



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2022 – 06:20 am GMT

PDB ID : 7PU5
Title : Structure of SFPQ-NONO complex
Authors : Fribourg, S.
Deposited on : 2021-09-28
Resolution : 3.00 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.27
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

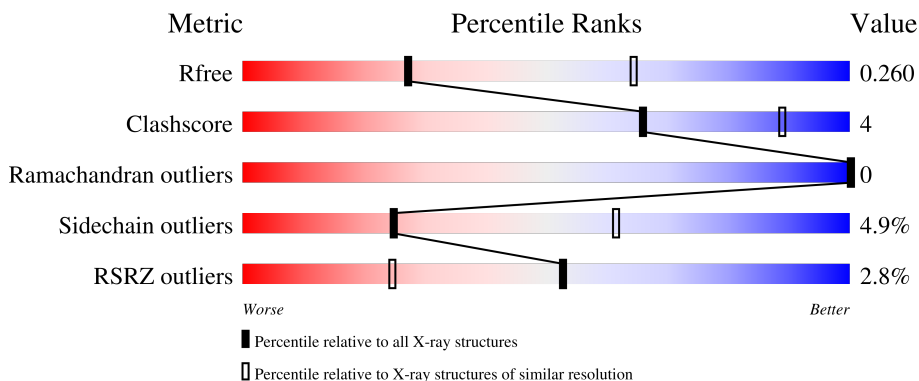
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 $\leq 5\%$. 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	259	 2% 81% 12% 7%
1	C	259	 4% 79% 14% 7%
1	E	259	 % 76% 15% 8%
1	G	259	 4% 78% 13% 8%
1	I	259	 4% 81% 10% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	259	<p>6% 77% 15% • 8%</p>
2	B	259	<p>79% 12% • 8%</p>
2	D	259	<p>2% 81% 11% • 8%</p>
2	F	259	<p>1% 80% 11% • 8%</p>
2	H	259	<p>81% 11% • 7%</p>
2	J	259	<p>3% 80% 12% • 8%</p>
2	L	259	<p>2% 80% 12% 7%</p>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 23285 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 Non-POU domain-containing octamer-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	242	Total 1963	C 1234	N 352	O 368	S 9	0	0	0
1	C	242	Total 1968	C 1238	N 353	O 368	S 9	0	0	0
1	E	237	Total 1929	C 1215	N 346	O 359	S 9	0	0	0
1	G	237	Total 1934	C 1219	N 347	O 359	S 9	0	0	0
1	I	238	Total 1930	C 1215	N 344	O 362	S 9	0	0	0
1	K	239	Total 1939	C 1221	N 346	O 363	S 9	0	0	0

- Molecule 2 is a protein called Splicing factor, proline- and glutamine-rich.

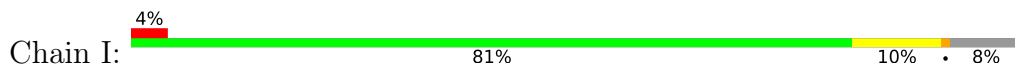
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	239	Total 1932	C 1215	N 339	O 371	S 7	0	0	0
2	D	239	Total 1932	C 1215	N 339	O 371	S 7	0	0	0
2	F	239	Total 1932	C 1215	N 339	O 371	S 7	0	0	0
2	H	240	Total 1941	C 1220	N 340	O 374	S 7	0	0	0
2	J	239	Total 1932	C 1215	N 339	O 371	S 7	0	0	0
2	L	240	Total 1941	C 1220	N 340	O 374	S 7	0	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

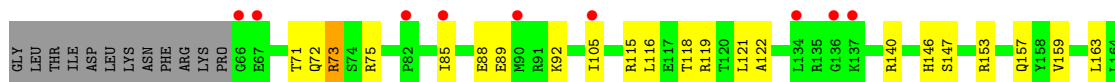
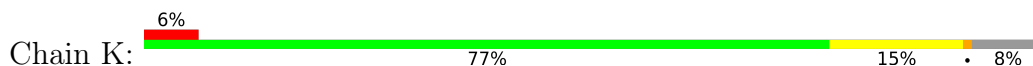
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Mg 2 2	0	0
3	D	2	Total Mg 2 2	0	0
3	F	2	Total Mg 2 2	0	0
3	H	2	Total Mg 2 2	0	0
3	J	2	Total Mg 2 2	0	0
3	K	1	Total Mg 1 1	0	0
3	L	1	Total Mg 1 1	0	0



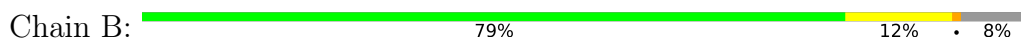
- Molecule 1: Non-POU domain-containing octamer-binding protein



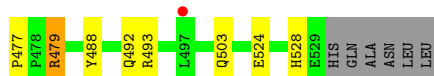
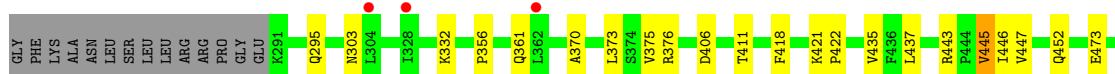
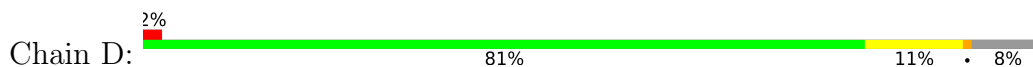
- Molecule 1: Non-POU domain-containing octamer-binding protein



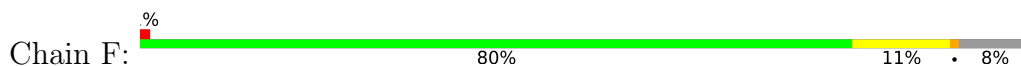
- Molecule 2: Splicing factor, proline- and glutamine-rich



- Molecule 2: Splicing factor, proline- and glutamine-rich

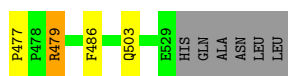
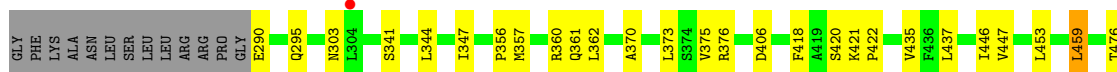
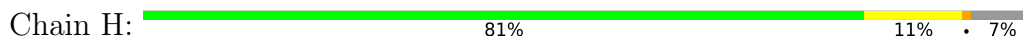


- Molecule 2: Splicing factor, proline- and glutamine-rich

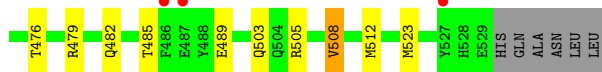
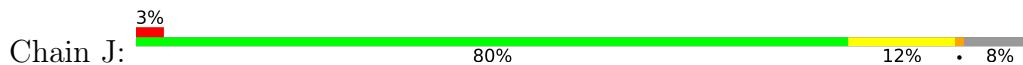




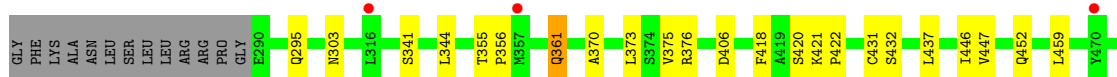
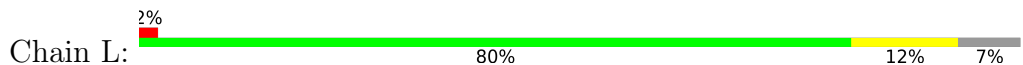
• Molecule 2: Splicing factor, proline- and glutamine-rich



• Molecule 2: Splicing factor, proline- and glutamine-rich



• Molecule 2: Splicing factor, proline- and glutamine-rich



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	467.00Å 66.82Å 126.89Å 90.00° 105.89° 90.00°	Depositor
Resolution (Å)	48.45 – 3.00 48.45 – 3.00	Depositor EDS
% Data completeness (in resolution range)	77.2 (48.45-3.00) 77.2 (48.45-3.00)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.15	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.53 (at 3.01Å)	Xtrriage
Refinement program	BUSTER 2.10.4 (16-JUL-2021)	Depositor
R, R_{free}	0.223 , 0.255 0.226 , 0.260	Depositor DCC
R_{free} test set	2903 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	80.7	Xtrriage
Anisotropy	0.046	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.022 for -h-2*1,-k,l	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	23285	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 59.10 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8551e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 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	0.37	0/2001	0.54	0/2686
1	C	0.39	0/2006	0.53	0/2692
1	E	0.41	0/1966	0.55	0/2639
1	G	0.40	0/1971	0.54	0/2645
1	I	0.38	0/1965	0.55	0/2636
1	K	0.37	0/1975	0.56	0/2650
2	B	0.38	0/1970	0.52	0/2648
2	D	0.37	0/1970	0.50	0/2648
2	F	0.39	0/1970	0.52	0/2648
2	H	0.39	0/1979	0.52	0/2660
2	J	0.38	0/1970	0.52	0/2648
2	L	0.37	0/1979	0.52	0/2660
All	All	0.39	0/23722	0.53	0/31860

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1963	0	1945	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1968	0	1957	15	0
1	E	1929	0	1917	20	0
1	G	1934	0	1927	19	0
1	I	1930	0	1923	15	0
1	K	1939	0	1937	22	0
2	B	1932	0	1903	24	0
2	D	1932	0	1903	17	0
2	F	1932	0	1903	22	0
2	H	1941	0	1909	18	0
2	J	1932	0	1903	21	0
2	L	1941	0	1909	22	0
3	B	2	0	0	0	0
3	D	2	0	0	0	0
3	F	2	0	0	0	0
3	H	2	0	0	0	0
3	J	2	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
All	All	23285	0	23036	183	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 4.

The worst 5 of 183 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:A:216:THR:O	2:B:479:ARG:NH2	1.96	0.98
1:I:251:ARG:NH1	2:J:381:TYR:O	2.10	0.85
2:B:479:ARG:NH1	2:B:487:GLU:OE2	2.11	0.83
1:C:85:ILE:HG23	1:C:105:ILE:HD11	1.65	0.77
1:E:85:ILE:HG23	1:E:105:ILE:HD11	1.65	0.77

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	240/259 (93%)	236 (98%)	4 (2%)	0	100	100
1	C	240/259 (93%)	235 (98%)	5 (2%)	0	100	100
1	E	233/259 (90%)	230 (99%)	3 (1%)	0	100	100
1	G	233/259 (90%)	229 (98%)	4 (2%)	0	100	100
1	I	234/259 (90%)	231 (99%)	3 (1%)	0	100	100
1	K	237/259 (92%)	234 (99%)	3 (1%)	0	100	100
2	B	237/259 (92%)	232 (98%)	5 (2%)	0	100	100
2	D	237/259 (92%)	230 (97%)	7 (3%)	0	100	100
2	F	237/259 (92%)	231 (98%)	6 (2%)	0	100	100
2	H	238/259 (92%)	232 (98%)	6 (2%)	0	100	100
2	J	237/259 (92%)	232 (98%)	5 (2%)	0	100	100
2	L	238/259 (92%)	233 (98%)	5 (2%)	0	100	100
All	All	2841/3108 (91%)	2785 (98%)	56 (2%)	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	210/227 (92%)	200 (95%)	10 (5%)	25	62
1	C	211/227 (93%)	196 (93%)	15 (7%)	14	46
1	E	207/227 (91%)	191 (92%)	16 (8%)	13	42
1	G	208/227 (92%)	196 (94%)	12 (6%)	20	55
1	I	207/227 (91%)	197 (95%)	10 (5%)	25	62
1	K	208/227 (92%)	195 (94%)	13 (6%)	18	51
2	B	205/221 (93%)	194 (95%)	11 (5%)	22	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	205/221 (93%)	198 (97%)	7 (3%)	37	72
2	F	205/221 (93%)	199 (97%)	6 (3%)	42	76
2	H	206/221 (93%)	199 (97%)	7 (3%)	37	72
2	J	205/221 (93%)	197 (96%)	8 (4%)	32	69
2	L	206/221 (93%)	199 (97%)	7 (3%)	37	72
All	All	2483/2688 (92%)	2361 (95%)	122 (5%)	25	61

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	301	GLU
1	K	183	ASP
1	G	183	ASP
1	K	165	GLU
2	L	432	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	J	295	GLN
2	L	295	GLN
2	D	295	GLN
1	E	72	GLN
2	F	295	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	242/259 (93%)	0.27	6 (2%) 57 29	67, 89, 125, 138	0
1	C	242/259 (93%)	0.28	10 (4%) 37 14	57, 85, 126, 141	0
1	E	237/259 (91%)	0.16	3 (1%) 77 51	60, 81, 124, 136	0
1	G	237/259 (91%)	0.28	10 (4%) 36 14	56, 83, 138, 151	0
1	I	238/259 (91%)	0.39	11 (4%) 32 12	64, 93, 135, 162	0
1	K	239/259 (92%)	0.41	16 (6%) 17 5	65, 102, 139, 213	0
2	B	239/259 (92%)	0.17	1 (0%) 92 79	46, 91, 120, 148	0
2	D	239/259 (92%)	0.10	4 (1%) 70 41	50, 84, 115, 143	0
2	F	239/259 (92%)	0.12	3 (1%) 77 51	43, 80, 112, 139	0
2	H	240/259 (92%)	0.15	1 (0%) 92 79	55, 85, 117, 142	0
2	J	239/259 (92%)	0.25	8 (3%) 46 20	52, 85, 133, 140	0
2	L	240/259 (92%)	0.25	6 (2%) 57 29	63, 98, 139, 144	0
All	All	2871/3108 (92%)	0.24	79 (2%) 53 25	43, 88, 130, 213	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	109	LYS	6.3
1	C	85	ILE	5.5
1	I	67	GLU	4.3
1	I	134	LEU	4.3
1	K	66	GLY	4.3

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

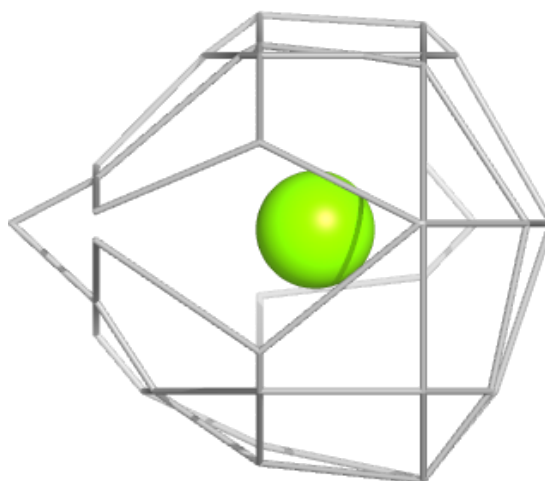
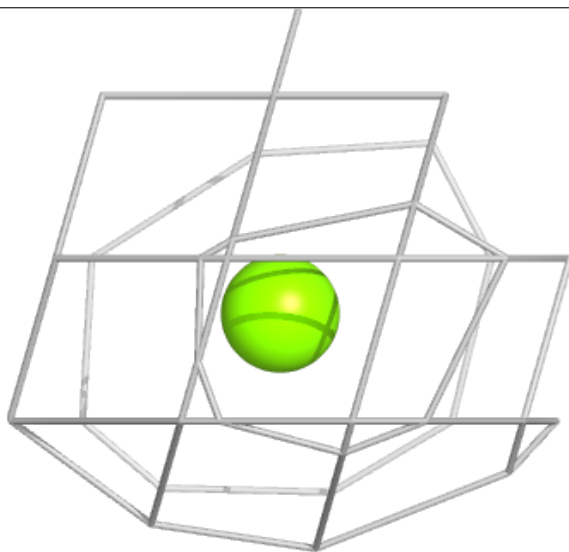
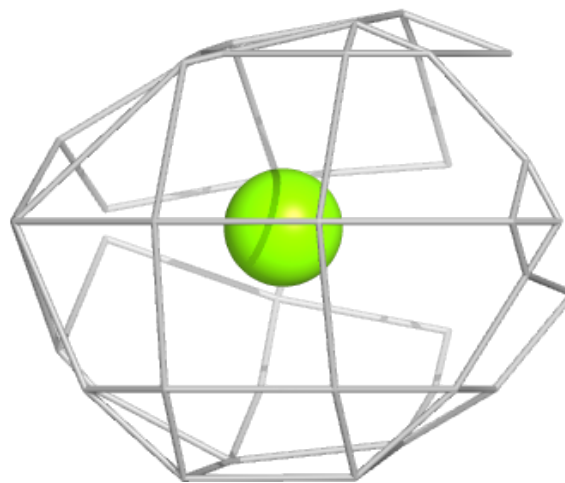
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	MG	F	601	1/1	0.90	0.14	52,52,52,52	0
3	MG	H	602	1/1	0.90	0.19	62,62,62,62	0
3	MG	H	601	1/1	0.91	0.19	60,60,60,60	0
3	MG	D	601	1/1	0.91	0.16	70,70,70,70	0
3	MG	L	601	1/1	0.91	0.17	78,78,78,78	0
3	MG	F	602	1/1	0.92	0.12	48,48,48,48	0
3	MG	J	601	1/1	0.92	0.11	71,71,71,71	0
3	MG	B	602	1/1	0.92	0.21	41,41,41,41	0
3	MG	K	401	1/1	0.93	0.20	62,62,62,62	0
3	MG	J	602	1/1	0.95	0.18	45,45,45,45	0
3	MG	D	602	1/1	0.95	0.35	41,41,41,41	0
3	MG	B	601	1/1	0.95	0.14	82,82,82,82	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.

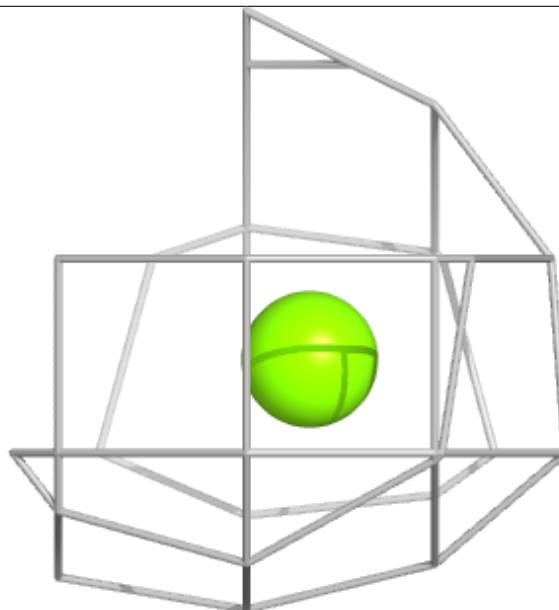
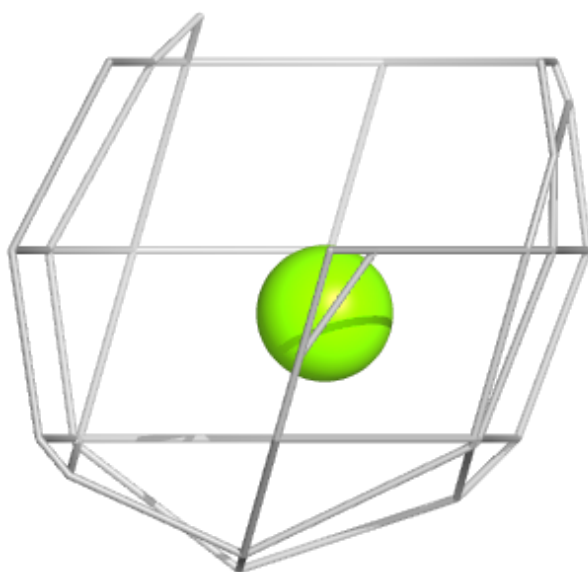
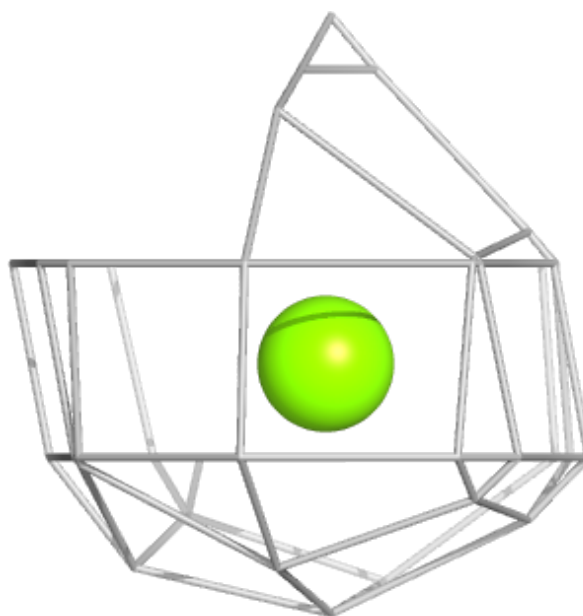
Electron density around MG F 601:

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



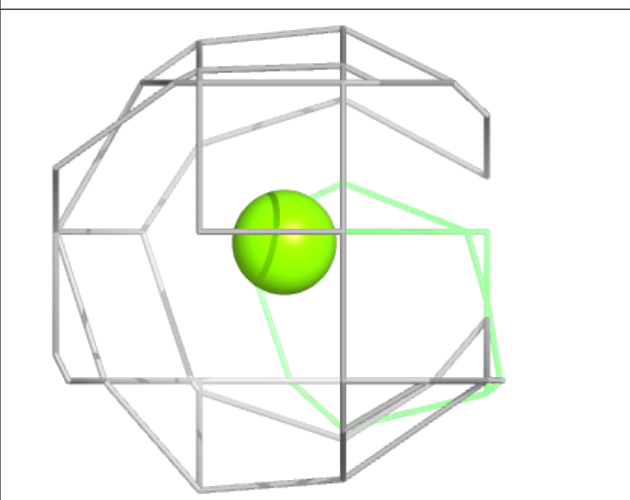
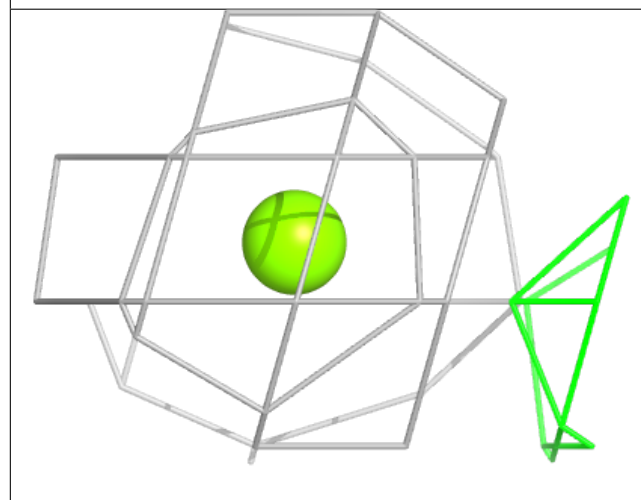
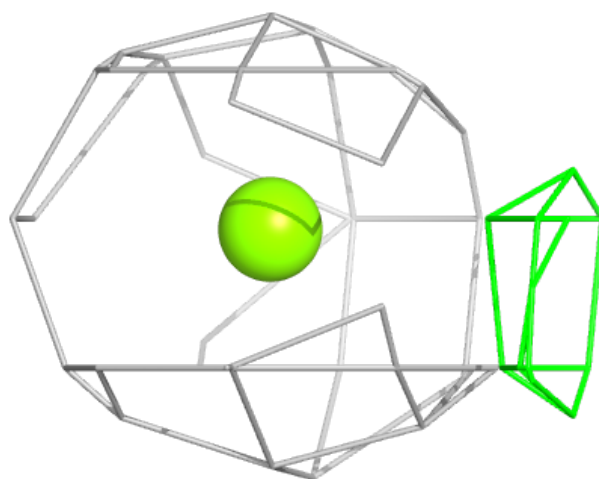
Electron density around MG H 602:

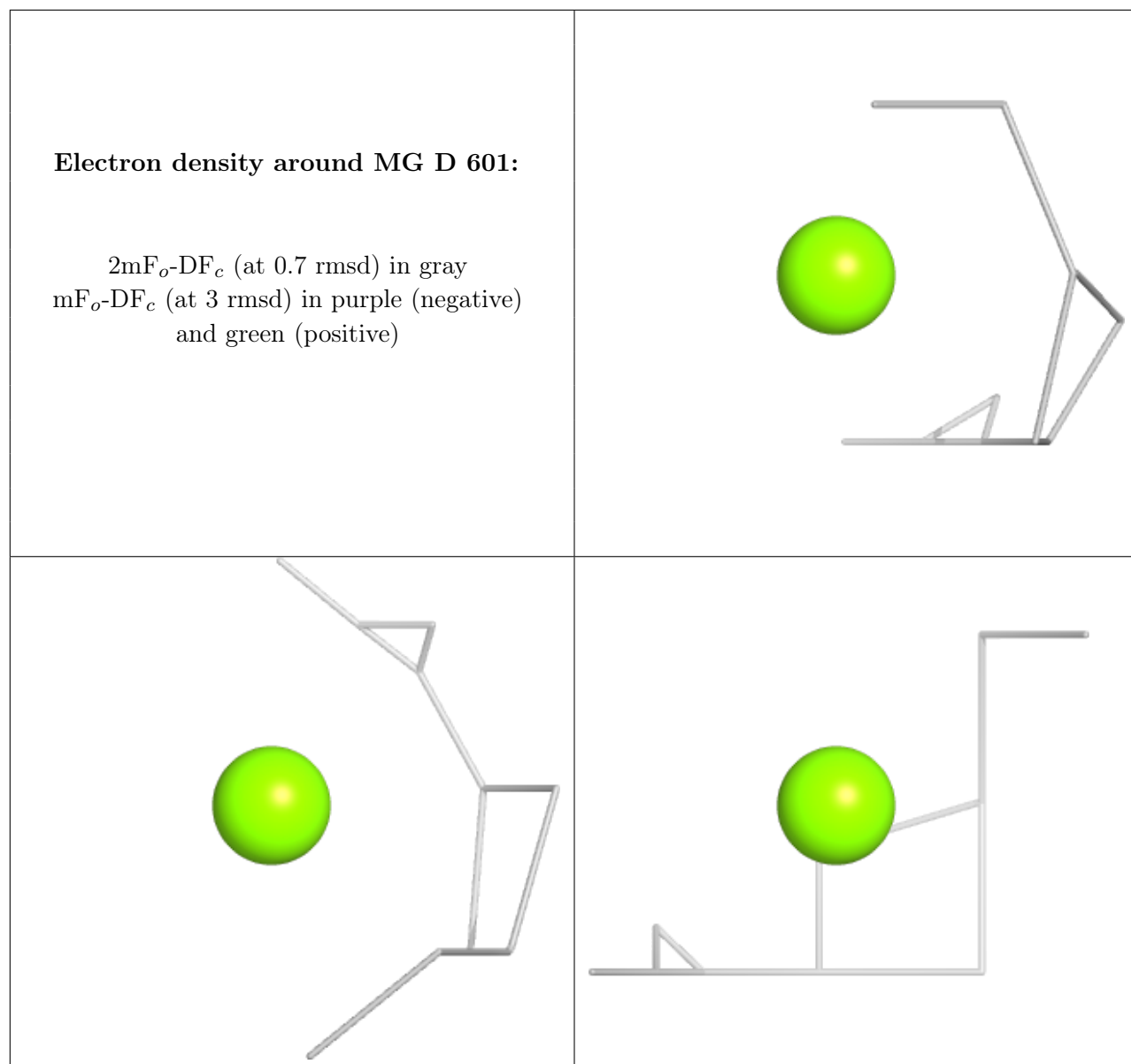
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MG H 601:

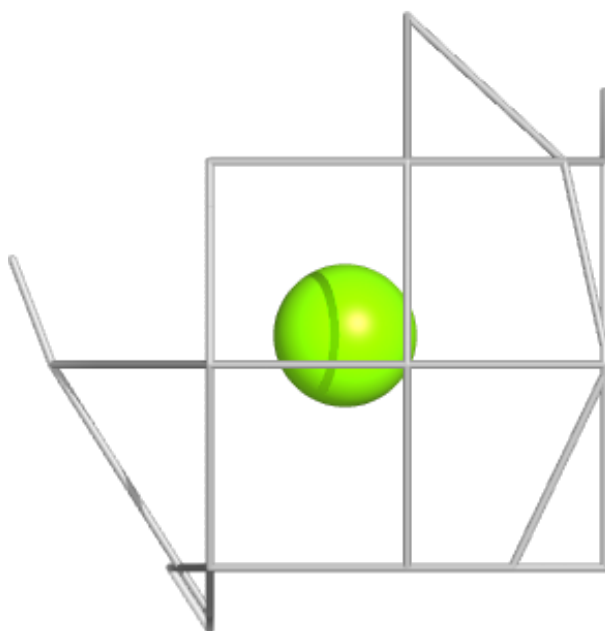
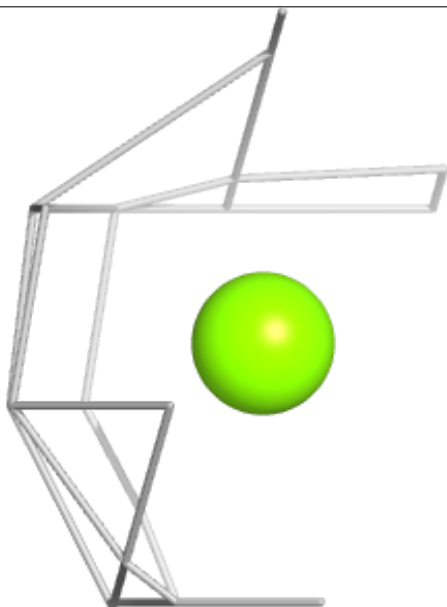
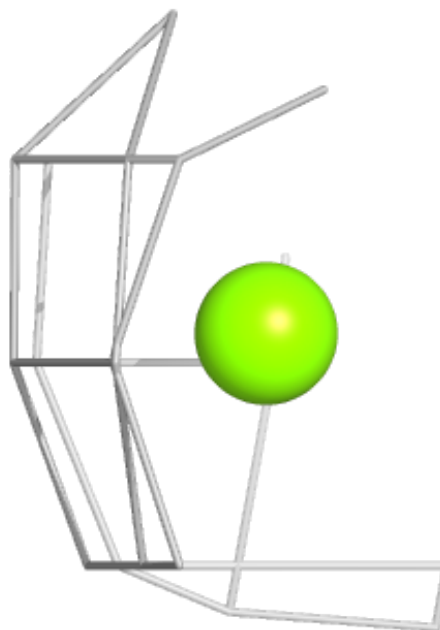
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





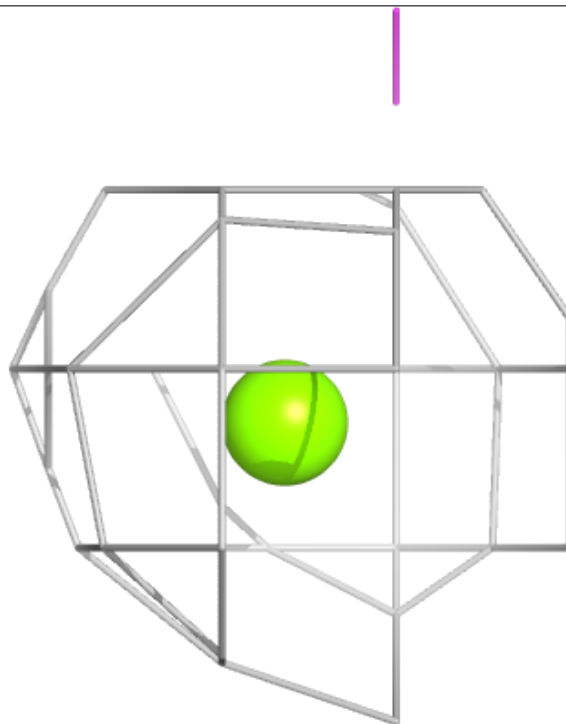
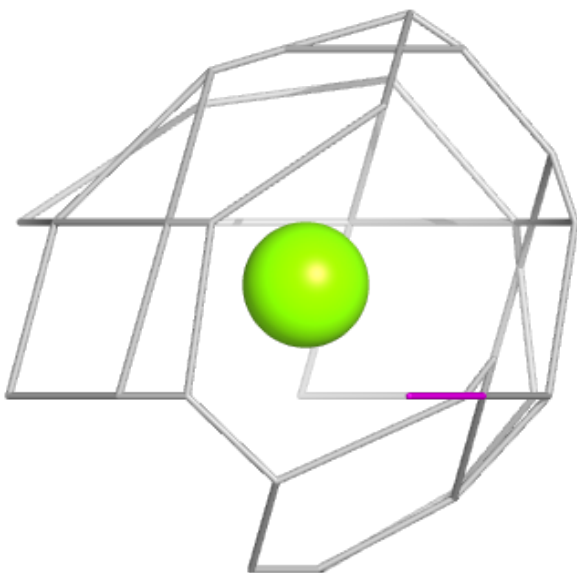
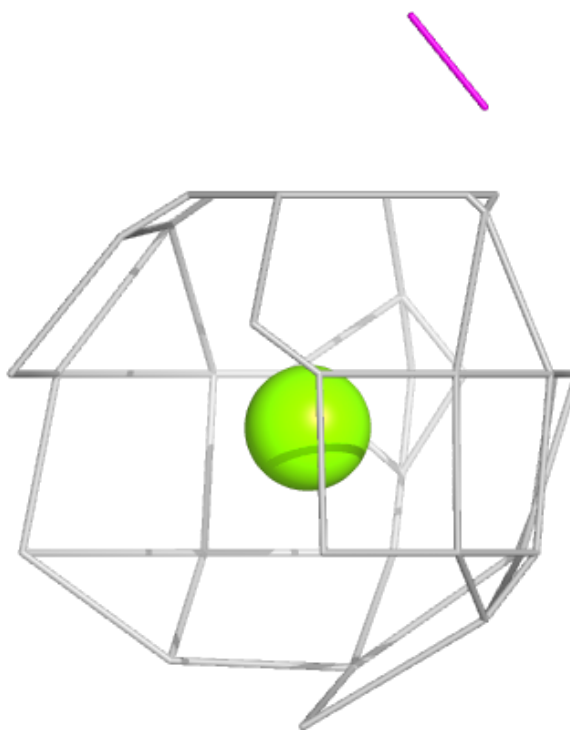
Electron density around MG L 601:

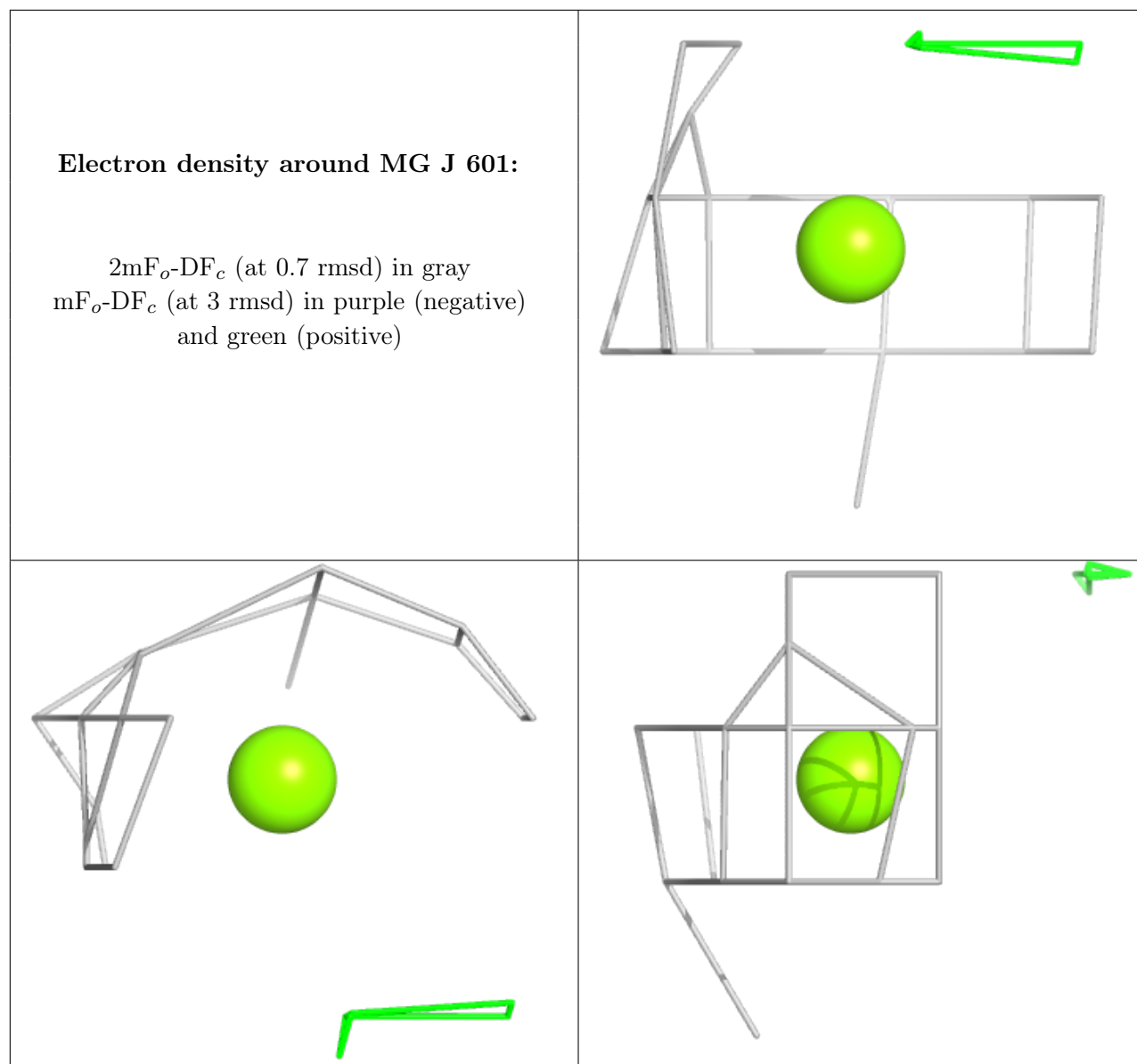
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MG F 602:

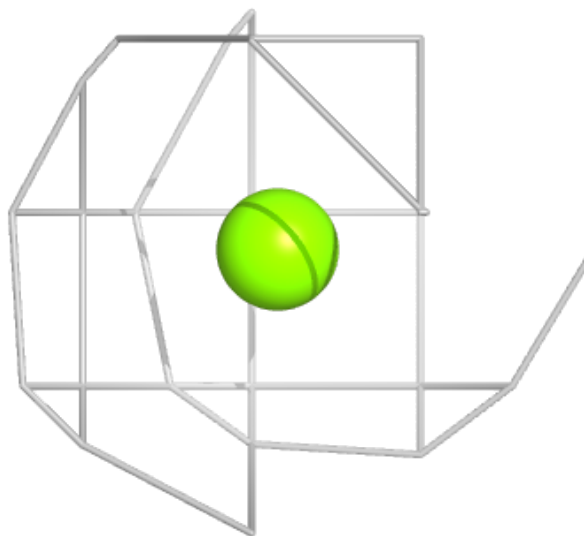
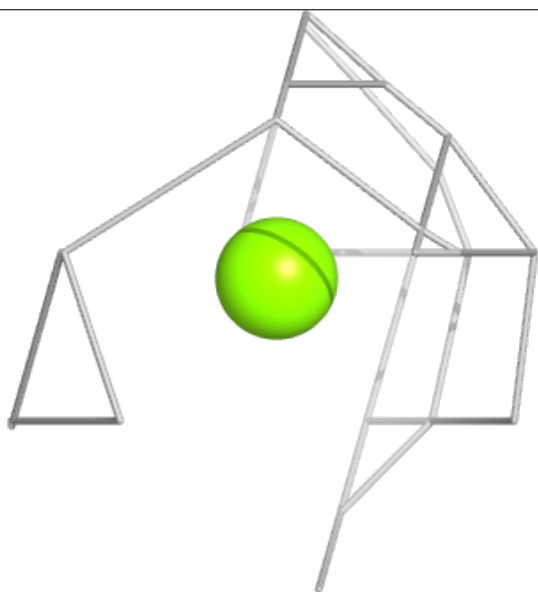
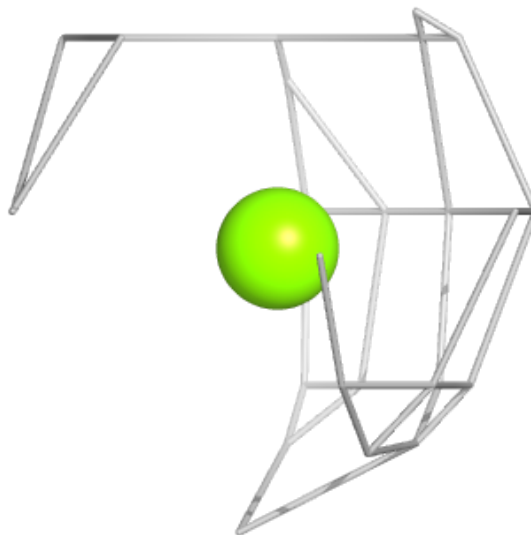
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





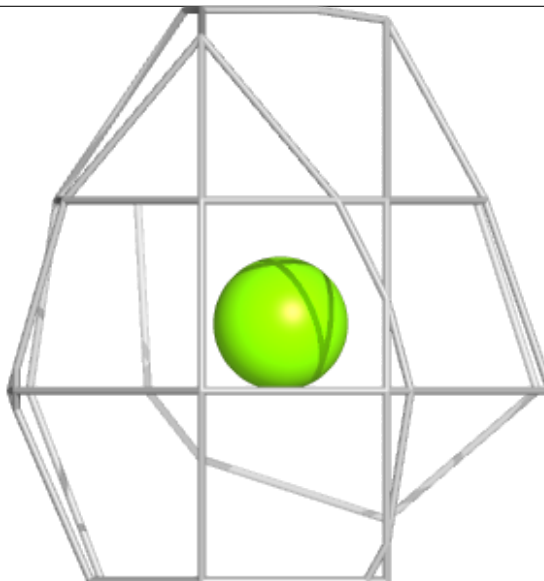
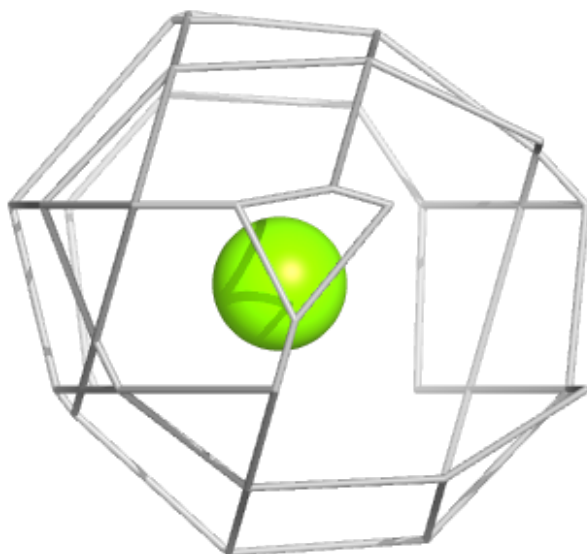
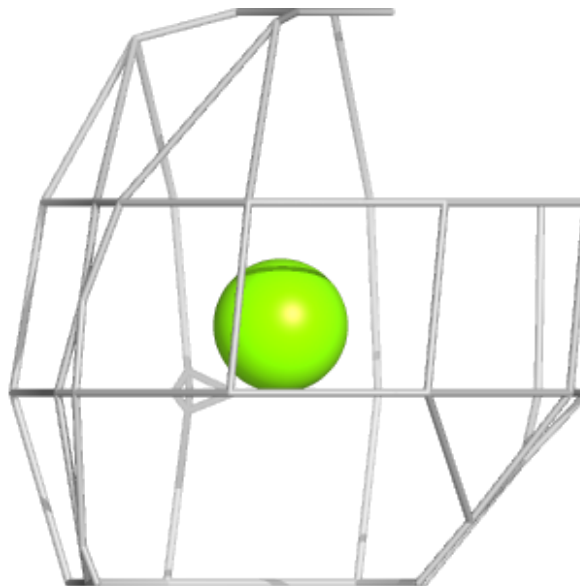
Electron density around MG B 602:

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



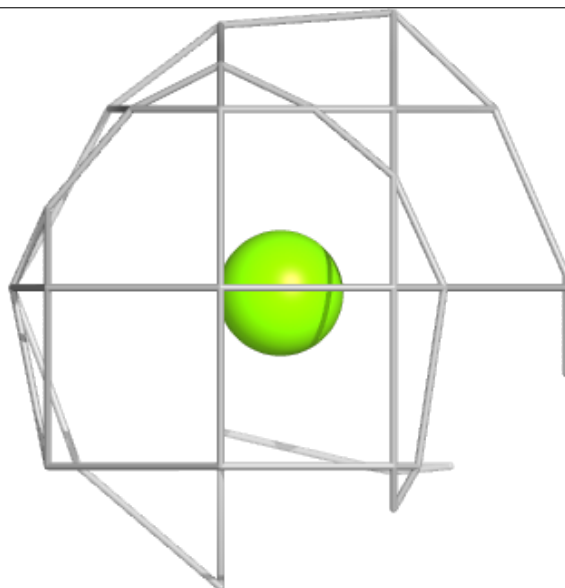
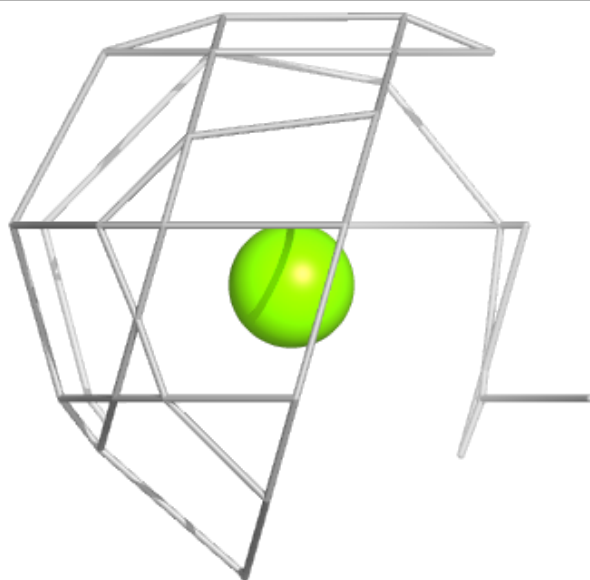
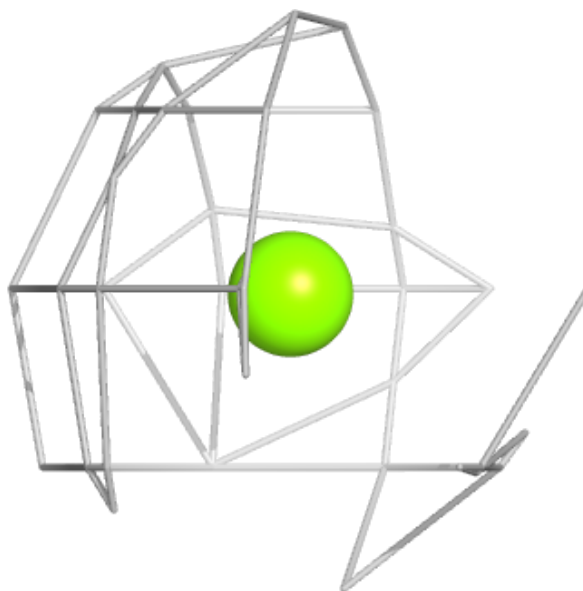
Electron density around MG K 401:

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



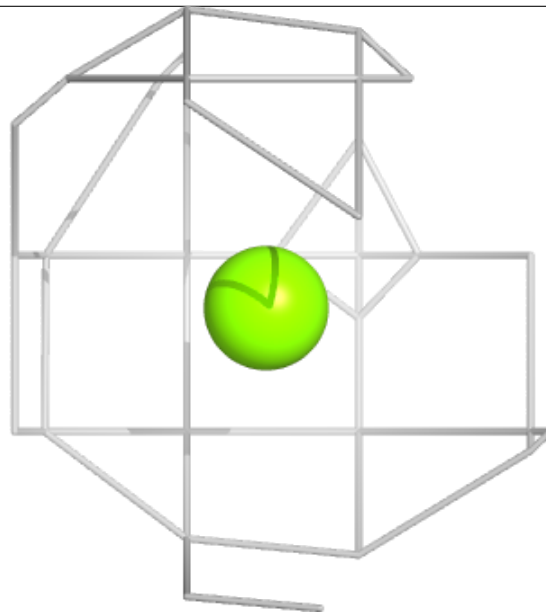
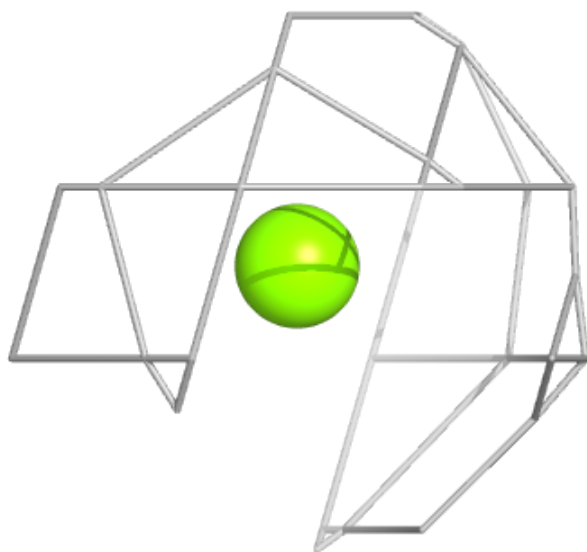
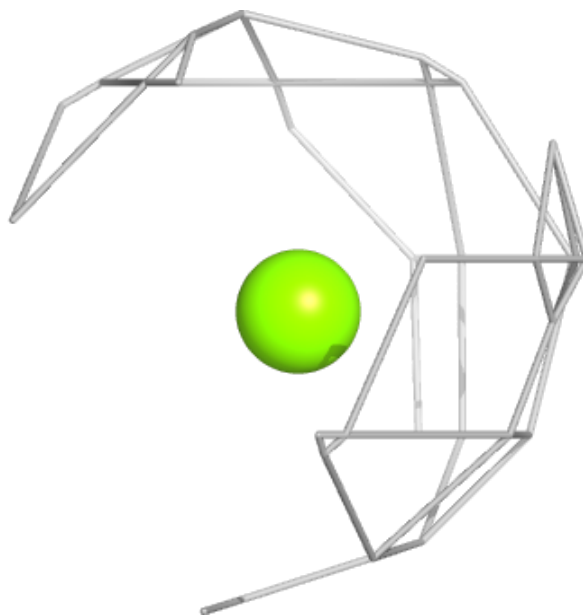
Electron density around MG J 602:

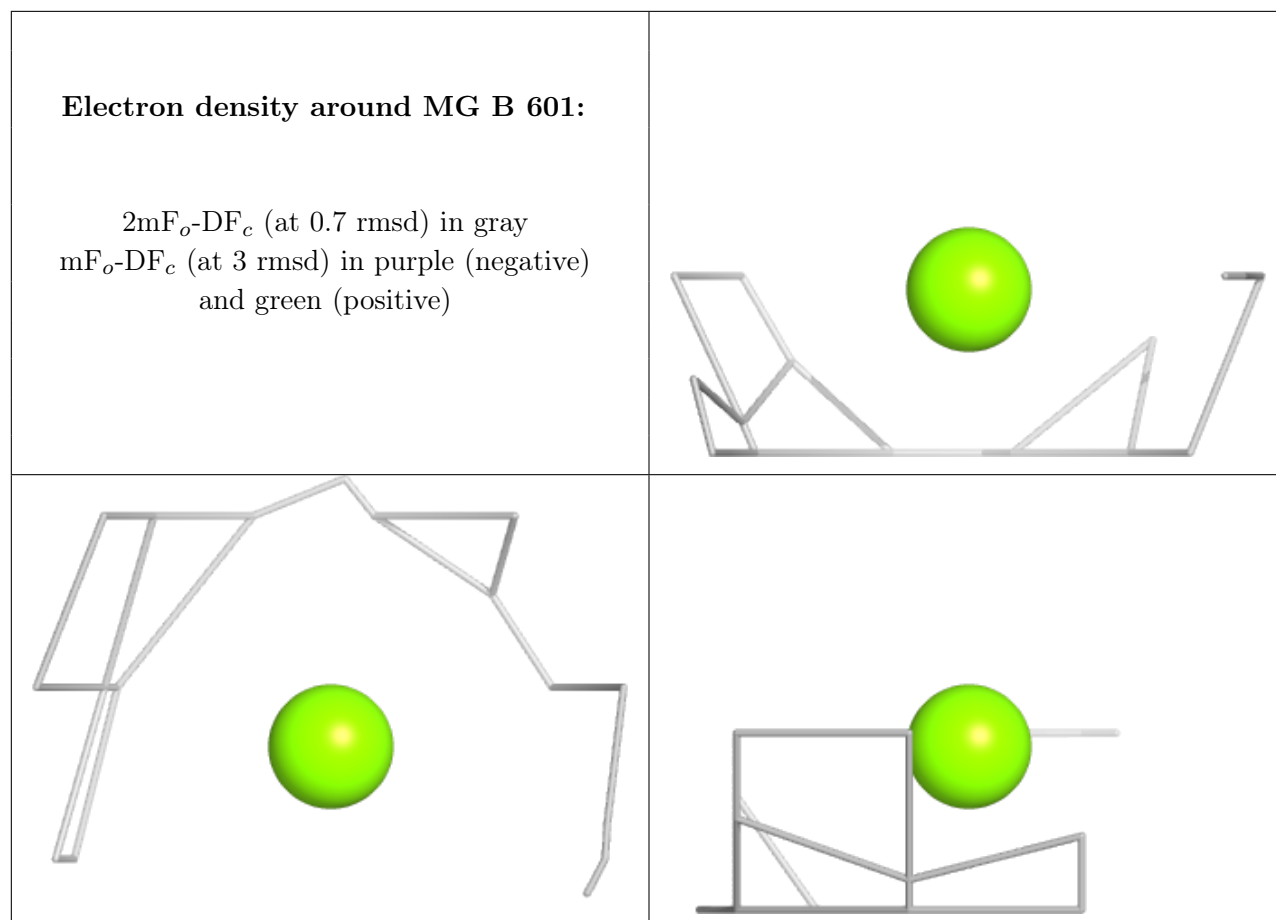
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MG D 602:

$2mF_o-DF_c$ (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.