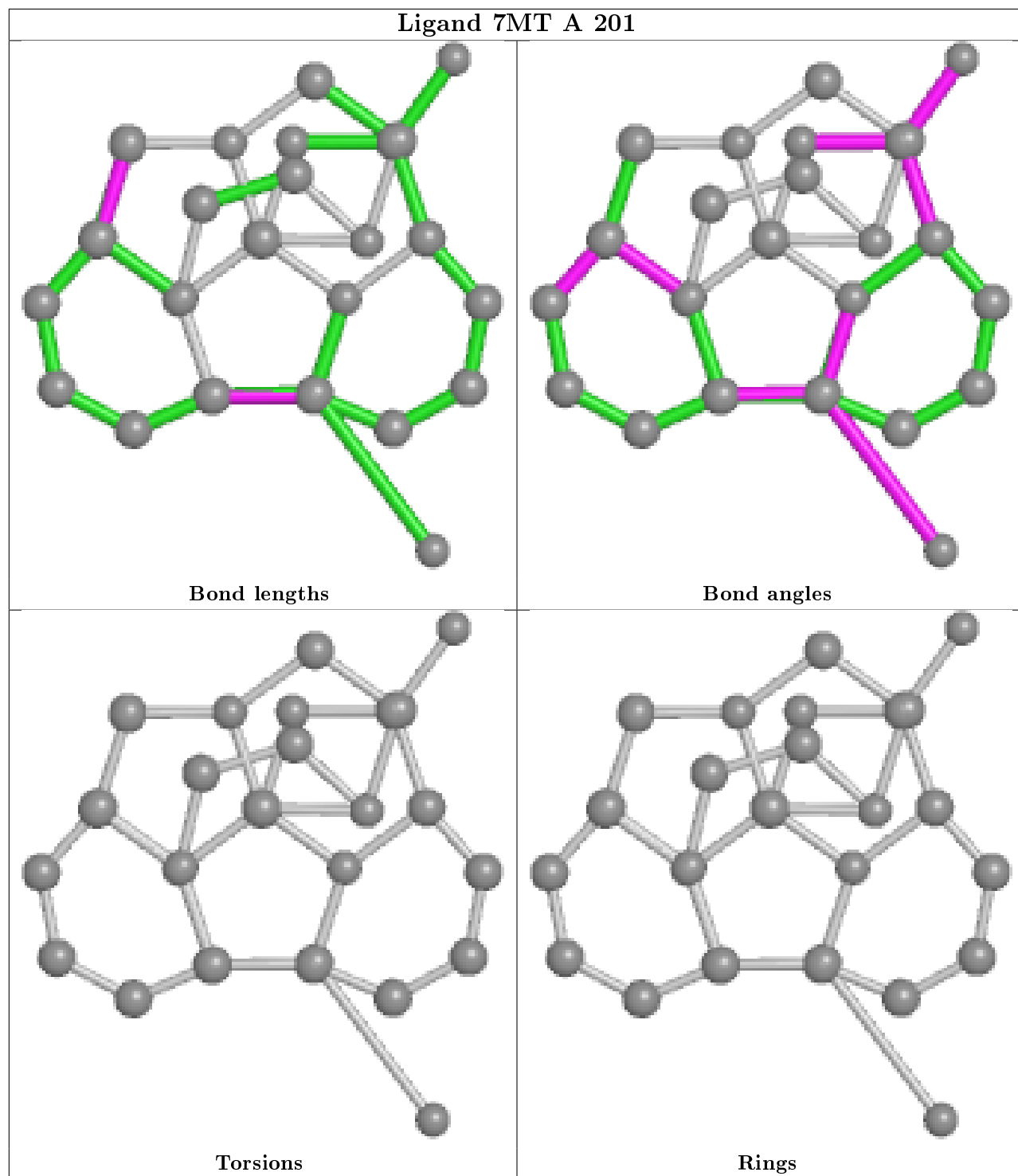
Ligand 7MT F 203

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT G 201

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT L 201

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT J 202

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT B 201

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT E 203

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT I 201

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT A 201

Bond lengths

Bond angles

Torsions

Rings

Ligand 7MT K 202

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

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	166/166 (100%)	-0.40	0	100	100	23, 33, 55, 84	0
1	B	166/166 (100%)	-0.37	1 (0%)	89	88	22, 32, 55, 77	0
1	C	166/166 (100%)	-0.35	0	100	100	24, 34, 57, 91	0
1	D	166/166 (100%)	-0.41	1 (0%)	89	88	25, 34, 59, 73	0
1	E	166/166 (100%)	-0.38	1 (0%)	89	88	24, 34, 53, 80	0
1	F	166/166 (100%)	-0.35	0	100	100	25, 37, 66, 87	0
1	G	166/166 (100%)	-0.32	1 (0%)	89	88	24, 33, 60, 86	0
1	H	166/166 (100%)	-0.26	1 (0%)	89	88	29, 43, 69, 94	0
1	I	166/166 (100%)	-0.18	1 (0%)	89	88	35, 45, 68, 85	0
1	J	166/166 (100%)	-0.31	2 (1%)	79	77	29, 39, 60, 93	0
1	K	166/166 (100%)	-0.33	1 (0%)	89	88	26, 35, 64, 104	0
1	L	166/166 (100%)	-0.24	1 (0%)	89	88	29, 41, 70, 94	0
All	All	1992/1992 (100%)	-0.33	10 (0%)	91	90	22, 37, 64, 104	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	71	ARG	2.9
1	D	166	LYS	2.8
1	L	166	LYS	2.7
1	E	71	ARG	2.6
1	I	71	ARG	2.6

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

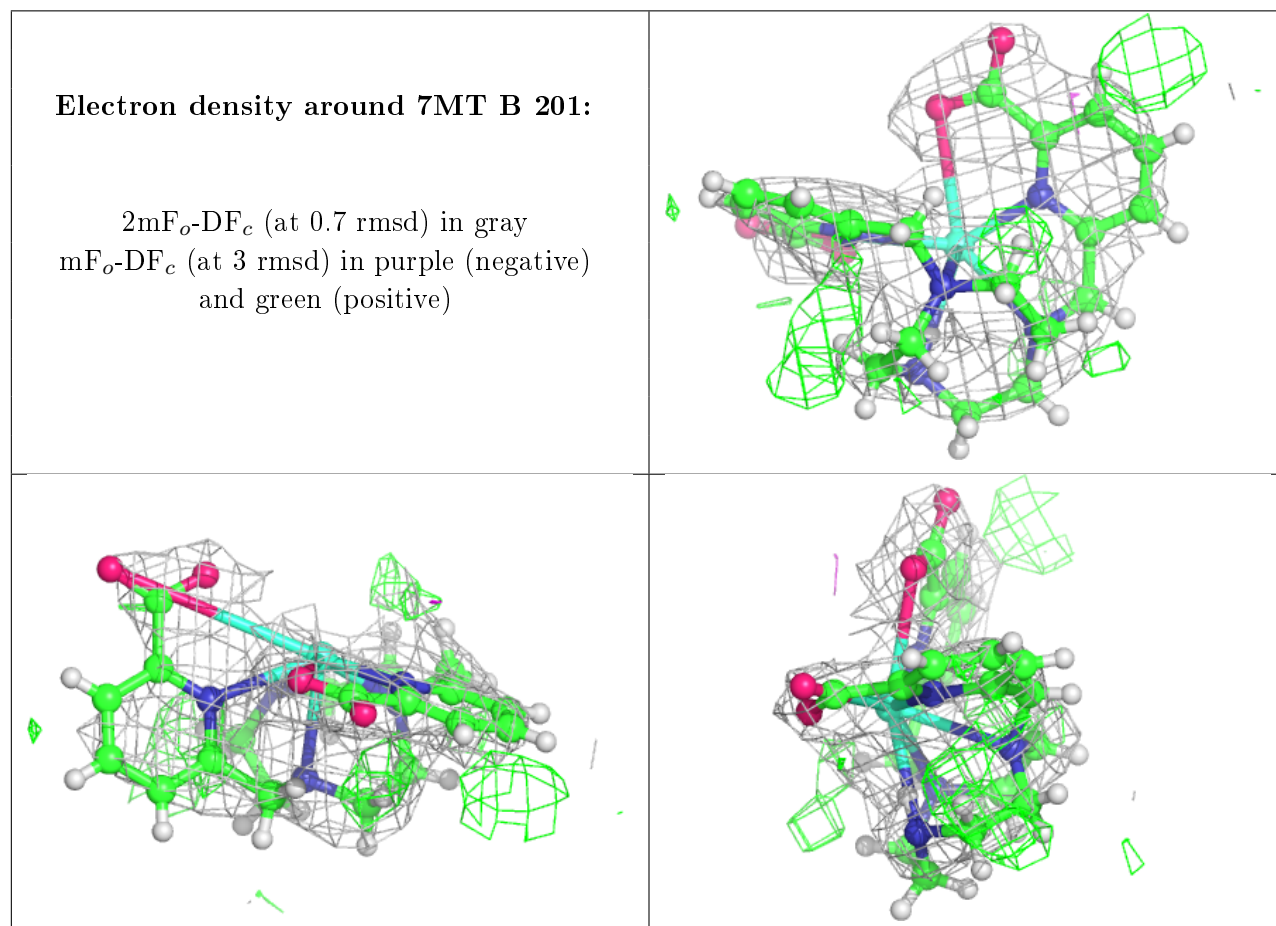## 6.3 Carbohydrates (i)

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(Å^2)	Q<0.9
3	IOD	C	202	1/1	0.41	0.17	129,129,129,129	1
3	IOD	B	202	1/1	0.55	0.14	141,141,141,141	1
3	IOD	L	202	1/1	0.57	0.10	186,186,186,186	1
3	IOD	I	202	1/1	0.76	0.17	139,139,139,139	1
4	TB	E	202	1/1	0.83	0.40	42,42,42,42	1
3	IOD	K	203	1/1	0.87	0.12	132,132,132,132	1
3	IOD	G	202	1/1	0.89	0.08	117,117,117,117	1
3	IOD	A	202	1/1	0.91	0.06	133,133,133,133	1
4	TB	F	202	1/1	0.93	0.13	60,60,60,60	1
3	IOD	J	203	1/1	0.93	0.05	116,116,116,116	1
2	7MT	B	201	30/30	0.94	0.15	25,43,49,83	53
3	IOD	H	202	1/1	0.94	0.07	109,109,109,109	1
3	IOD	F	204	1/1	0.94	0.06	134,134,134,134	1
4	TB	F	201	1/1	0.94	0.52	46,46,46,46	1
4	TB	C	201	1/1	0.95	0.08	70,70,70,70	1
2	7MT	I	201	30/30	0.96	0.13	37,43,54,72	53
2	7MT	E	203	30/30	0.96	0.13	20,33,46,62	53
2	7MT	L	201	30/30	0.96	0.12	32,45,55,80	53
4	TB	D	201	1/1	0.96	0.18	66,66,66,66	1
2	7MT	K	202	30/30	0.96	0.14	28,42,46,73	53
3	IOD	E	204	1/1	0.97	0.07	114,114,114,114	1
4	TB	E	201	1/1	0.98	0.23	30,30,30,30	1
2	7MT	J	202	30/30	0.98	0.13	21,34,42,56	53
2	7MT	G	201	30/30	0.98	0.16	21,31,36,56	53
2	7MT	A	201	30/30	0.98	0.12	17,35,41,61	53
2	7MT	F	203	30/30	0.98	0.11	30,40,45,66	53
4	TB	H	201	1/1	0.99	0.13	75,75,75,75	1
4	TB	J	201	1/1	0.99	0.20	61,61,61,61	1
4	TB	K	201	1/1	0.99	0.06	82,82,82,82	1

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.
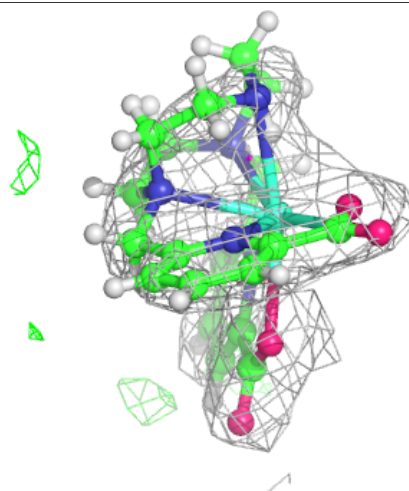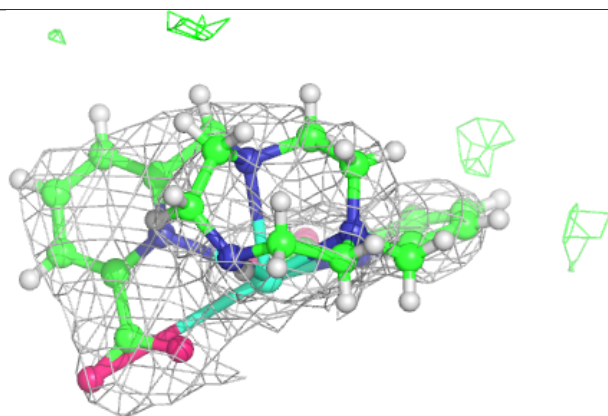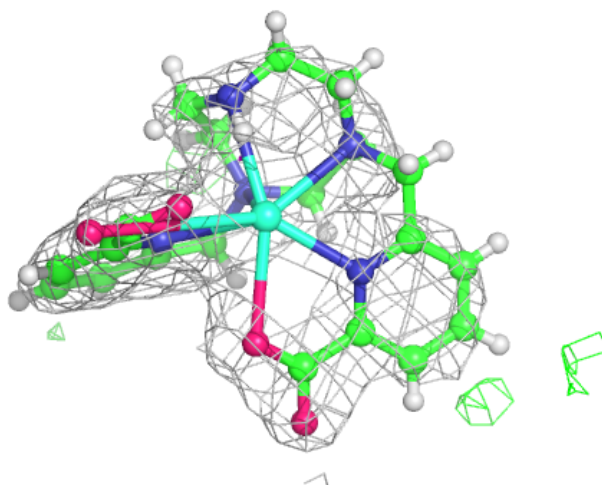
Electron density around 7MT B 201:

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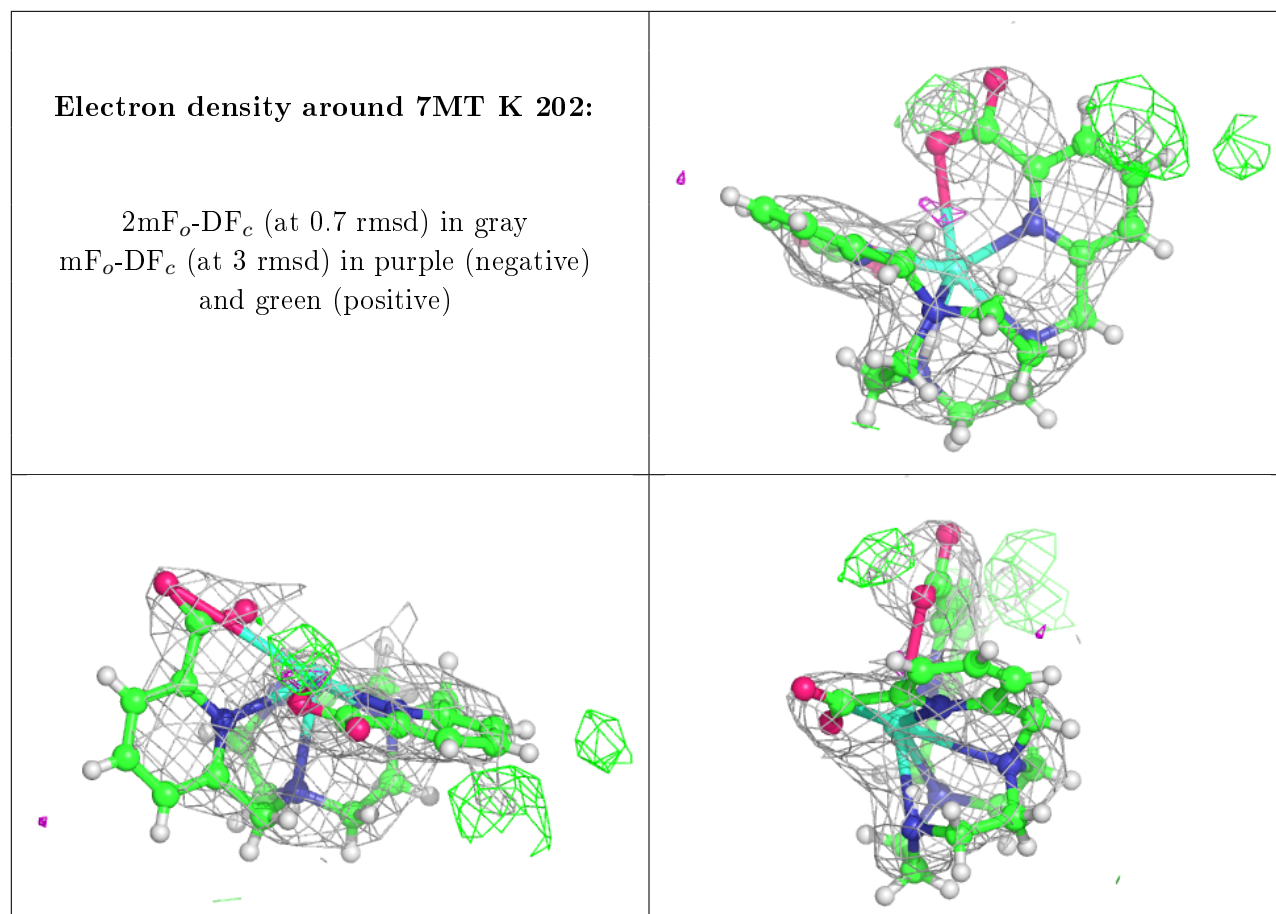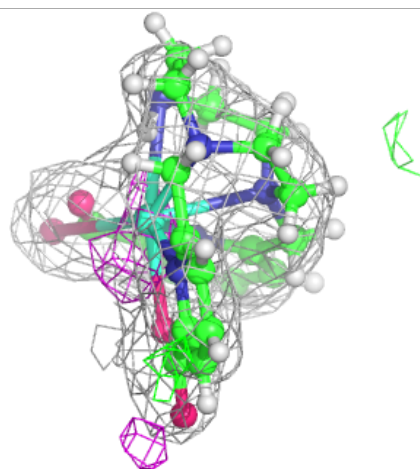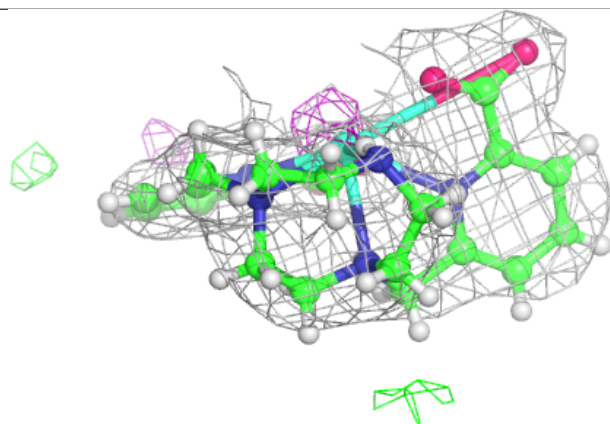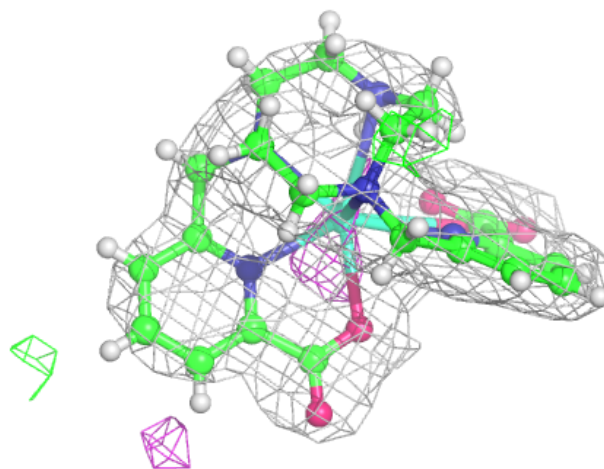$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around 7MT I 201:

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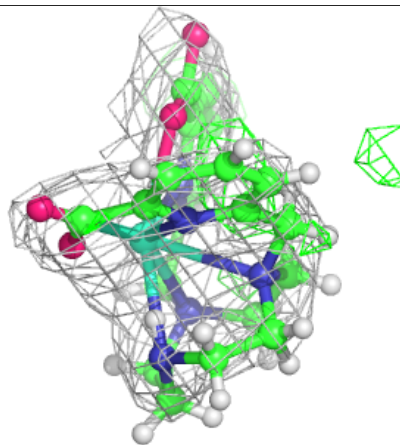$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
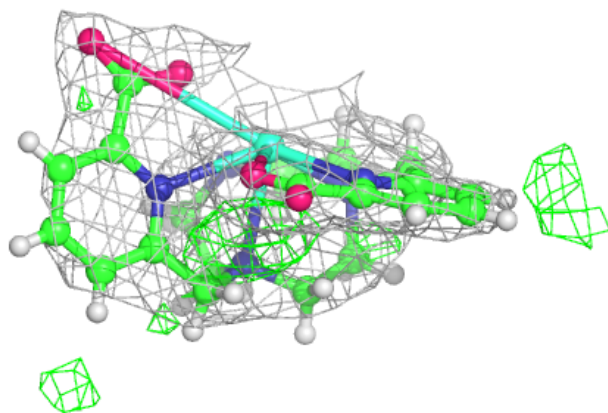and green (positive)

Electron density around 7MT E 203:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
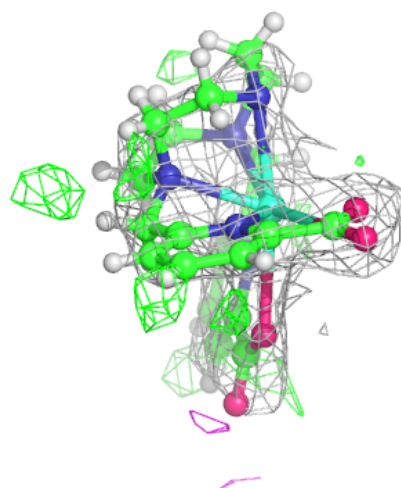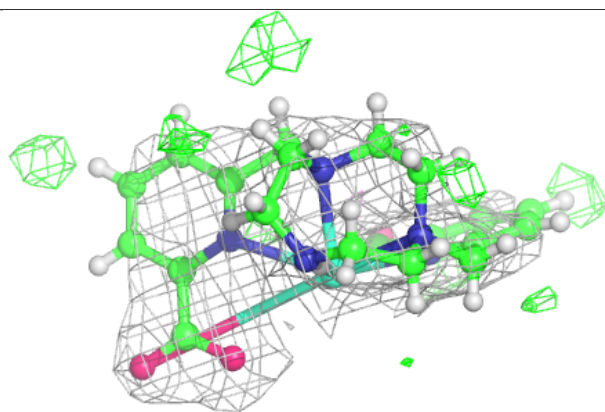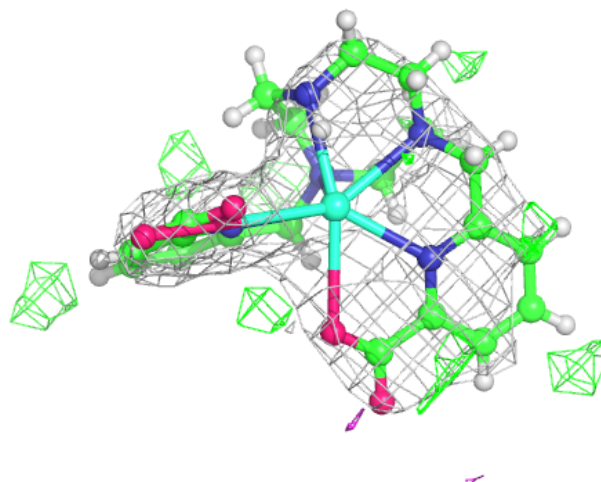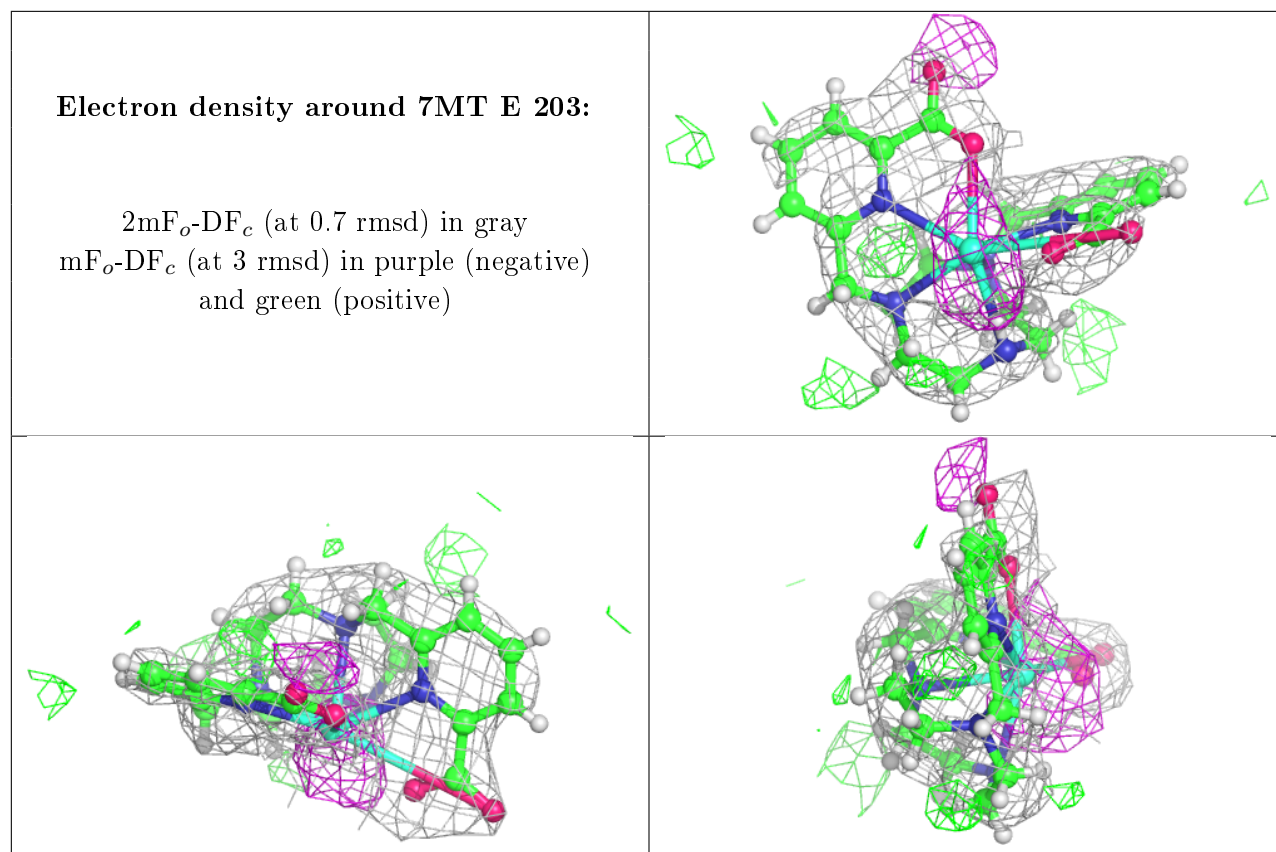
Electron density around 7MT L 201:

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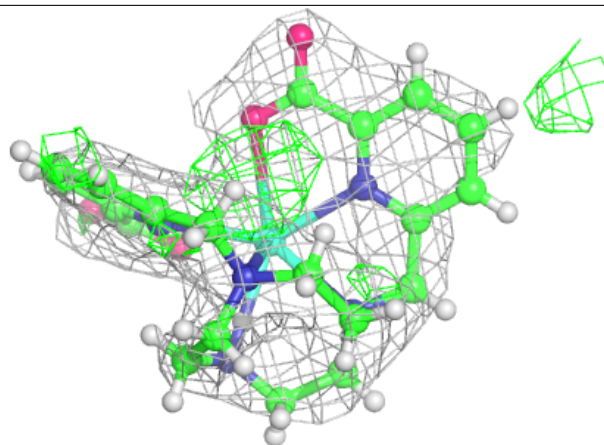and green (positive)

Electron density around 7MT K 202:

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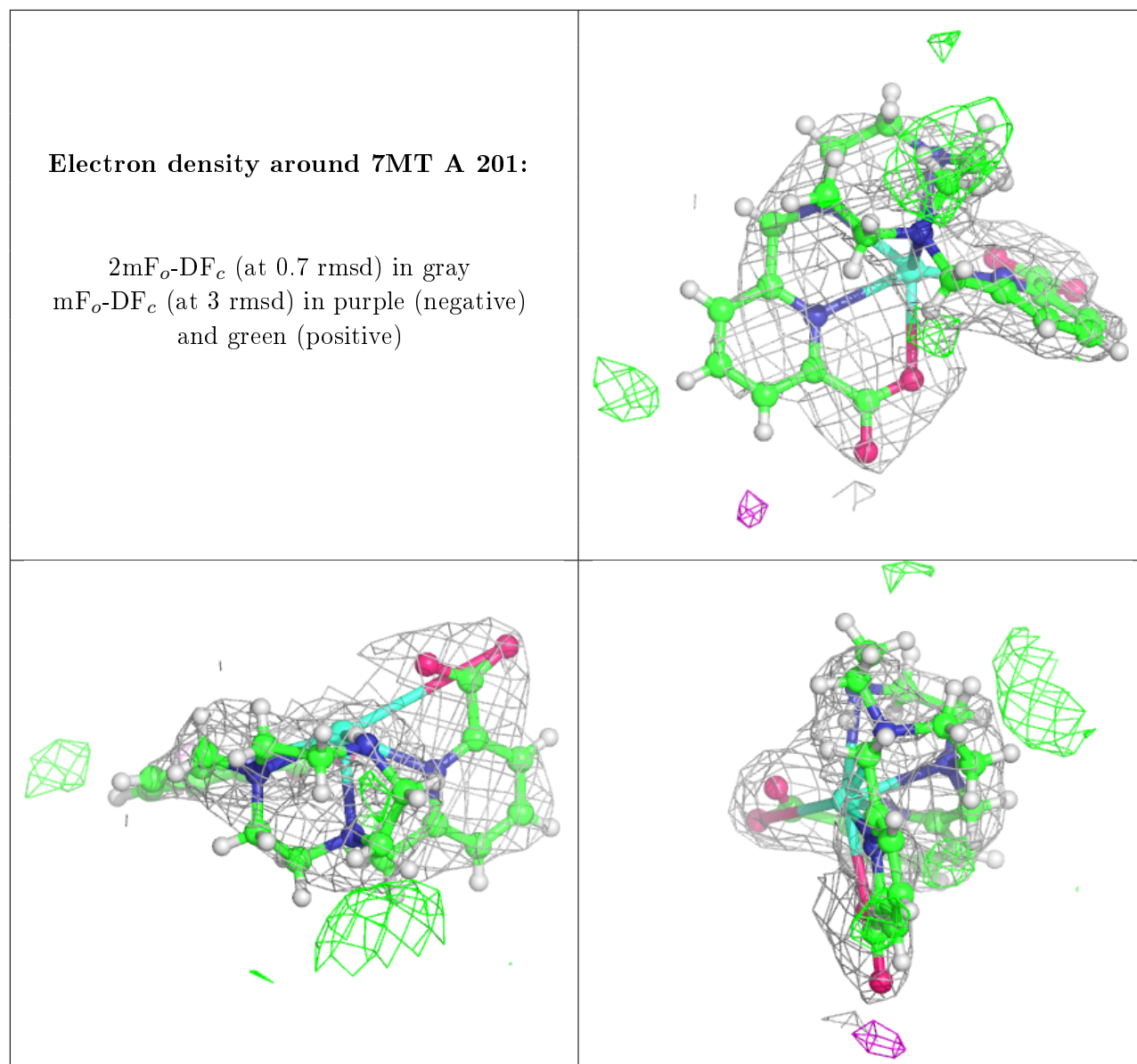and green (positive)

Electron density around 7MT J 202:

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

Electron density around 7MT G 201:

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

Electron density around 7MT A 201:

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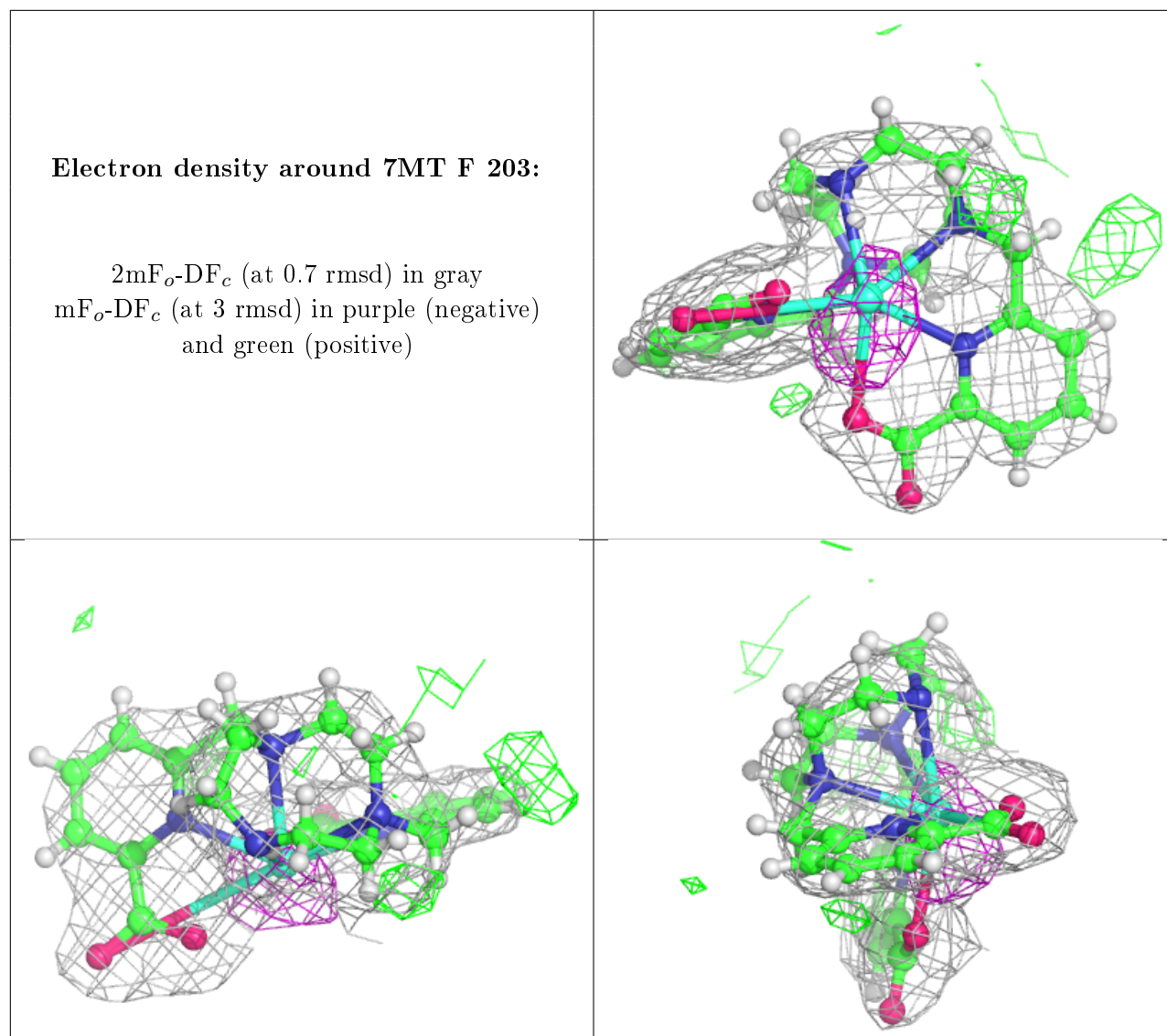$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around 7MT F 203:

$2mF_o$-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 ⓘ

There are no such residues in this entry.