wwPDB X-ray Structure Validation Summary Report (i)

Nov 22, 2023 – 08:10 PM JST

PDB ID	:	7XD9
Title	:	Crystal Structure of Dengue Virus serotype 2 (DENV2) Polymerase Elongation Complex (CTP Form)
Authors	:	Wu, J.; Wang, X.; Gong, P.
Deposited on	:	2022-03-26
Resolution	:	2.58 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

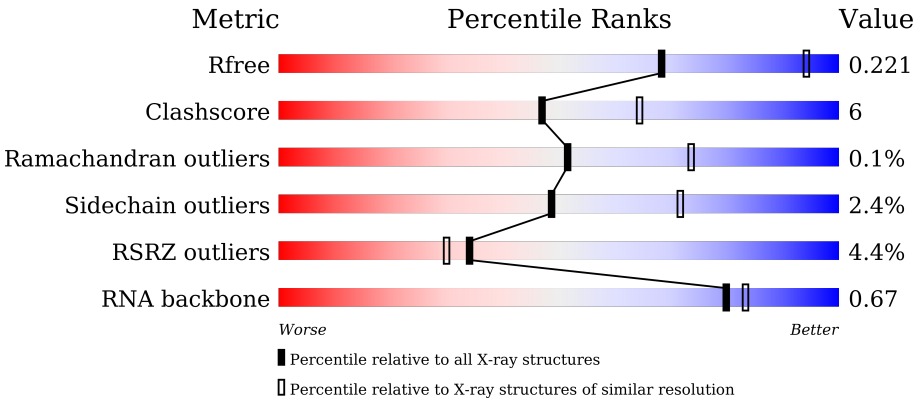
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.58 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.



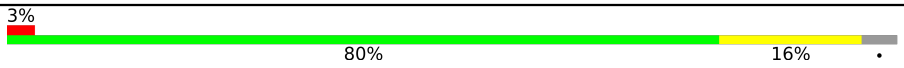
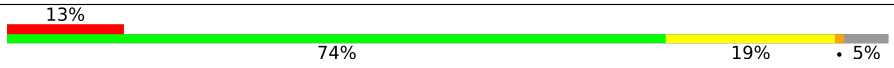
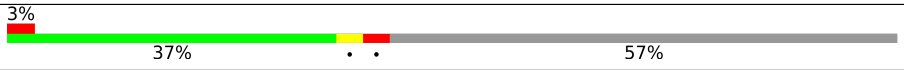
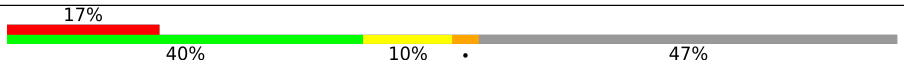
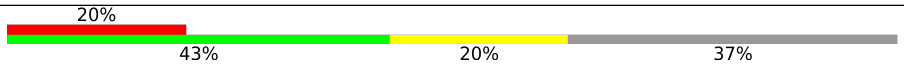
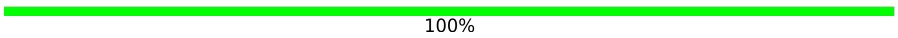
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)
RNA backbone	3102	1075 (2.90-2.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	647	2% 83% 13% ·
1	D	647	% 82% 13% ·
1	G	647	3% 83% 13% ·
1	J	647	3% 79% 17% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	647	
1	P	647	
2	B	30	
2	E	30	
2	H	30	
2	K	30	
2	N	30	
2	Q	30	
3	C	9	
3	F	9	
3	I	9	
3	L	9	
3	O	9	
3	R	9	

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
6	GOL	H	1101	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 32878 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 NS5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	621	Total	C	N	O	S	0	0	0
			4913	3095	878	907	33			
1	D	622	Total	C	N	O	S	0	0	0
			4973	3131	886	922	34			
1	G	623	Total	C	N	O	S	0	0	0
			4913	3095	880	904	34			
1	J	625	Total	C	N	O	S	0	0	0
			4917	3100	872	911	34			
1	M	624	Total	C	N	O	S	0	0	0
			4929	3101	879	915	34			
1	P	616	Total	C	N	O	S	0	0	0
			4537	2851	818	834	34			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	901	GLY	-	expression tag	UNP Q91H74
A	902	SER	-	expression tag	UNP Q91H74
A	903	SER	-	expression tag	UNP Q91H74
A	904	SER	-	expression tag	UNP Q91H74
A	905	HIS	-	expression tag	UNP Q91H74
A	906	HIS	-	expression tag	UNP Q91H74
A	907	HIS	-	expression tag	UNP Q91H74
A	908	HIS	-	expression tag	UNP Q91H74
A	909	HIS	-	expression tag	UNP Q91H74
A	910	HIS	-	expression tag	UNP Q91H74
D	901	GLY	-	expression tag	UNP Q91H74
D	902	SER	-	expression tag	UNP Q91H74
D	903	SER	-	expression tag	UNP Q91H74
D	904	SER	-	expression tag	UNP Q91H74
D	905	HIS	-	expression tag	UNP Q91H74
D	906	HIS	-	expression tag	UNP Q91H74
D	907	HIS	-	expression tag	UNP Q91H74

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	908	HIS	-	expression tag	UNP Q91H74
D	909	HIS	-	expression tag	UNP Q91H74
D	910	HIS	-	expression tag	UNP Q91H74
G	901	GLY	-	expression tag	UNP Q91H74
G	902	SER	-	expression tag	UNP Q91H74
G	903	SER	-	expression tag	UNP Q91H74
G	904	SER	-	expression tag	UNP Q91H74
G	905	HIS	-	expression tag	UNP Q91H74
G	906	HIS	-	expression tag	UNP Q91H74
G	907	HIS	-	expression tag	UNP Q91H74
G	908	HIS	-	expression tag	UNP Q91H74
G	909	HIS	-	expression tag	UNP Q91H74
G	910	HIS	-	expression tag	UNP Q91H74
J	901	GLY	-	expression tag	UNP Q91H74
J	902	SER	-	expression tag	UNP Q91H74
J	903	SER	-	expression tag	UNP Q91H74
J	904	SER	-	expression tag	UNP Q91H74
J	905	HIS	-	expression tag	UNP Q91H74
J	906	HIS	-	expression tag	UNP Q91H74
J	907	HIS	-	expression tag	UNP Q91H74
J	908	HIS	-	expression tag	UNP Q91H74
J	909	HIS	-	expression tag	UNP Q91H74
J	910	HIS	-	expression tag	UNP Q91H74
M	901	GLY	-	expression tag	UNP Q91H74
M	902	SER	-	expression tag	UNP Q91H74
M	903	SER	-	expression tag	UNP Q91H74
M	904	SER	-	expression tag	UNP Q91H74
M	905	HIS	-	expression tag	UNP Q91H74
M	906	HIS	-	expression tag	UNP Q91H74
M	907	HIS	-	expression tag	UNP Q91H74
M	908	HIS	-	expression tag	UNP Q91H74
M	909	HIS	-	expression tag	UNP Q91H74
M	910	HIS	-	expression tag	UNP Q91H74
P	901	GLY	-	expression tag	UNP Q91H74
P	902	SER	-	expression tag	UNP Q91H74
P	903	SER	-	expression tag	UNP Q91H74
P	904	SER	-	expression tag	UNP Q91H74
P	905	HIS	-	expression tag	UNP Q91H74
P	906	HIS	-	expression tag	UNP Q91H74
P	907	HIS	-	expression tag	UNP Q91H74
P	908	HIS	-	expression tag	UNP Q91H74
P	909	HIS	-	expression tag	UNP Q91H74

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	910	HIS	-	expression tag	UNP Q91H74

- Molecule 2 is a RNA chain called RNA (30-mer).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	12	Total	C	N	O	P	0	0	0
			251	113	42	84	12			
2	E	13	Total	C	N	O	P	0	0	0
			273	123	47	90	13			
2	H	12	Total	C	N	O	P	0	0	0
			251	113	42	84	12			
2	K	16	Total	C	N	O	P	0	0	0
			336	152	60	109	15			
2	N	19	Total	C	N	O	P	0	0	0
			401	181	73	129	18			
2	Q	12	Total	C	N	O	P	0	0	0
			251	113	42	84	12			

- Molecule 3 is a RNA chain called RNA (9-mer).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	P	0	0	0
			195	87	36	63	9			
3	F	9	Total	C	N	O	P	0	0	0
			195	87	36	63	9			
3	I	9	Total	C	N	O	P	0	0	0
			195	87	36	63	9			
3	L	9	Total	C	N	O	P	0	0	0
			194	87	36	62	9			
3	O	9	Total	C	N	O	P	0	0	0
			195	87	36	63	9			
3	R	9	Total	C	N	O	P	0	0	0
			195	87	36	63	9			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	0
			2	2		
4	D	2	Total	Mg	0	0
			2	2		

Continued on next page...

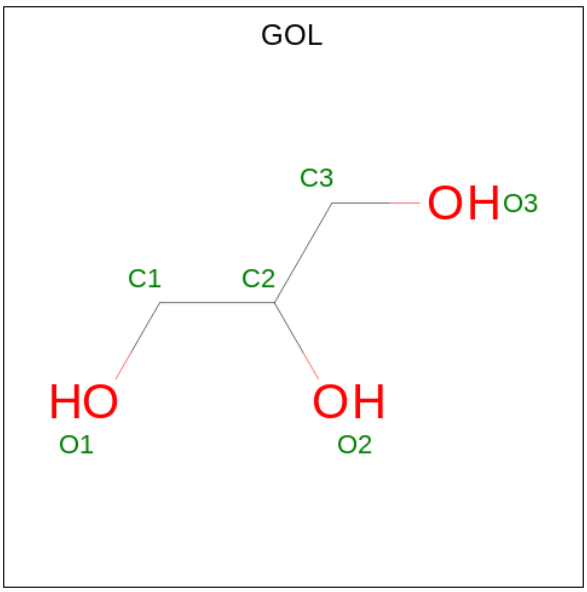
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	2	Total Mg 2 2	0	0
4	J	2	Total Mg 2 2	0	0
4	M	2	Total Mg 2 2	0	0
4	P	2	Total Mg 2 2	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

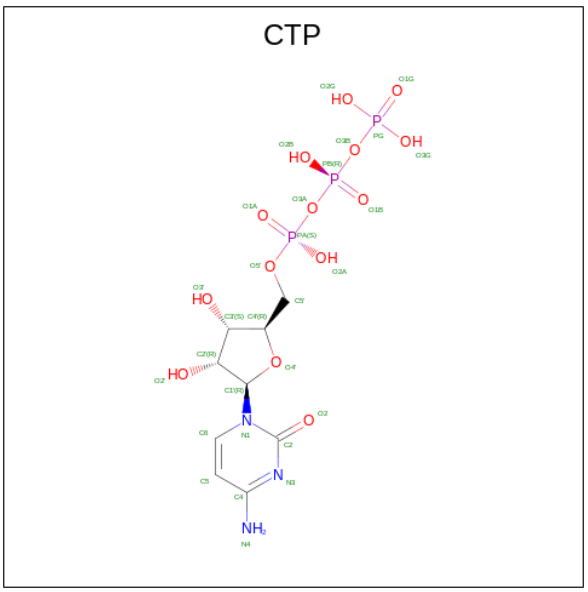
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Zn 2 2	0	0
5	D	2	Total Zn 2 2	0	0
5	G	2	Total Zn 2 2	0	0
5	J	2	Total Zn 2 2	0	0
5	M	2	Total Zn 2 2	0	0
5	P	2	Total Zn 2 2	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	A	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	E	1	Total C O 6 3 3	0	0
6	E	1	Total C O 6 3 3	0	0
6	H	1	Total C O 6 3 3	0	0
6	J	1	Total C O 6 3 3	0	0
6	K	1	Total C O 6 3 3	0	0
6	M	1	Total C O 6 3 3	0	0

- Molecule 7 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
7	D	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
7	G	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
7	J	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
7	M	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
7	P	1	Total	C	N	O	P	0	0
			29	9	3	14	3		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	98	Total	O	0	0
			98	98		
8	B	6	Total	O	0	0
			6	6		
8	C	3	Total	O	0	0
			3	3		
8	D	111	Total	O	0	0
			111	111		
8	E	8	Total	O	0	0
			8	8		
8	F	5	Total	O	0	0
			5	5		
8	G	74	Total	O	0	0
			74	74		
8	H	2	Total	O	0	0
			2	2		
8	I	2	Total	O	0	0
			2	2		
8	J	68	Total	O	0	0
			68	68		
8	K	2	Total	O	0	0
			2	2		
8	L	1	Total	O	0	0
			1	1		
8	M	84	Total	O	0	0
			84	84		
8	N	7	Total	O	0	0
			7	7		

Continued on next page...

Continued from previous page...

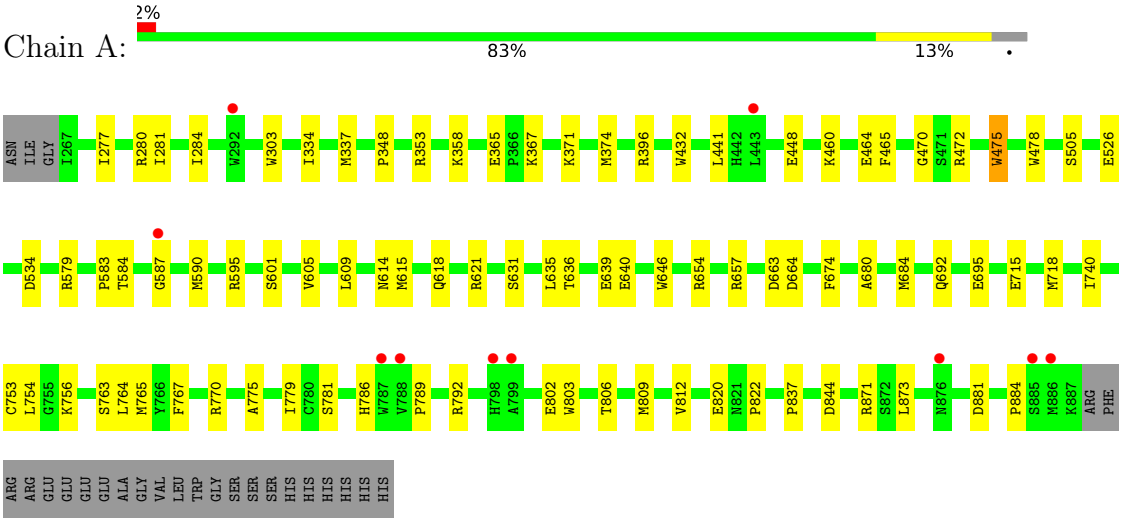
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	O	4	Total O 4 4	0	0
8	P	18	Total O 18 18	0	0
8	R	1	Total O 1 1	0	0

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.

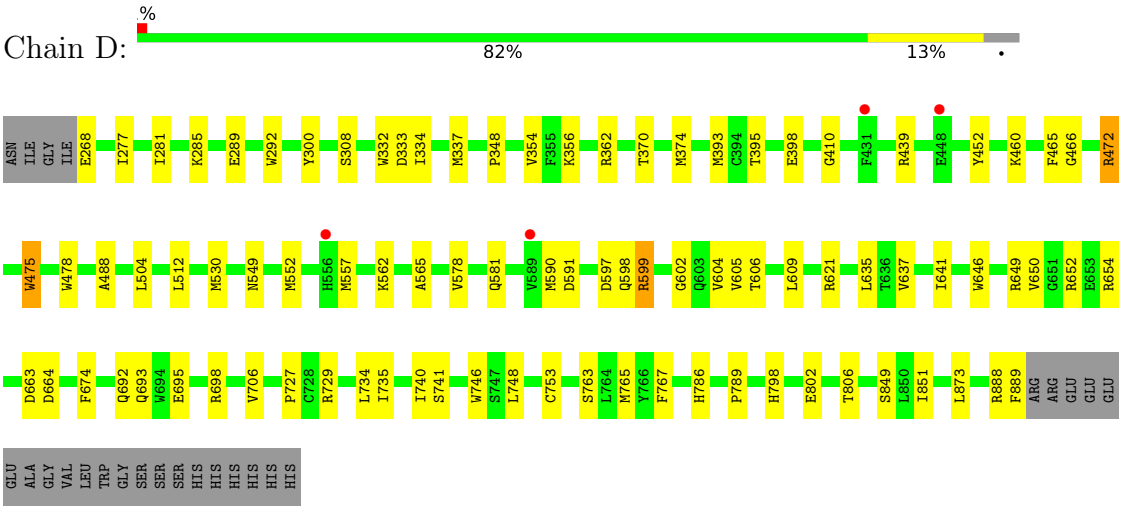
• Molecule 1: NS5

Chain A:



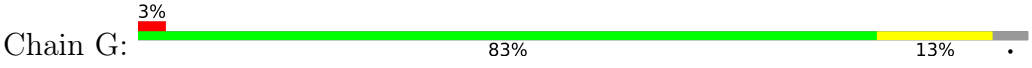
• Molecule 1: NS5

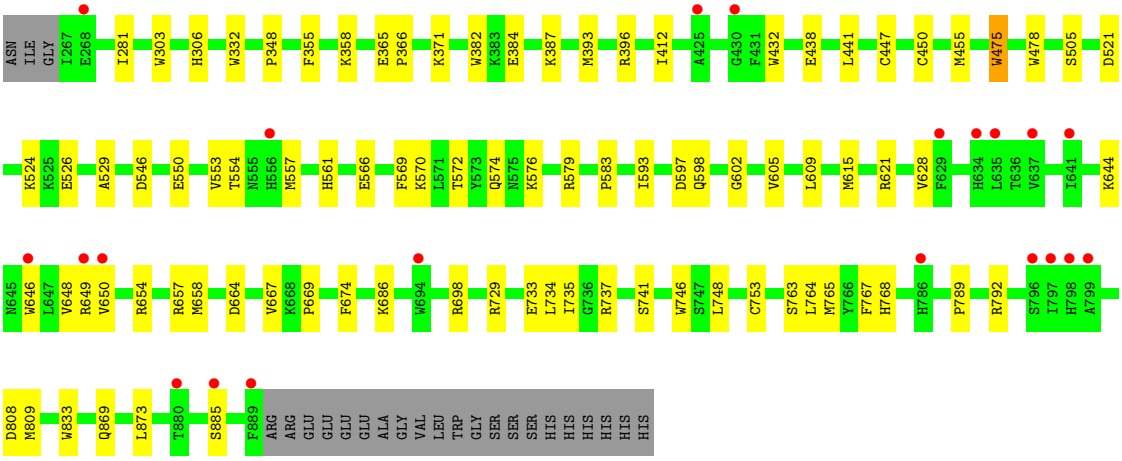
Chain D:



• Molecule 1: NS5

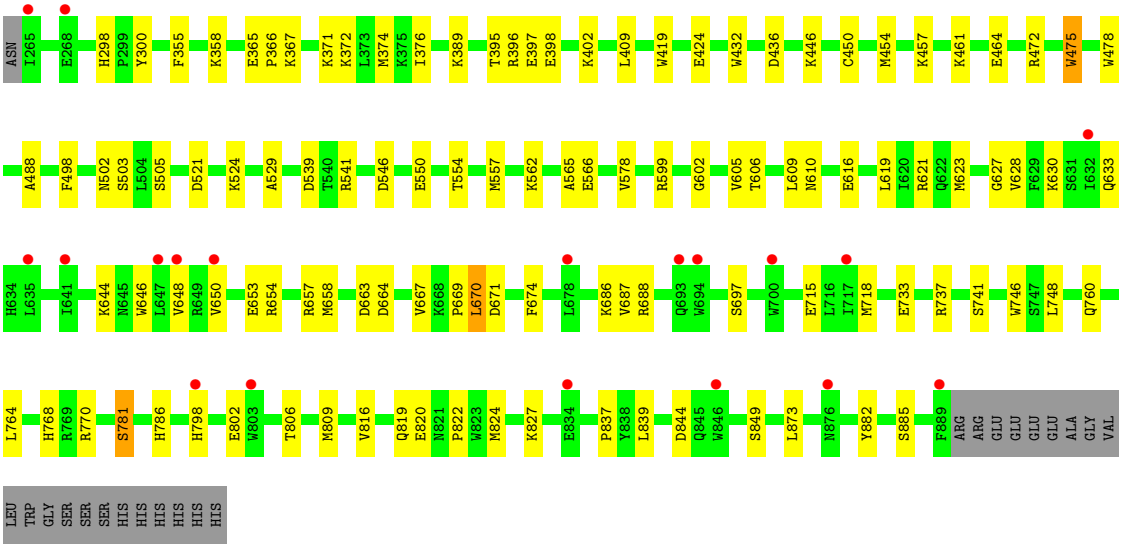
Chain G:





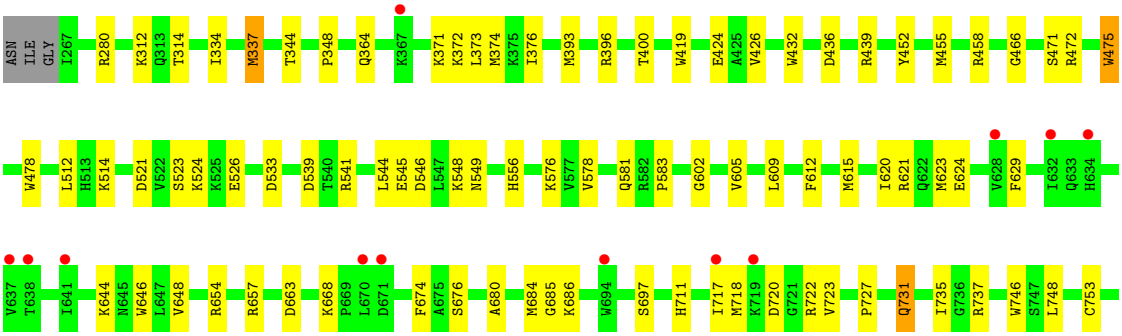
• Molecule 1: NS5

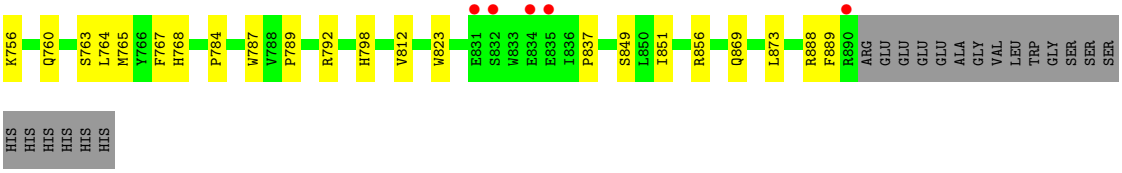
Chain J:



• Molecule 1: NS5

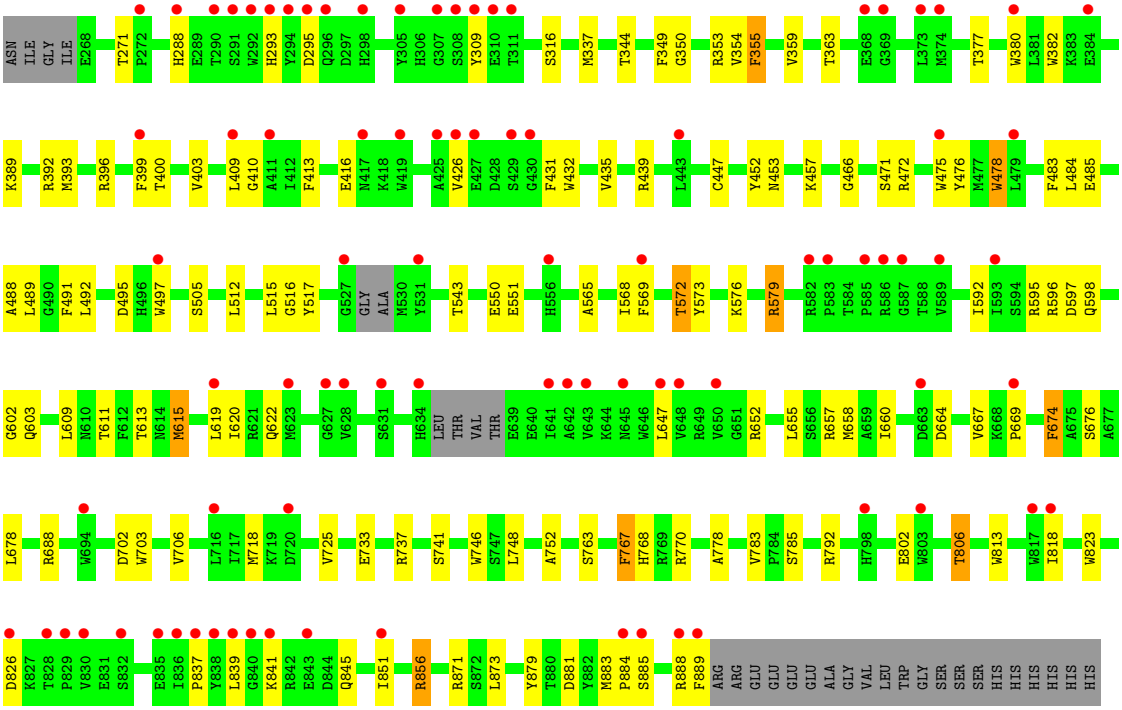
Chain M:





• Molecule 1: NS5

Chain P:



• Molecule 2: RNA (30-mer)

Chain B:

• Molecule 2: RNA (30-mer)

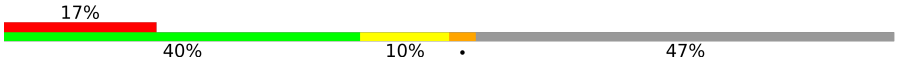
Chain E:

• Molecule 2: RNA (30-mer)

Chain H:

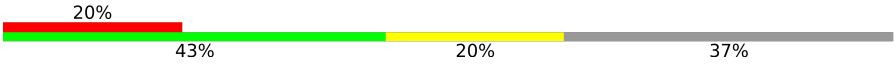
• Molecule 2: RNA (30-mer)

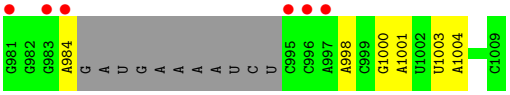
Chain K:



• Molecule 2: RNA (30-mer)

Chain N:





• Molecule 2: RNA (30-mer)

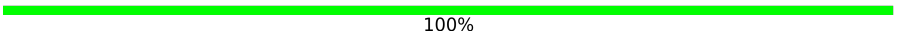
Chain Q:

• Molecule 3: RNA (9-mer)

Chain C:

• Molecule 3: RNA (9-mer)

Chain F:



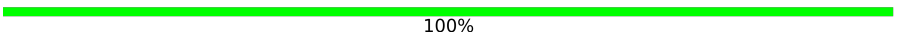
There are no outlier residues recorded for this chain.

• Molecule 3: RNA (9-mer)

Chain I:

• Molecule 3: RNA (9-mer)

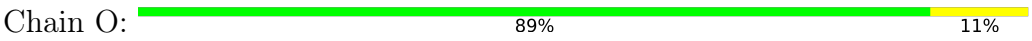
Chain L:



There are no outlier residues recorded for this chain.

- Molecule 3: RNA (9-mer)

Chain O:



- Molecule 3: RNA (9-mer)

Chain R:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	139.43Å 154.58Å 172.35Å 90.00° 94.68° 90.00°	Depositor
Resolution (Å)	48.84 – 2.58 48.84 – 2.57	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.84-2.58) 99.5 (48.84-2.57)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.58Å)	Xtriage
Refinement program	PHENIX 1.19	Depositor
R, R_{free}	0.191 , 0.221 0.192 , 0.221	Depositor DCC
R_{free} test set	11578 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	58.3	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	32878	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% 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: ZN, CTP, GOL, 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	0.43	0/5031	0.65	0/6822
1	D	0.44	0/5093	0.66	0/6898
1	G	0.40	0/5033	0.62	0/6832
1	J	0.43	1/5037 (0.0%)	0.62	0/6836
1	M	0.44	0/5047	0.63	0/6846
1	P	0.35	0/4649	0.57	0/6345
2	B	0.82	0/279	1.45	3/431 (0.7%)
2	E	0.75	0/304	1.25	2/470 (0.4%)
2	H	0.79	1/279 (0.4%)	1.04	0/431
2	K	1.54	6/374 (1.6%)	1.52	8/578 (1.4%)
2	N	0.59	0/447	1.03	0/692
2	Q	0.58	0/279	1.03	0/431
3	C	0.85	0/218	1.37	2/336 (0.6%)
3	F	0.80	0/218	1.13	0/336
3	I	0.85	0/218	1.16	1/336 (0.3%)
3	L	0.73	0/217	1.12	0/334
3	O	0.83	0/218	1.08	0/336
3	R	0.85	0/218	1.31	0/336
All	All	0.48	8/33159 (0.0%)	0.72	16/45626 (0.0%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	996	C	C2-O2	14.76	1.37	1.24
2	K	996	C	C4-N4	13.36	1.46	1.33
2	K	996	C	N1-C6	11.00	1.43	1.37
2	K	996	C	C5-C6	8.89	1.41	1.34
2	K	996	C	N3-C4	8.61	1.40	1.33

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	996	C	C5-C6-N1	13.90	127.95	121.00
2	K	996	C	C4-C5-C6	-12.43	111.18	117.40
2	K	996	C	C6-N1-C2	-9.76	116.40	120.30
2	E	1000	G	N3-C2-N2	7.39	125.08	119.90
2	E	1000	G	N1-C6-O6	-7.25	115.55	119.90

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	4913	0	4699	46	0
1	D	4973	0	4775	57	0
1	G	4913	0	4657	51	0
1	J	4917	0	4659	69	0
1	M	4929	0	4688	67	0
1	P	4537	0	3973	92	0
2	B	251	0	129	2	0
2	E	273	0	140	2	0
2	H	251	0	129	2	0
2	K	336	0	175	3	0
2	N	401	0	208	2	0
2	Q	251	0	129	1	0
3	C	195	0	97	1	0
3	F	195	0	97	0	0
3	I	195	0	97	0	0
3	L	194	0	97	0	0
3	O	195	0	97	1	0
3	R	195	0	97	1	0
4	A	2	0	0	0	0
4	D	2	0	0	0	0
4	G	2	0	0	0	0
4	J	2	0	0	0	0
4	M	2	0	0	0	0
4	P	2	0	0	0	0
5	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	2	0	0	0	0
5	G	2	0	0	0	0
5	J	2	0	0	0	0
5	M	2	0	0	0	0
5	P	2	0	0	0	0
6	A	12	0	16	0	0
6	D	24	0	32	0	0
6	E	12	0	16	1	0
6	H	6	0	8	1	0
6	J	6	0	8	0	0
6	K	6	0	8	0	0
6	M	6	0	8	2	0
7	A	29	0	12	2	0
7	D	29	0	12	1	0
7	G	29	0	12	0	0
7	J	29	0	12	1	0
7	M	29	0	12	0	0
7	P	29	0	12	0	0
8	A	98	0	0	1	0
8	B	6	0	0	0	0
8	C	3	0	0	0	0
8	D	111	0	0	2	0
8	E	8	0	0	0	0
8	F	5	0	0	0	0
8	G	74	0	0	1	0
8	H	2	0	0	0	0
8	I	2	0	0	0	0
8	J	68	0	0	1	0
8	K	2	0	0	0	0
8	L	1	0	0	0	0
8	M	84	0	0	1	0
8	N	7	0	0	0	0
8	O	4	0	0	0	0
8	P	18	0	0	1	0
8	R	1	0	0	0	0
All	All	32878	0	29111	382	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.

The worst 5 of 382 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:J:389:LYS:HE2	1:J:502:ASN:HD22	1.17	1.08
1:J:521:ASP:HA	1:J:524:LYS:HE2	1.59	0.84
1:M:344:THR:HG21	1:M:471:SER:HB2	1.59	0.83
1:P:309:TYR:HE1	1:P:592:ILE:HD11	1.46	0.81
1:P:475:TRP:HZ3	1:P:576:LYS:HD2	1.45	0.79

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	619/647 (96%)	595 (96%)	23 (4%)	1 (0%)	47 69
1	D	620/647 (96%)	598 (96%)	22 (4%)	0	100 100
1	G	621/647 (96%)	603 (97%)	18 (3%)	0	100 100
1	J	623/647 (96%)	589 (94%)	34 (6%)	0	100 100
1	M	622/647 (96%)	600 (96%)	22 (4%)	0	100 100
1	P	610/647 (94%)	565 (93%)	44 (7%)	1 (0%)	47 69
All	All	3715/3882 (96%)	3550 (96%)	163 (4%)	2 (0%)	51 73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	601	SER
1	P	767	PHE

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	507/564 (90%)	500 (99%)	7 (1%)	67	84
1	D	520/564 (92%)	512 (98%)	8 (2%)	65	82
1	G	502/564 (89%)	492 (98%)	10 (2%)	55	76
1	J	504/564 (89%)	491 (97%)	13 (3%)	46	69
1	M	509/564 (90%)	497 (98%)	12 (2%)	49	72
1	P	408/564 (72%)	388 (95%)	20 (5%)	25	46
All	All	2950/3384 (87%)	2880 (98%)	70 (2%)	49	72

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	447	CYS
1	P	572	THR
1	P	676	SER
1	G	885	SER
1	G	869	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	440	ASN
1	P	453	ASN
1	P	622	GLN
1	G	306	HIS
1	D	786	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	11/30 (36%)	0	0
2	E	12/30 (40%)	1 (8%)	0
2	H	11/30 (36%)	0	0
2	K	14/30 (46%)	0	0
2	N	17/30 (56%)	2 (11%)	0
2	Q	11/30 (36%)	0	0
3	C	8/9 (88%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	F	8/9 (88%)	0	0
3	I	8/9 (88%)	1 (12%)	0
3	L	8/9 (88%)	0	0
3	O	8/9 (88%)	0	0
3	R	8/9 (88%)	0	0
All	All	124/234 (52%)	4 (3%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	1000	G
3	I	1094	G
2	N	984	A
2	N	998	A

There are no RNA pucker outliers to report.

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 42 ligands modelled in this entry, 24 are monoatomic - leaving 18 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	GOL	E	1101	-	5,5,5	0.09	0	5,5,5	0.34	0
7	CTP	A	1007	4	26,30,30	1.11	1 (3%)	39,47,47	1.18	2 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
6	GOL	H	1101	-	5,5,5	1.06	0	5,5,5	1.08	0
6	GOL	D	1008	-	5,5,5	0.07	0	5,5,5	0.28	0
7	CTP	J	1006	4	26,30,30	0.55	0	39,47,47	0.59	0
7	CTP	M	1006	4	26,30,30	0.54	0	39,47,47	0.53	0
6	GOL	A	1005	-	5,5,5	1.01	0	5,5,5	1.00	0
6	GOL	J	1005	-	5,5,5	0.08	0	5,5,5	0.34	0
7	CTP	D	1009	4	26,30,30	0.53	0	39,47,47	0.53	0
7	CTP	P	1005	4	26,30,30	0.55	0	39,47,47	0.52	0
6	GOL	M	1005	-	5,5,5	0.12	0	5,5,5	0.36	0
6	GOL	D	1007	-	5,5,5	0.09	0	5,5,5	0.32	0
7	CTP	G	1005	4	26,30,30	0.55	0	39,47,47	0.52	0
6	GOL	D	1006	-	5,5,5	0.10	0	5,5,5	0.28	0
6	GOL	D	1005	-	5,5,5	0.08	0	5,5,5	0.33	0
6	GOL	A	1006	-	5,5,5	0.08	0	5,5,5	0.32	0
6	GOL	K	1101	-	5,5,5	0.09	0	5,5,5	0.30	0
6	GOL	E	1102	-	5,5,5	0.13	0	5,5,5	0.29	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GOL	E	1101	-	-	4/4/4/4	-
7	CTP	A	1007	4	-	4/22/38/38	0/2/2/2
6	GOL	H	1101	-	-	1/4/4/4	-
6	GOL	D	1008	-	-	0/4/4/4	-
7	CTP	J	1006	4	-	6/22/38/38	0/2/2/2
7	CTP	M	1006	4	-	6/22/38/38	0/2/2/2
6	GOL	A	1005	-	-	1/4/4/4	-
6	GOL	J	1005	-	-	2/4/4/4	-
7	CTP	D	1009	4	-	4/22/38/38	0/2/2/2
7	CTP	P	1005	4	-	4/22/38/38	0/2/2/2
6	GOL	M	1005	-	-	2/4/4/4	-
6	GOL	D	1007	-	-	4/4/4/4	-
7	CTP	G	1005	4	-	4/22/38/38	0/2/2/2
6	GOL	D	1006	-	-	3/4/4/4	-
6	GOL	D	1005	-	-	2/4/4/4	-
6	GOL	A	1006	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GOL	K	1101	-	-	2/4/4/4	-
6	GOL	E	1102	-	-	2/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1007	CTP	C5-C4	-2.52	1.37	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	1007	CTP	C2'-C1'-N1	-3.26	103.97	113.22
7	A	1007	CTP	O2-C2-N3	-2.27	118.64	122.33

There are no chirality outliers.

5 of 53 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1006	GOL	O1-C1-C2-C3
6	D	1005	GOL	C1-C2-C3-O3
6	D	1007	GOL	O1-C1-C2-O2
6	D	1007	GOL	O1-C1-C2-C3
6	D	1007	GOL	C1-C2-C3-O3

There are no ring outliers.

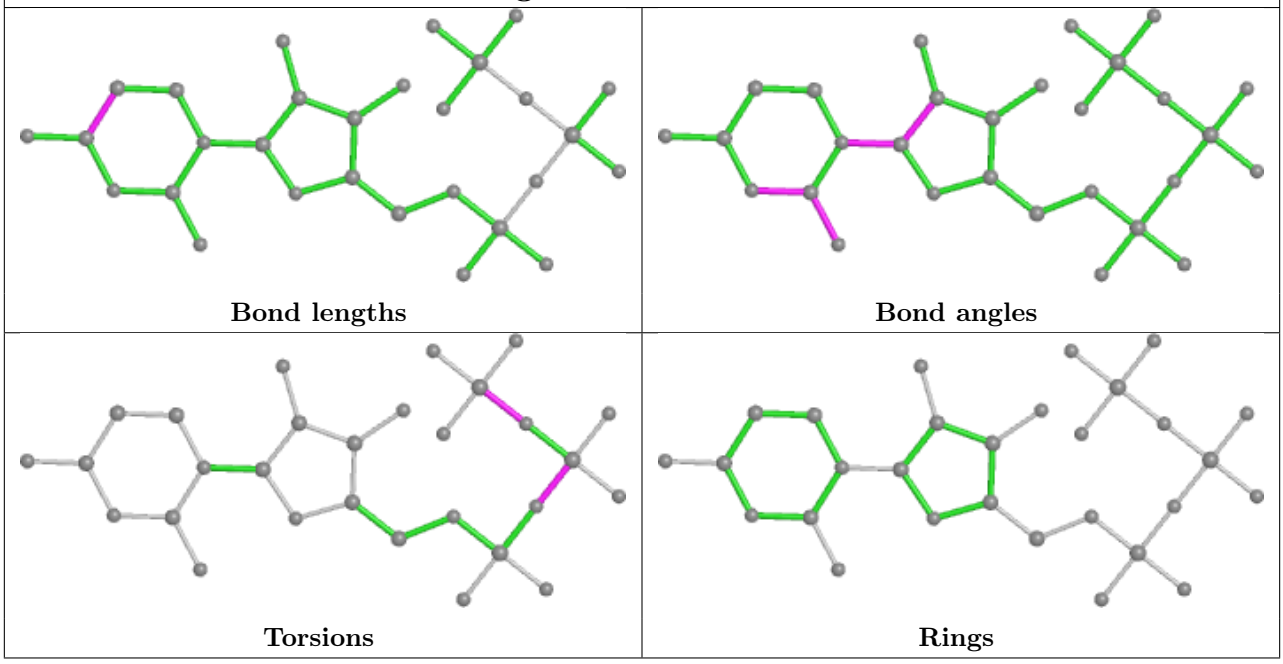
6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1007	CTP	2	0
6	H	1101	GOL	1	0
7	J	1006	CTP	1	0
7	D	1009	CTP	1	0
6	M	1005	GOL	2	0
6	E	1102	GOL	1	0

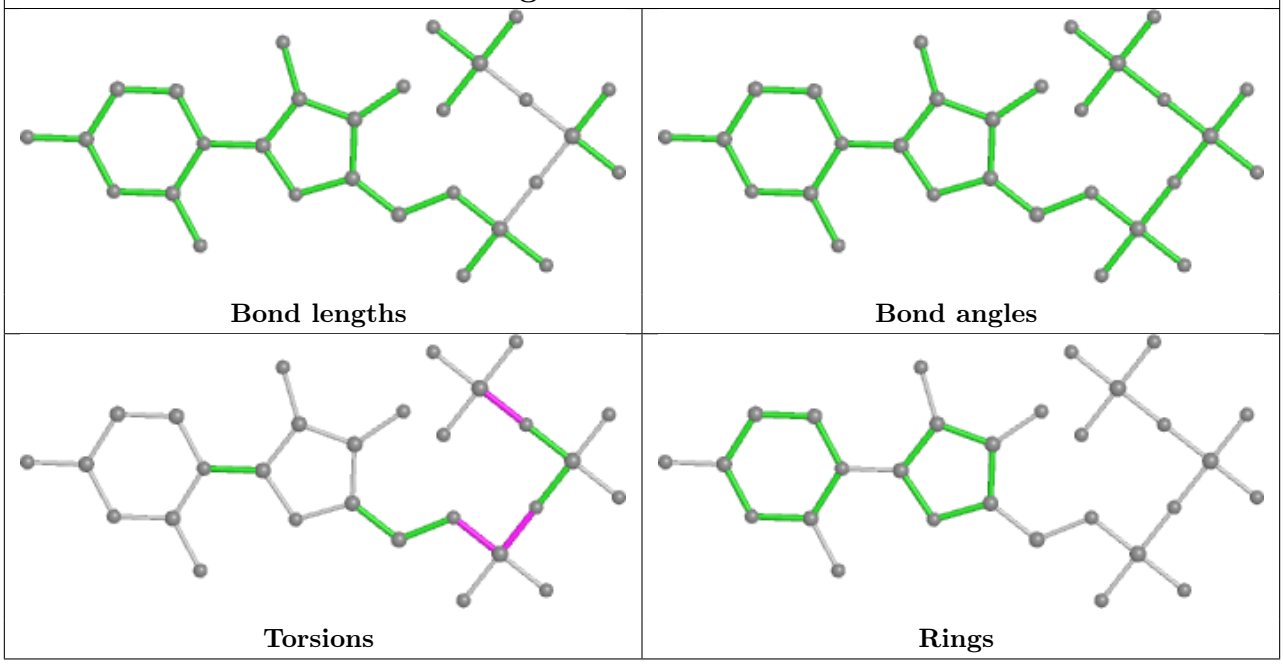
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be

highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

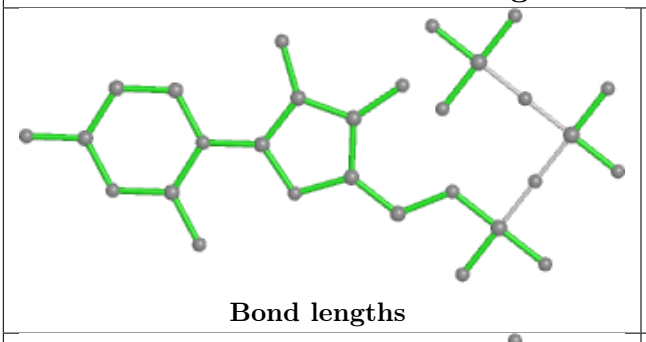
Ligand CTP A 1007



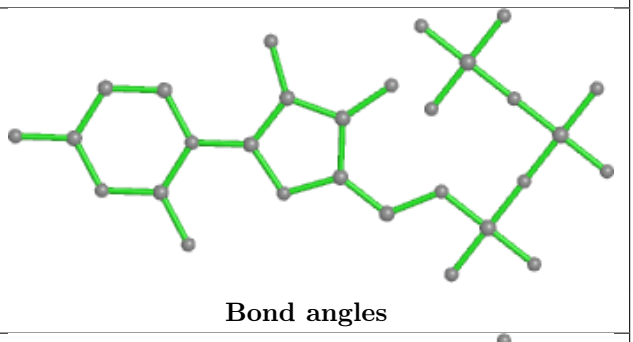
Ligand CTP J 1006



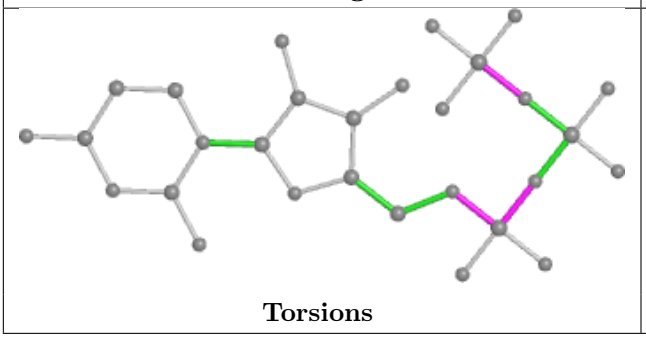
Ligand CTP M 1006



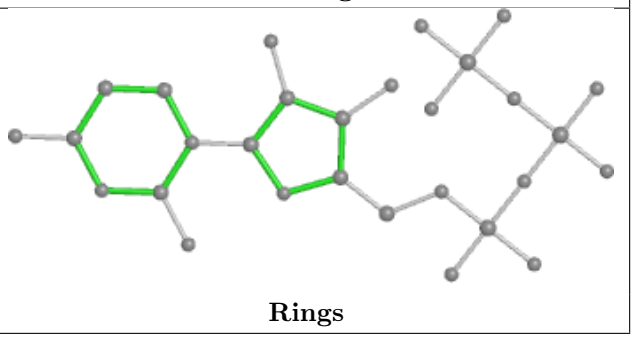
Bond lengths



Bond angles

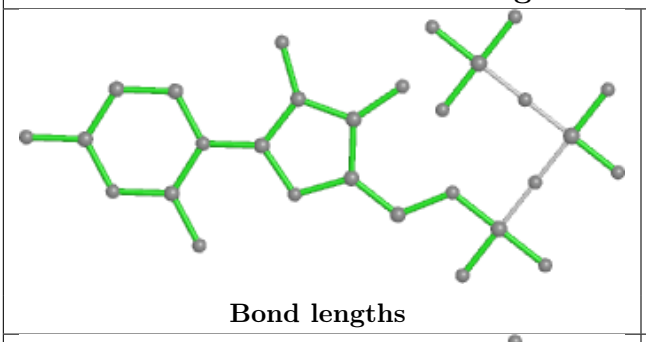


Torsions

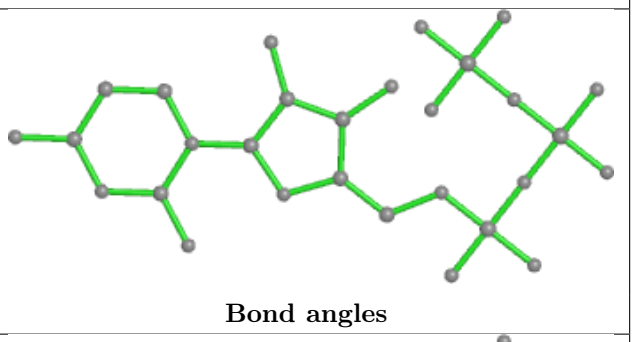


Rings

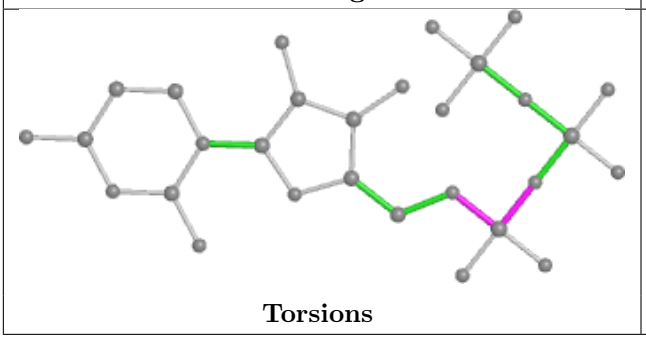
Ligand CTP D 1009



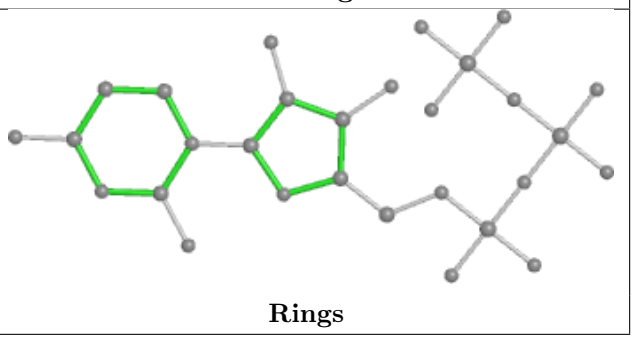
Bond lengths



Bond angles

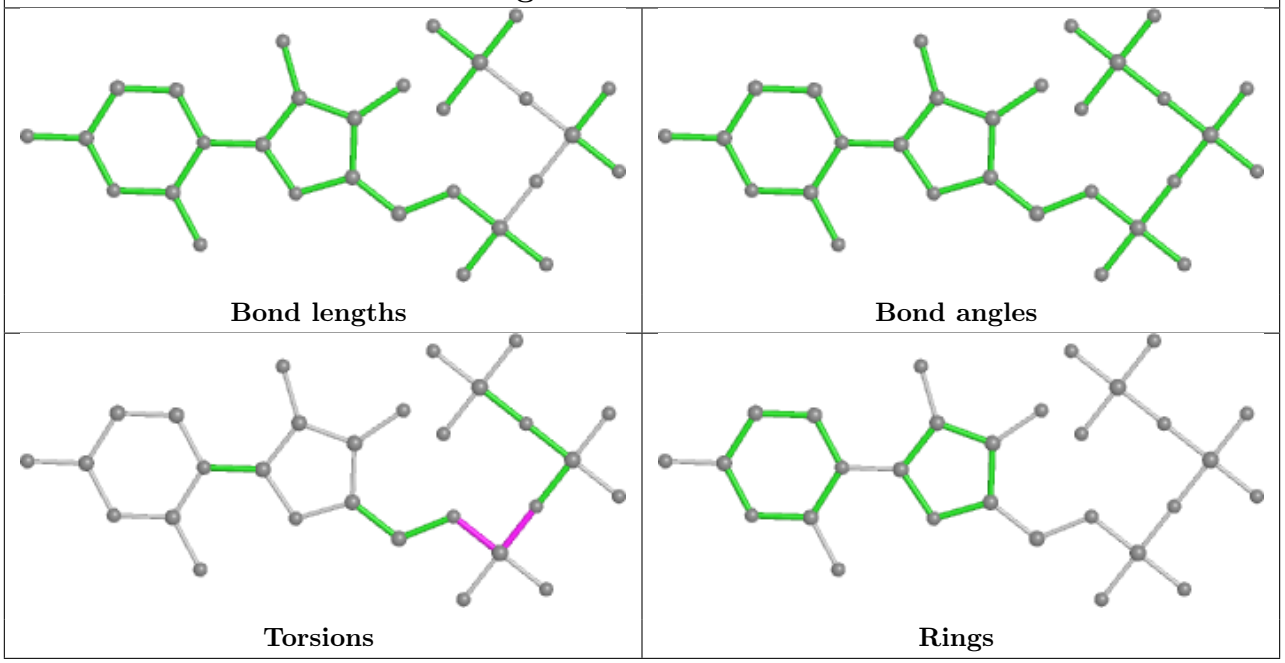


Torsions

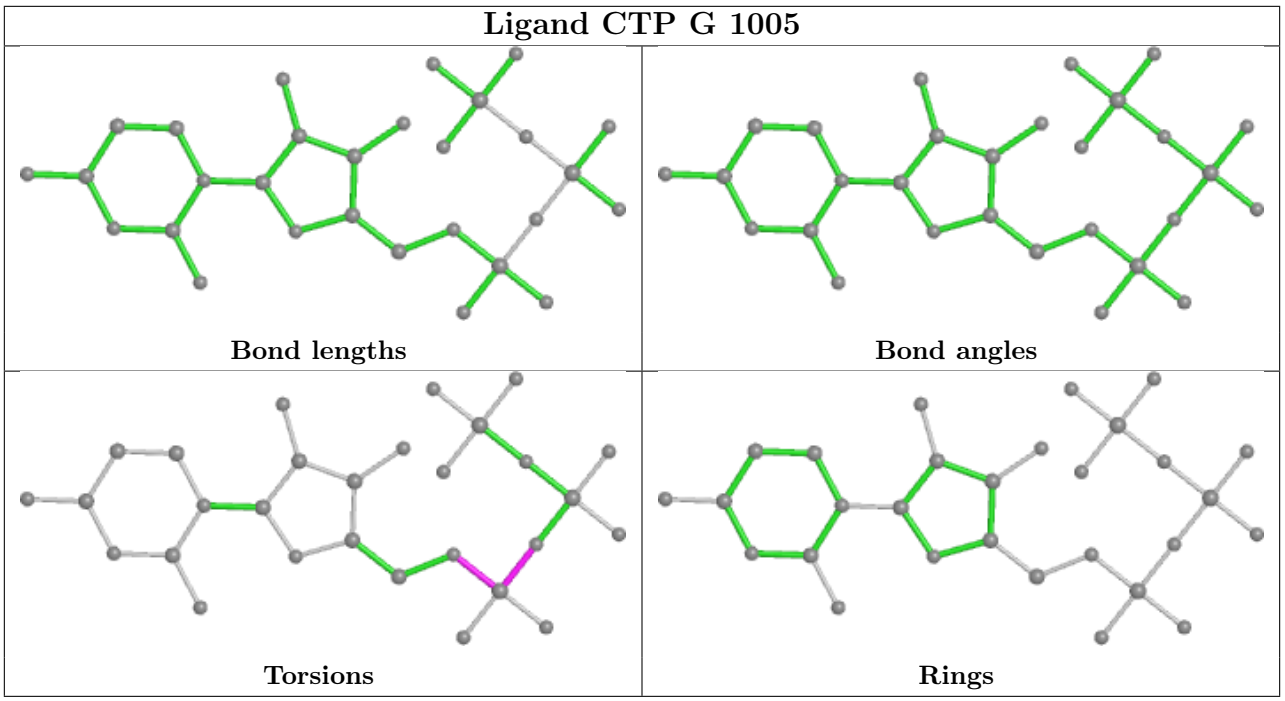


Rings

Ligand CTP P 1005



Ligand CTP G 1005



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	621/647 (95%)	0.06	10 (1%) 72 69	40, 57, 75, 97	0
1	D	622/647 (96%)	0.08	4 (0%) 89 89	38, 54, 75, 92	0
1	G	623/647 (96%)	0.16	21 (3%) 45 41	44, 62, 81, 105	0
1	J	625/647 (96%)	0.23	19 (3%) 50 46	45, 63, 91, 108	0
1	M	624/647 (96%)	0.16	17 (2%) 54 51	42, 56, 73, 87	0
1	P	616/647 (95%)	0.69	87 (14%) 2 2	51, 93, 121, 141	0
2	B	12/30 (40%)	-0.03	0 100 100	43, 48, 89, 107	0
2	E	13/30 (43%)	0.28	1 (7%) 13 11	47, 50, 112, 118	0
2	H	12/30 (40%)	-0.13	0 100 100	53, 56, 82, 97	2 (16%)
2	K	16/30 (53%)	1.04	5 (31%) 0 0	49, 61, 106, 107	6 (37%)
2	N	19/30 (63%)	1.38	6 (31%) 0 0	46, 53, 106, 108	9 (47%)
2	Q	12/30 (40%)	0.27	2 (16%) 1 1	67, 72, 107, 114	2 (16%)
3	C	9/9 (100%)	-0.24	0 100 100	46, 48, 50, 53	0
3	F	9/9 (100%)	0.04	0 100 100	47, 49, 58, 66	0
3	I	9/9 (100%)	0.01	0 100 100	52, 55, 58, 64	0
3	L	9/9 (100%)	0.21	0 100 100	51, 55, 63, 71	0
3	O	9/9 (100%)	-0.08	0 100 100	46, 49, 56, 64	0
3	R	9/9 (100%)	0.07	0 100 100	64, 71, 78, 78	0
All	All	3869/4116 (93%)	0.24	172 (4%) 34 30	38, 60, 104, 141	19 (0%)

The worst 5 of 172 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	996	C	6.1
2	N	995	C	6.0
2	N	984	A	5.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	N	996	C	5.2
2	E	997	A	4.9

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^{th} 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
6	GOL	H	1101	6/6	0.71	0.54	77,78,80,82	0
4	MG	J	1001	1/1	0.72	0.17	76,76,76,76	1
6	GOL	A	1006	6/6	0.78	0.26	72,74,80,80	0
4	MG	D	1001	1/1	0.81	0.17	69,69,69,69	1
4	MG	G	1002	1/1	0.82	0.09	74,74,74,74	0
6	GOL	D	1005	6/6	0.83	0.24	61,64,66,71	0
4	MG	A	1001	1/1	0.83	0.15	72,72,72,72	0
6	GOL	J	1005	6/6	0.83	0.33	69,71,75,79	0
4	MG	G	1001	1/1	0.84	0.27	78,78,78,78	0
6	GOL	E	1101	6/6	0.84	0.28	58,63,67,69	0
6	GOL	M	1005	6/6	0.84	0.28	47,49,53,53	6
7	CTP	D	1009	29/29	0.84	0.27	55,62,70,74	29
6	GOL	A	1005	6/6	0.85	0.31	63,69,70,80	0
6	GOL	K	1101	6/6	0.86	0.29	70,71,76,76	0
6	GOL	D	1006	6/6	0.87	0.33	57,65,69,70	0
6	GOL	D	1007	6/6	0.87	0.19	51,59,68,74	0
4	MG	P	1001	1/1	0.88	0.19	91,91,91,91	1
4	MG	D	1002	1/1	0.90	0.24	71,71,71,71	0
4	MG	A	1002	1/1	0.90	0.06	70,70,70,70	0
7	CTP	M	1006	29/29	0.90	0.23	52,61,73,77	29
4	MG	M	1001	1/1	0.91	0.16	69,69,69,69	1

Continued on next page...

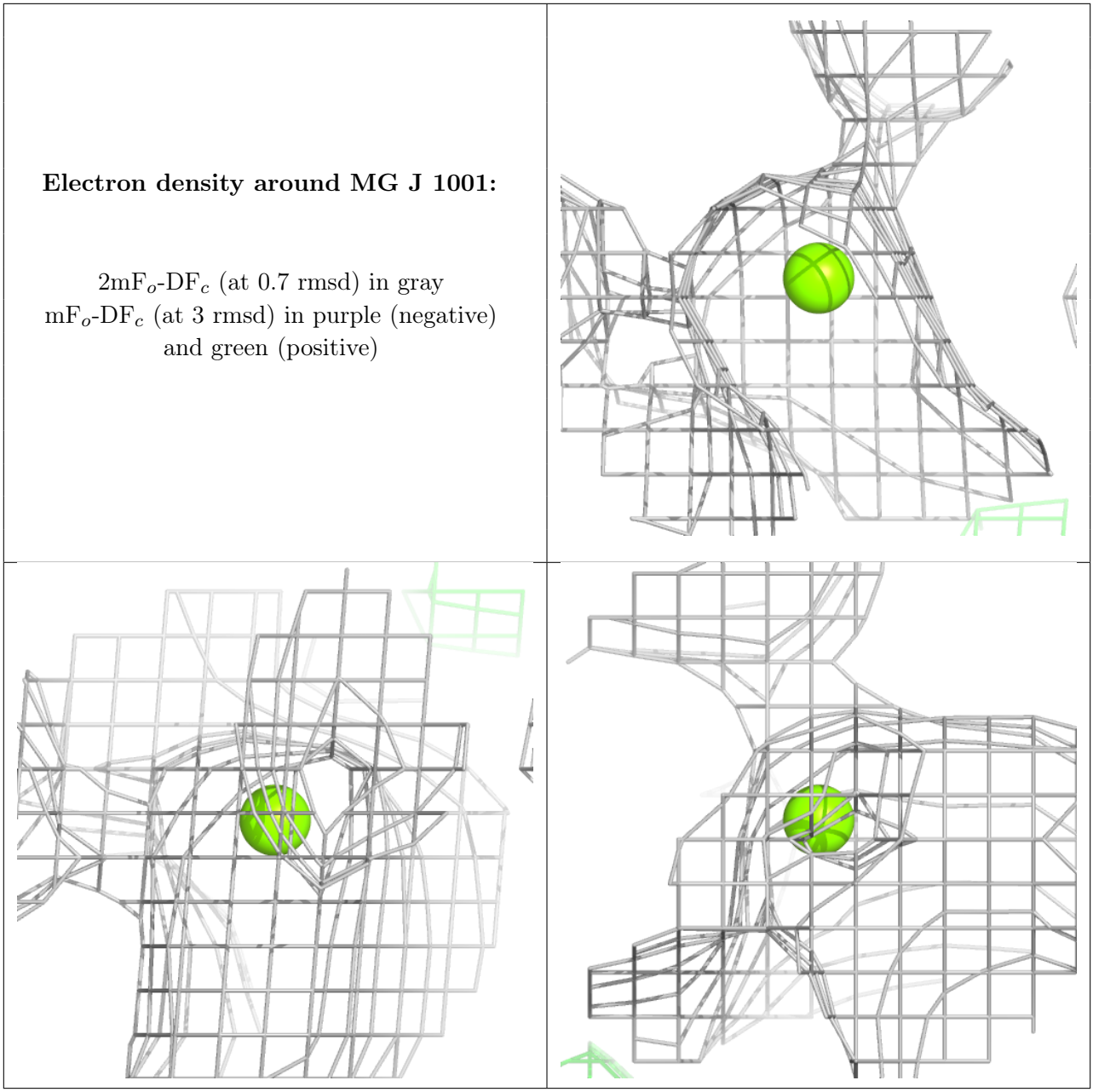
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	GOL	E	1102	6/6	0.91	0.28	55,59,67,67	0
7	CTP	A	1007	29/29	0.91	0.20	43,56,69,71	29
4	MG	M	1002	1/1	0.91	0.12	69,69,69,69	0
6	GOL	D	1008	6/6	0.91	0.29	66,67,71,74	0
4	MG	P	1002	1/1	0.92	0.12	86,86,86,86	0
7	CTP	P	1005	29/29	0.93	0.22	74,84,91,93	29
7	CTP	J	1006	29/29	0.94	0.17	49,63,72,74	29
5	ZN	P	1004	1/1	0.94	0.09	113,113,113,113	0
7	CTP	G	1005	29/29	0.94	0.18	53,60,70,75	29
4	MG	J	1002	1/1	0.95	0.07	72,72,72,72	0
5	ZN	P	1003	1/1	0.96	0.08	82,82,82,82	0
5	ZN	G	1004	1/1	0.98	0.13	72,72,72,72	0
5	ZN	D	1004	1/1	0.99	0.12	68,68,68,68	0
5	ZN	G	1003	1/1	0.99	0.16	52,52,52,52	0
5	ZN	A	1003	1/1	0.99	0.14	59,59,59,59	0
5	ZN	J	1003	1/1	0.99	0.09	58,58,58,58	0
5	ZN	J	1004	1/1	0.99	0.21	71,71,71,71	0
5	ZN	M	1003	1/1	0.99	0.09	59,59,59,59	0
5	ZN	M	1004	1/1	0.99	0.20	68,68,68,68	0
5	ZN	A	1004	1/1	0.99	0.17	63,63,63,63	0
5	ZN	D	1003	1/1	0.99	0.16	48,48,48,48	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.

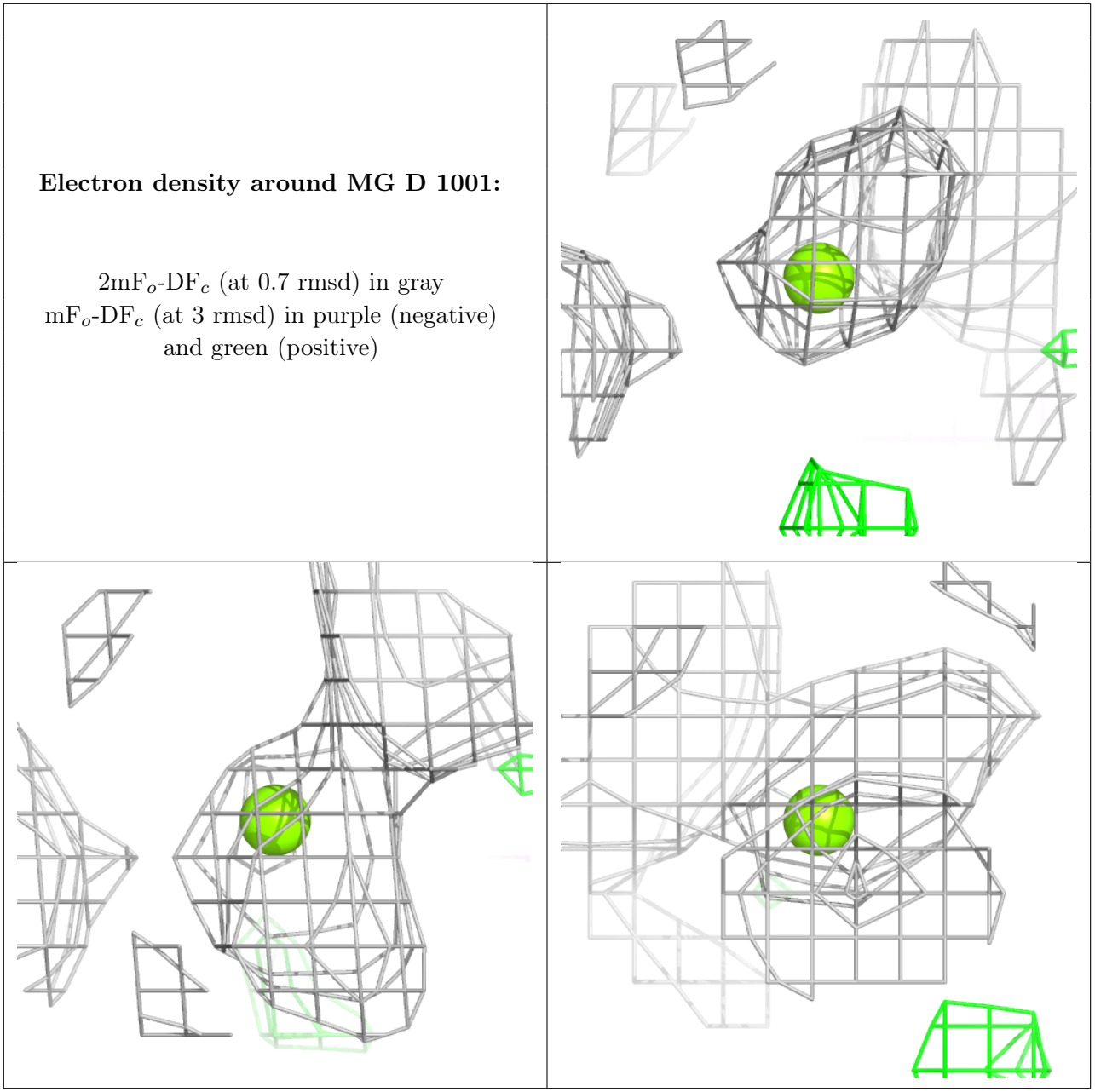
Electron density around MG J 1001:

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



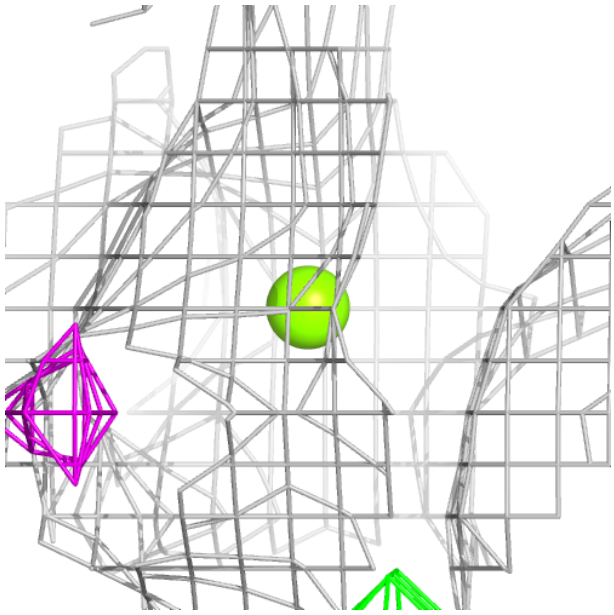
Electron density around MG D 1001:

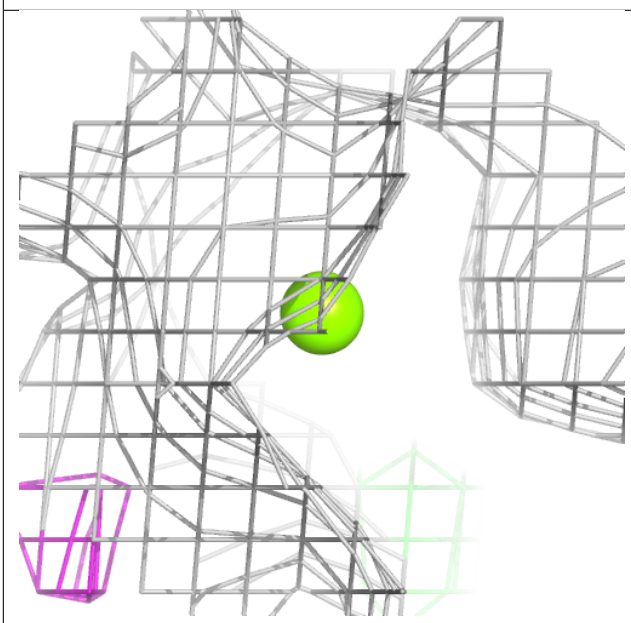
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

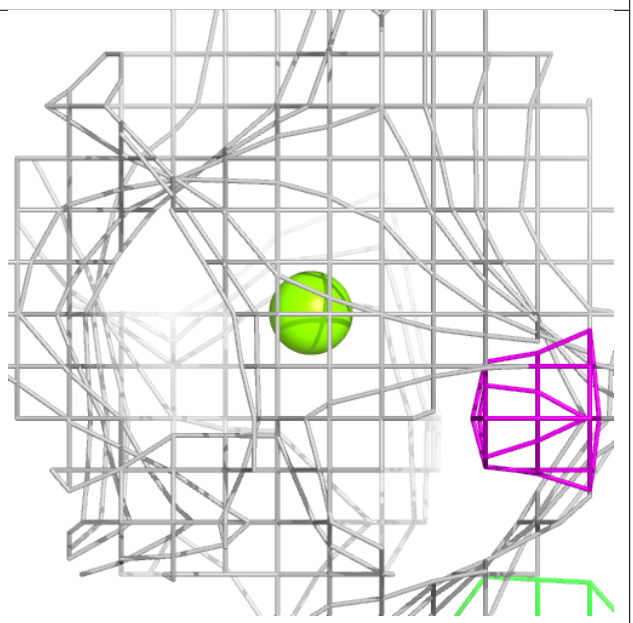


Electron density around MG G 1002:

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

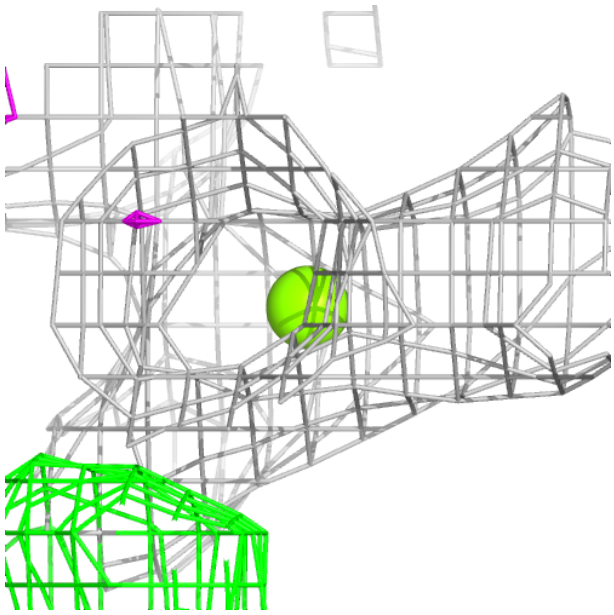


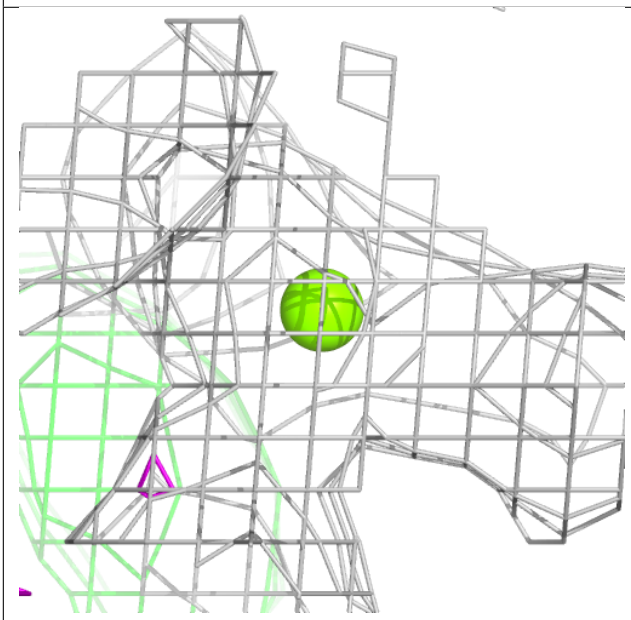


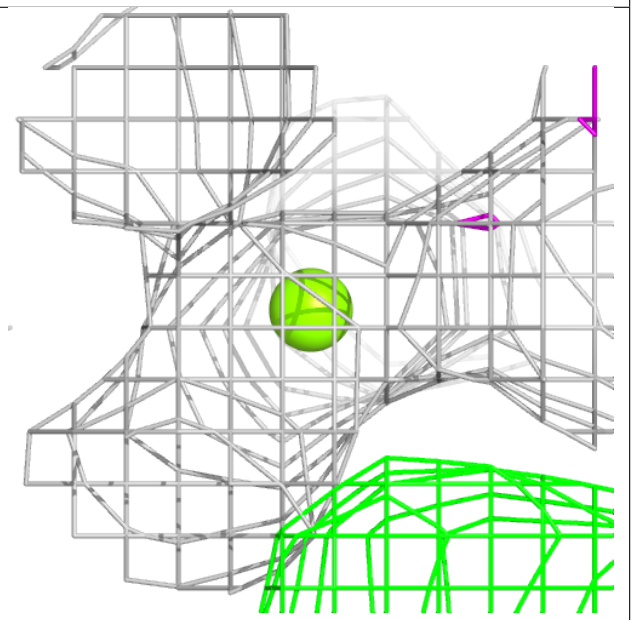


Electron density around MG A 1001:

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

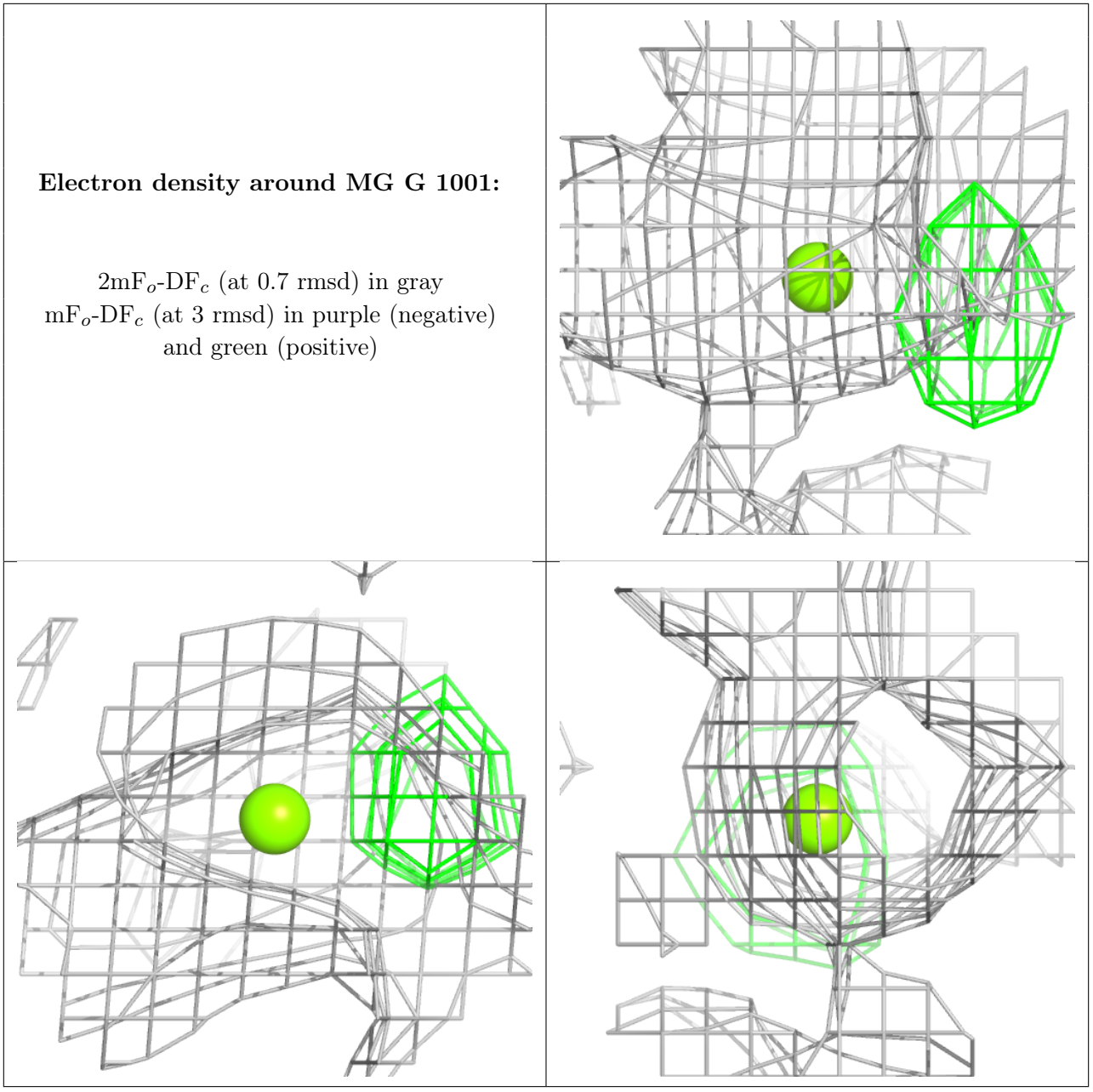






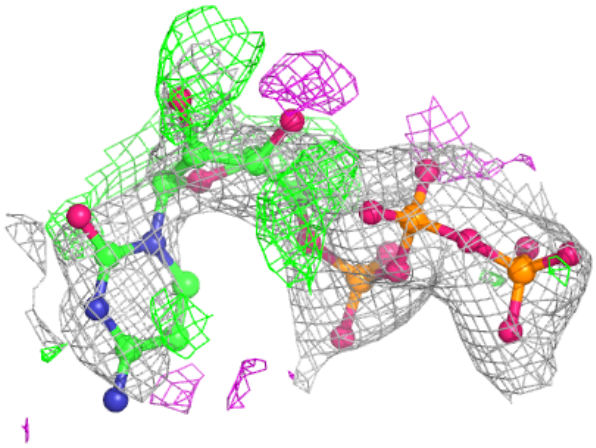
Electron density around MG G 1001:

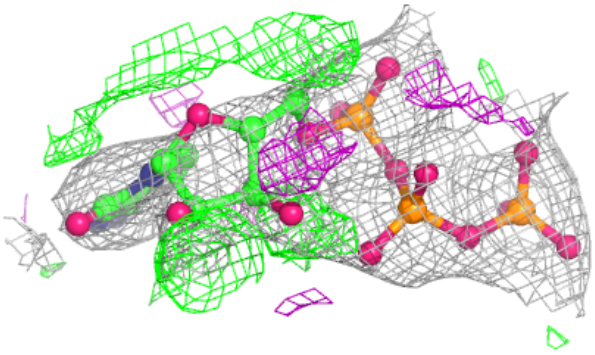
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

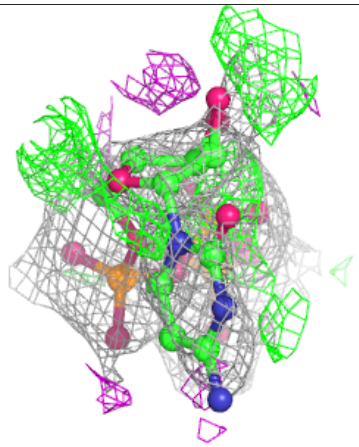


Electron density around CTP D 1009:

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

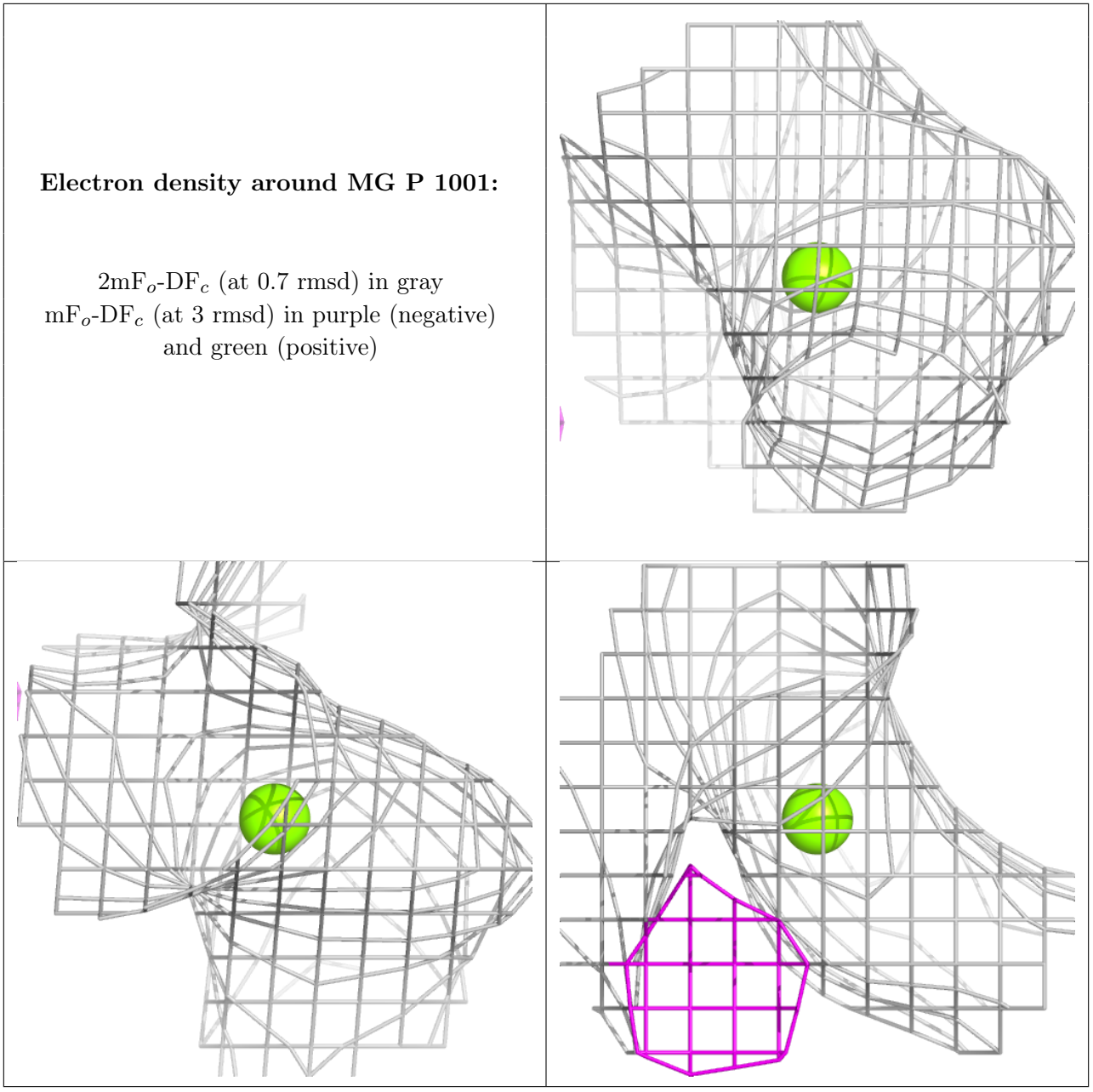






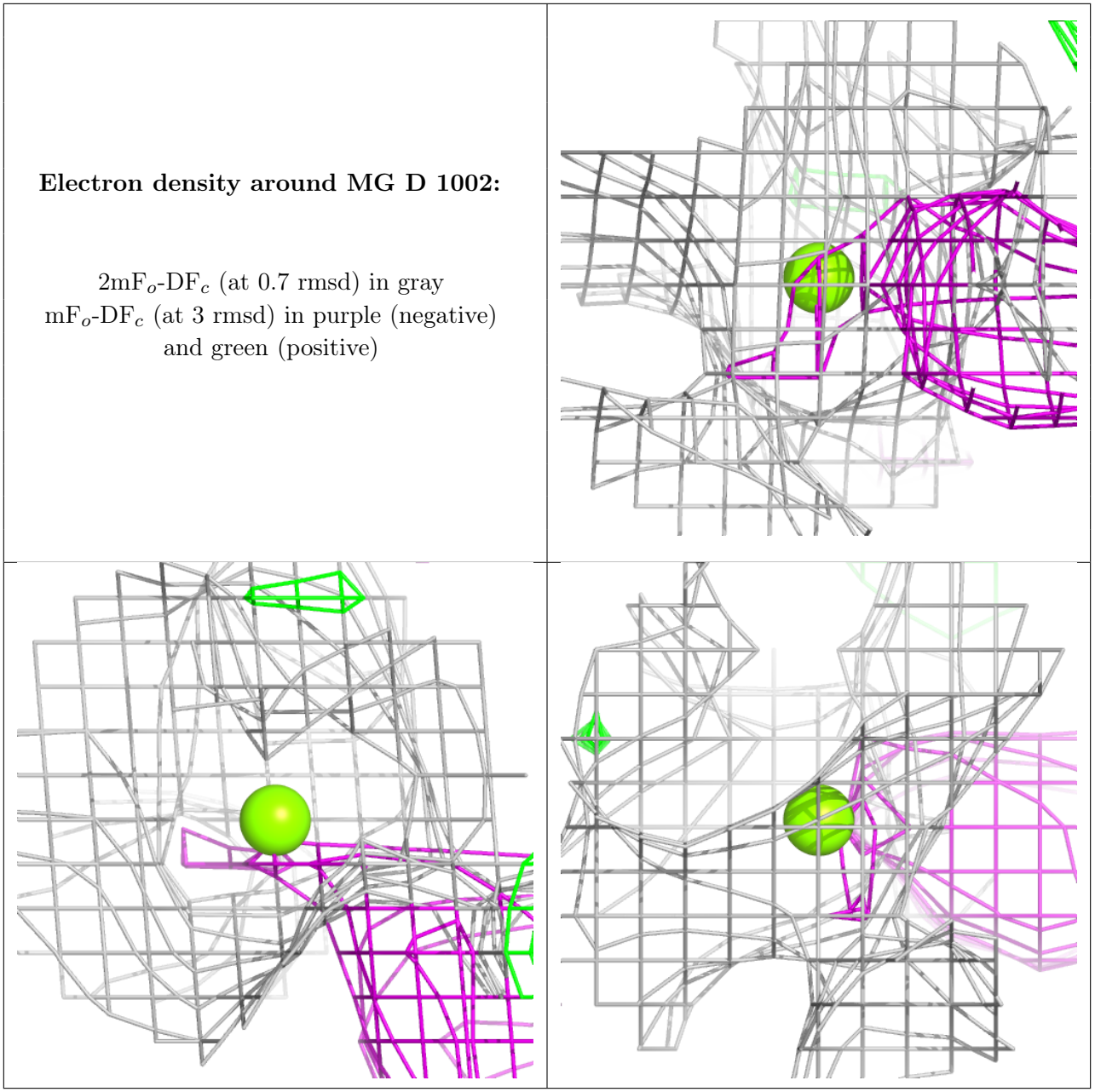
Electron density around MG P 1001:

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



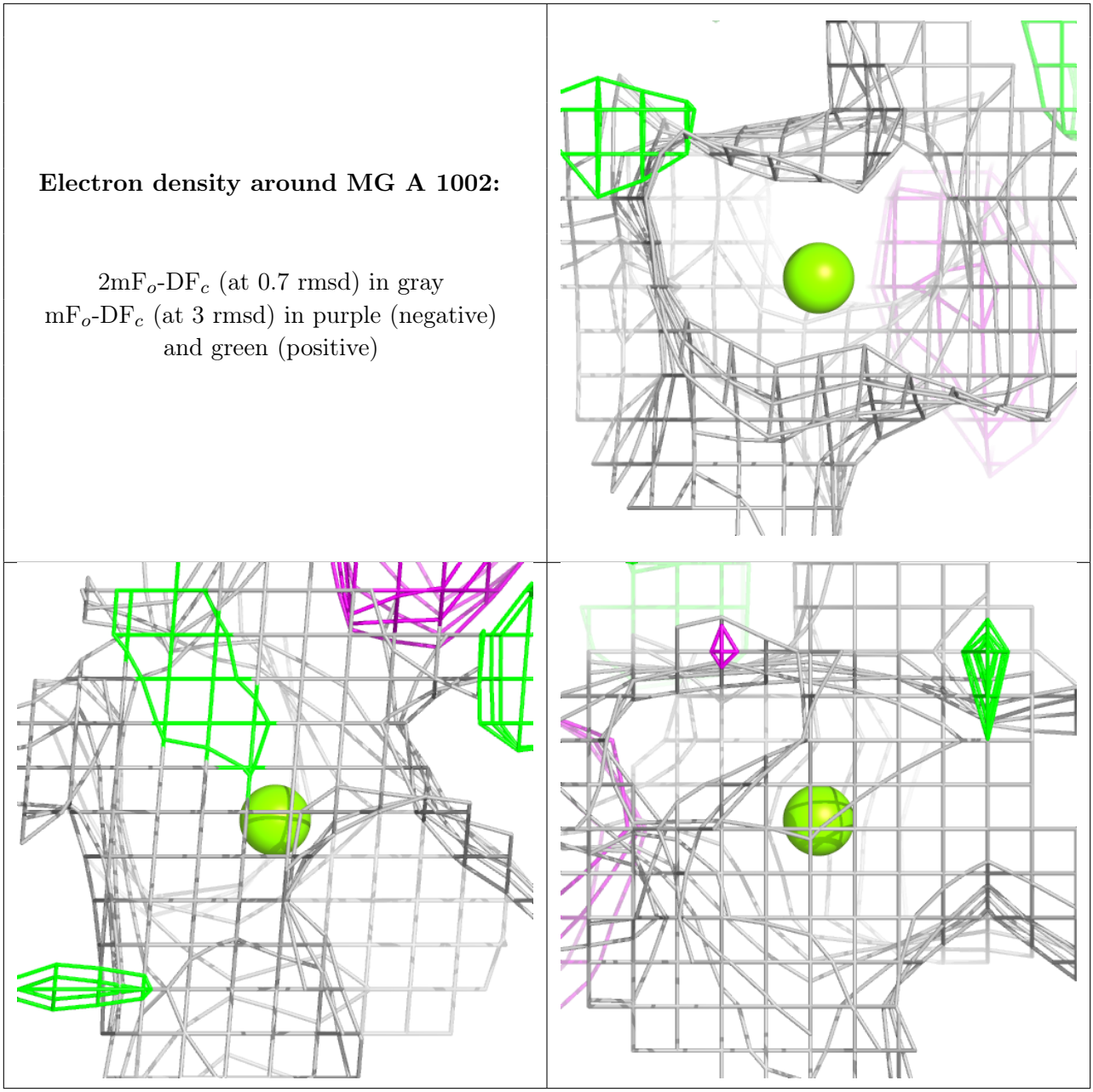
Electron density around MG D 1002:

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



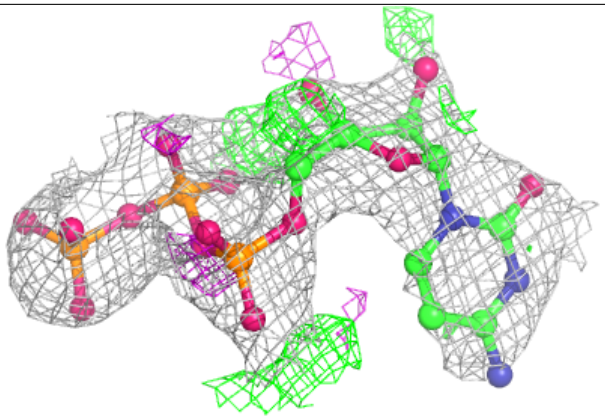
Electron density around MG A 1002:

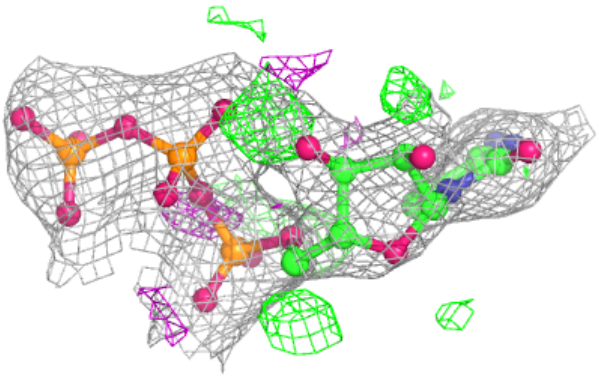
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

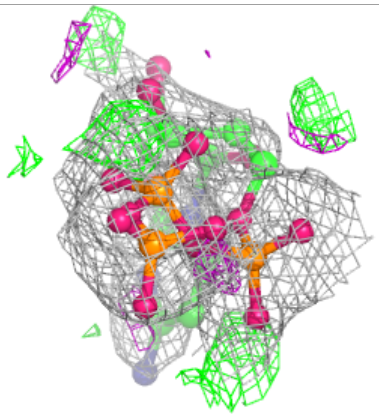


Electron density around CTP M 1006:

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

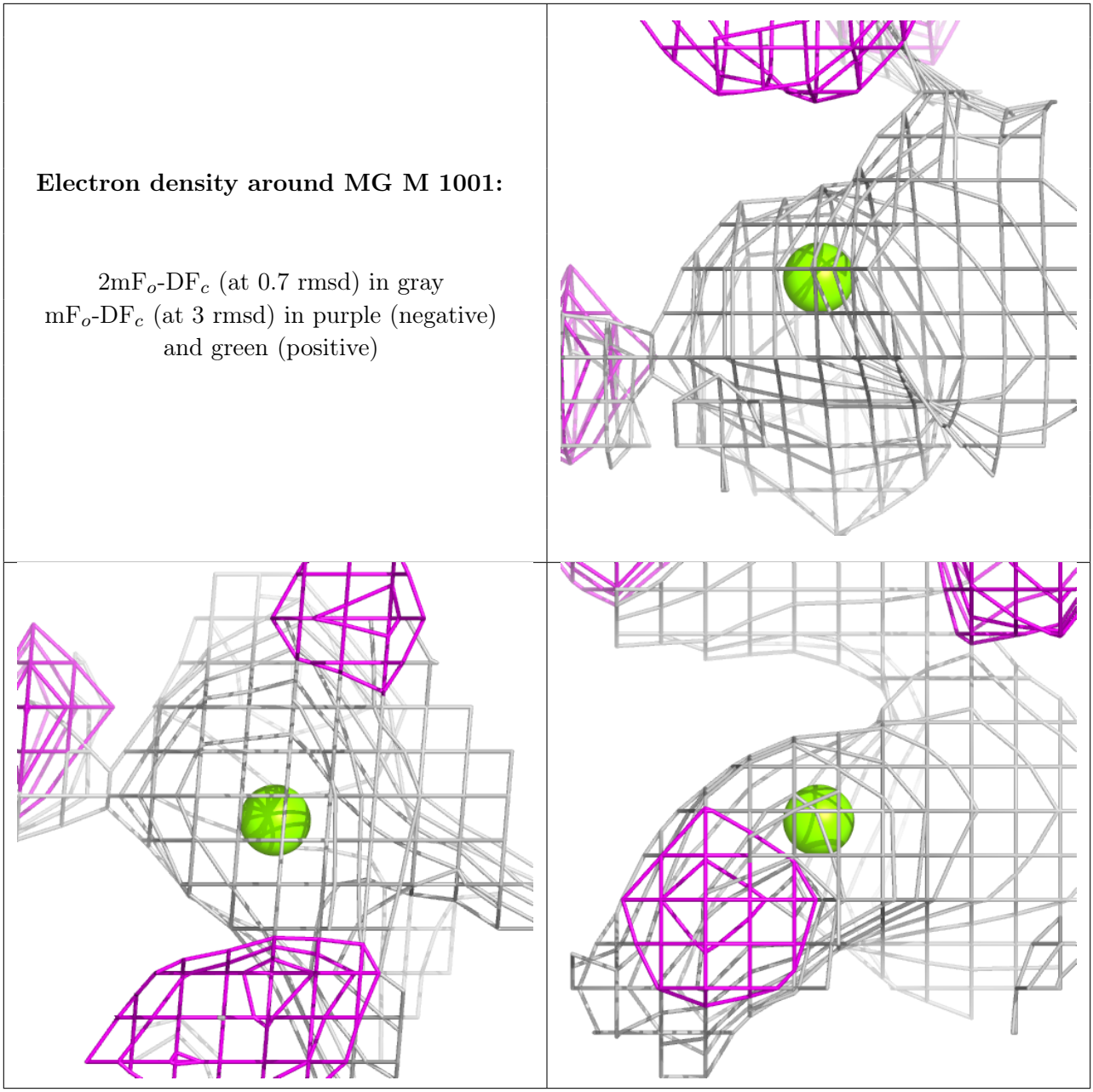






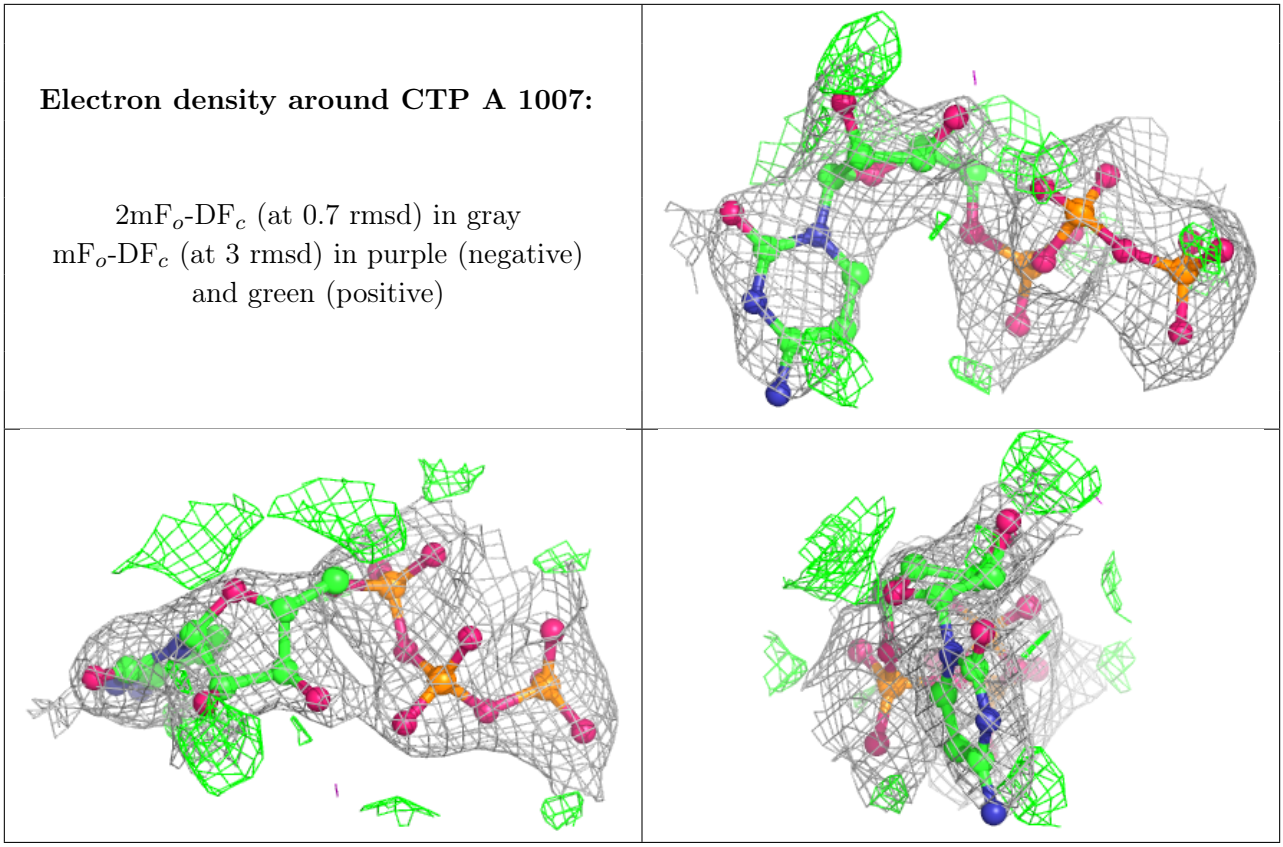
Electron density around MG M 1001:

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



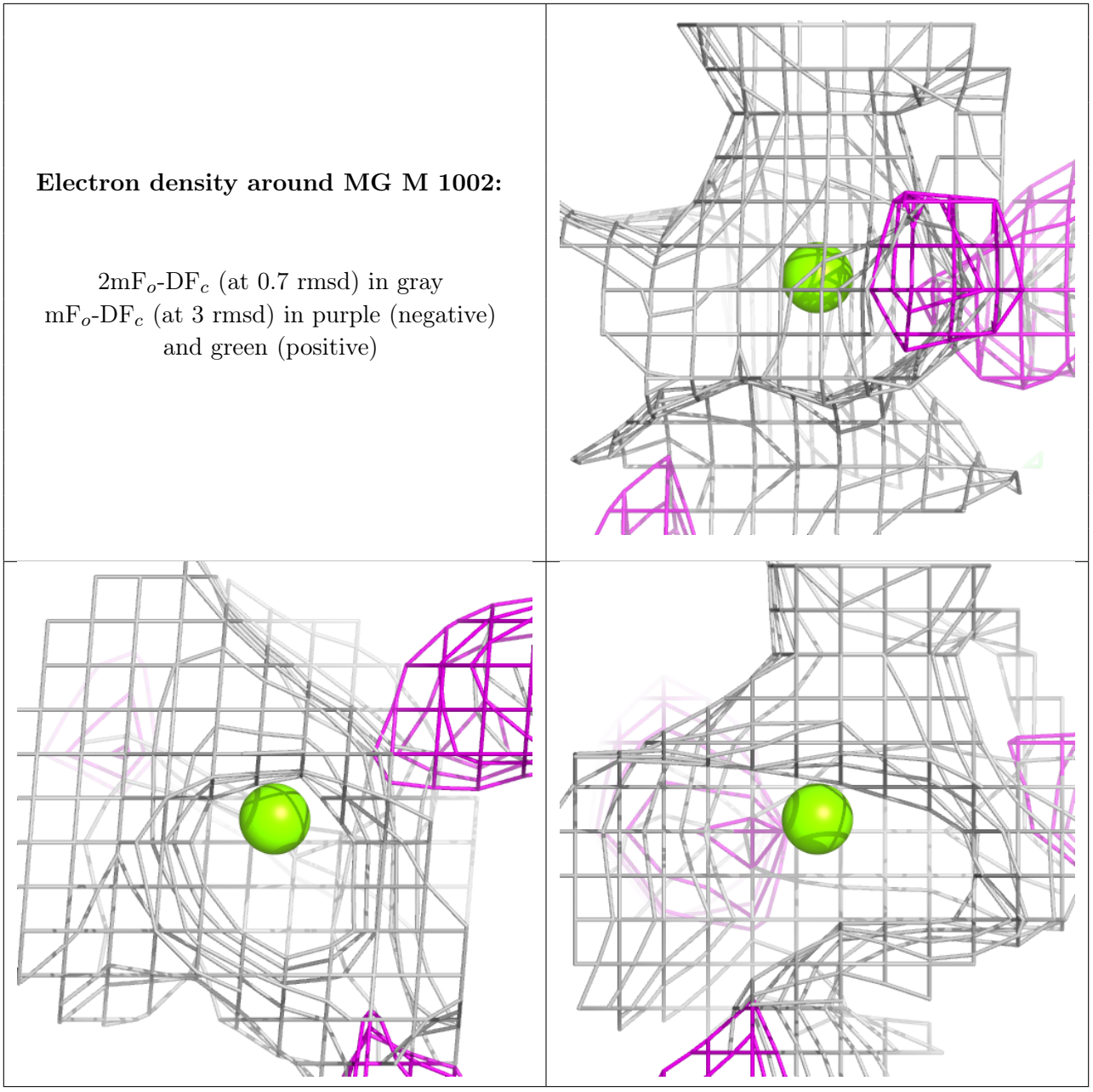
Electron density around CTP A 1007:

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



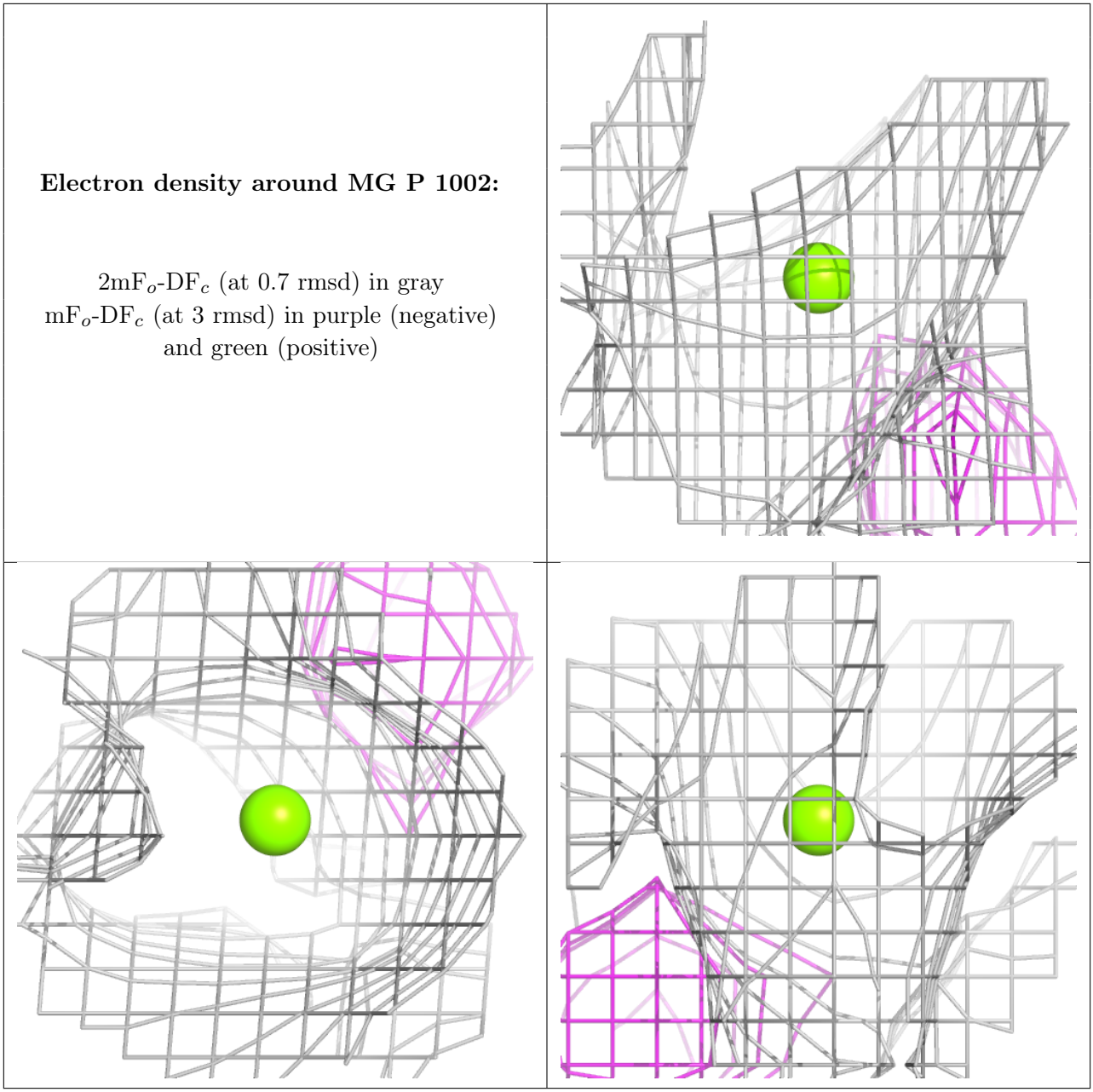
Electron density around MG M 1002:

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



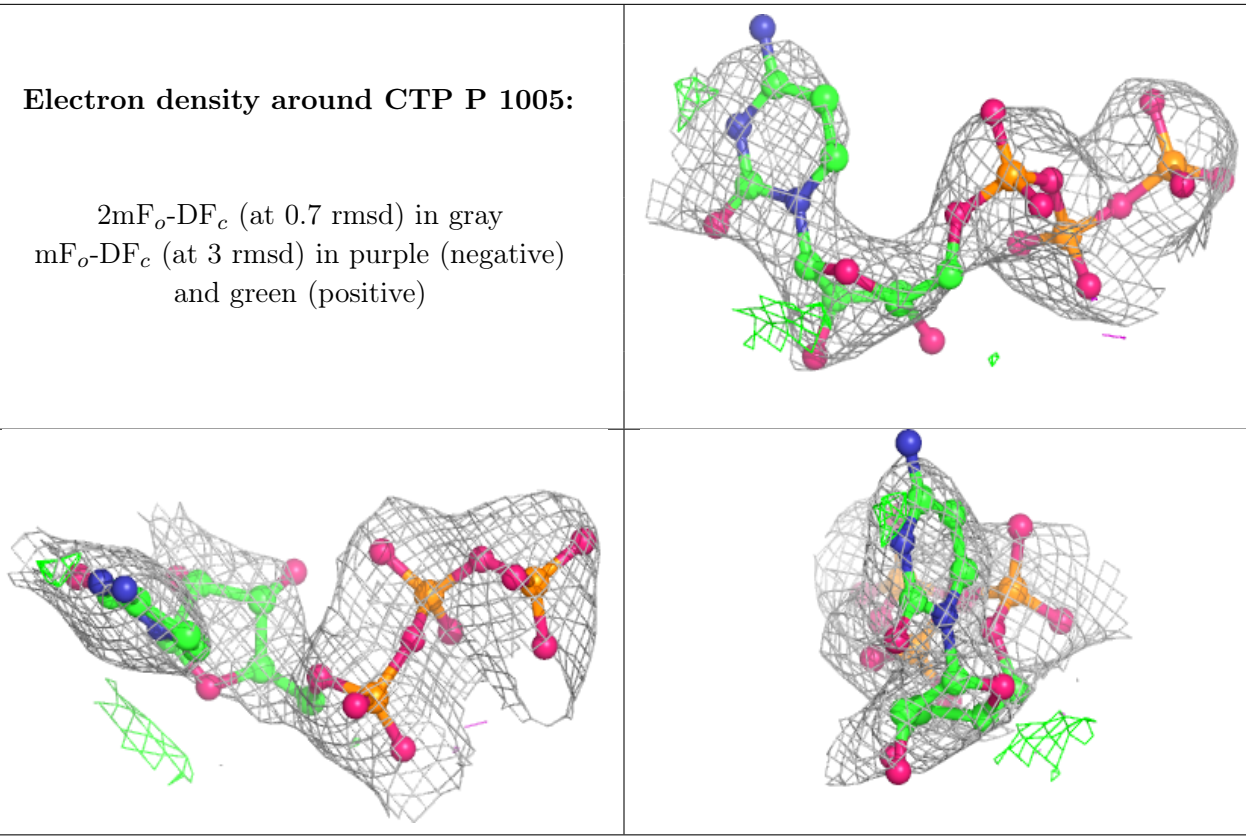
Electron density around MG P 1002:

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



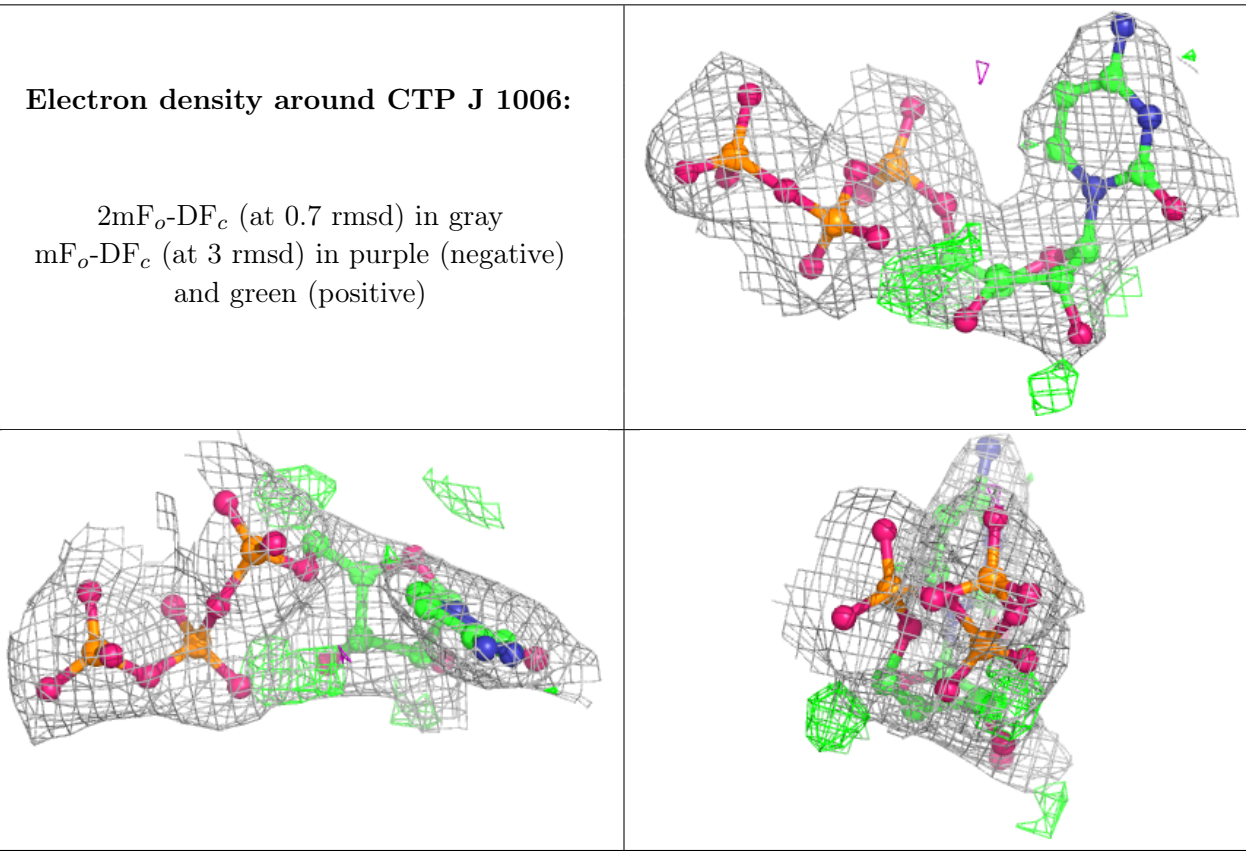
Electron density around CTP P 1005:

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



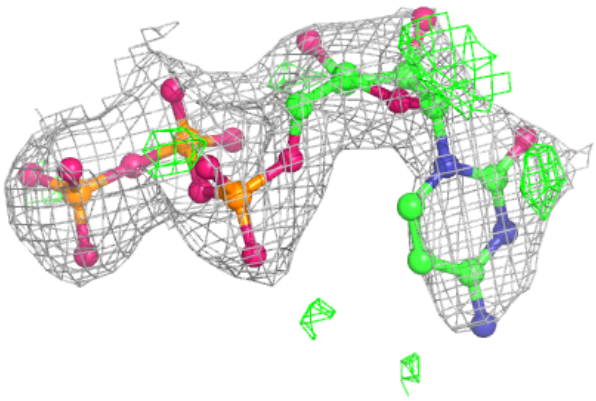
Electron density around CTP J 1006:

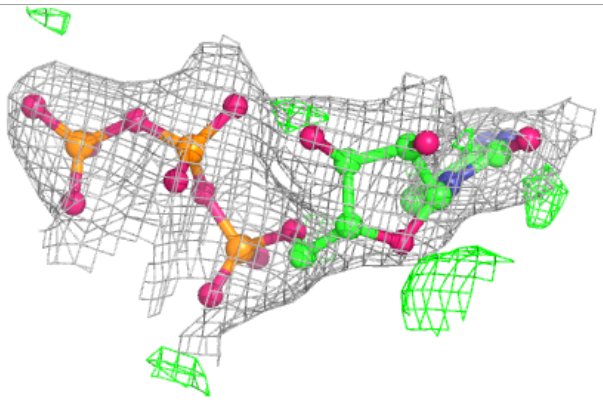
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

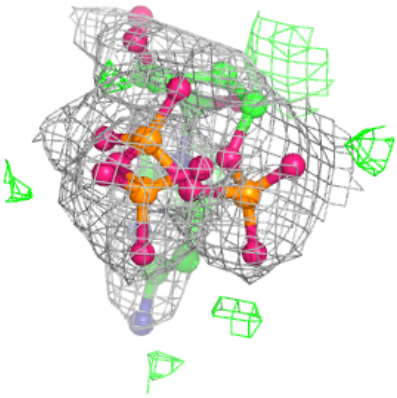


Electron density around CTP G 1005:

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

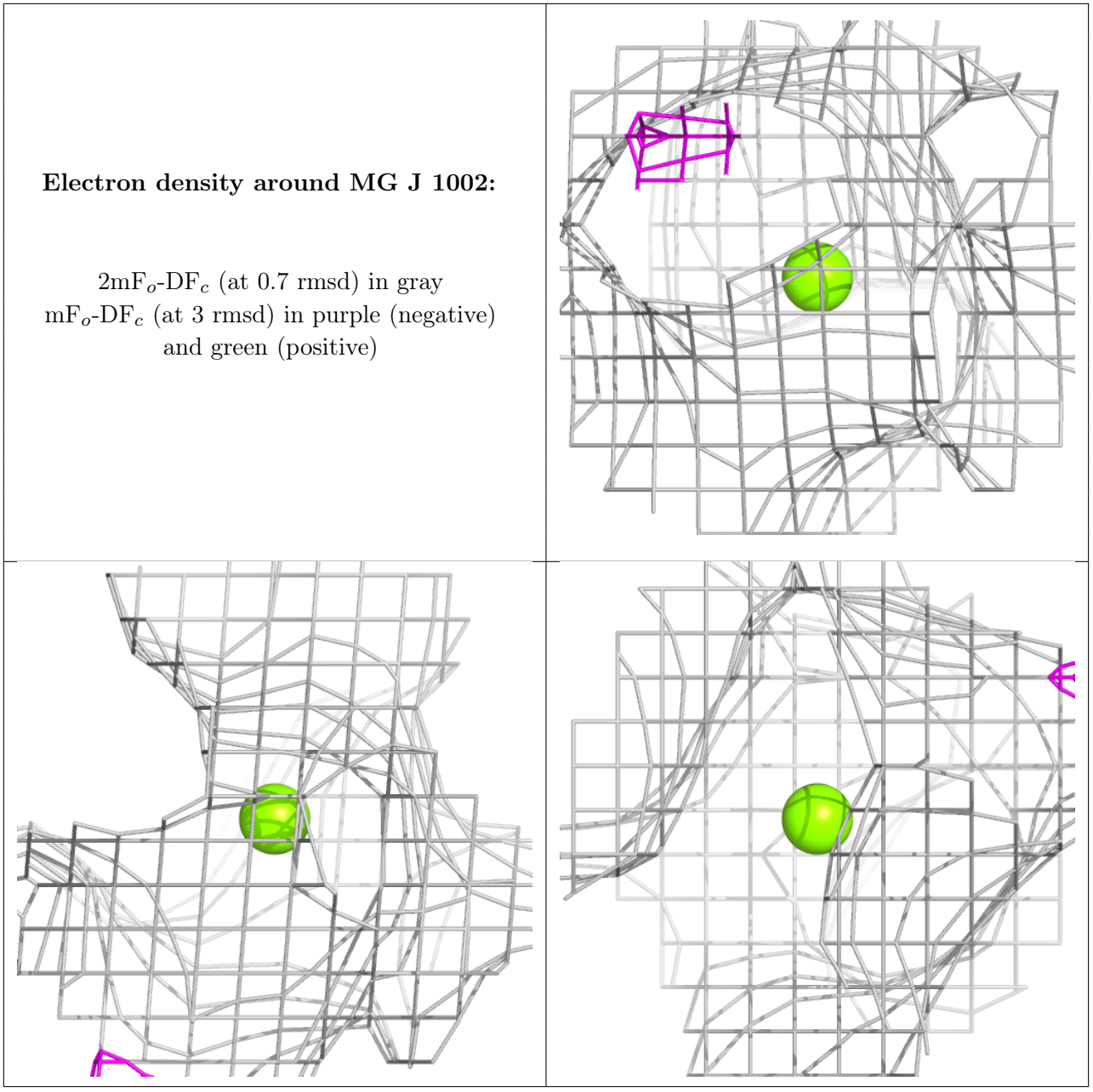






Electron density around MG J 1002:

$2mF_o$-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.