



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 18, 2023 – 10:28 am BST

PDB ID : 8CM5
Title : W-formate dehydrogenase C872A from *Desulfovibrio vulgaris*
Authors : Vilela-Alves, G.; Mota, C.; Klymanska, K.; Oliveira, A.R.; Manuel, R.R.;
Pereira, I.C.; Romao, M.J.
Deposited on : 2023-02-17
Resolution : 2.15 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 types of validation reports are described at

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

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.35.1
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) : 2.35.1

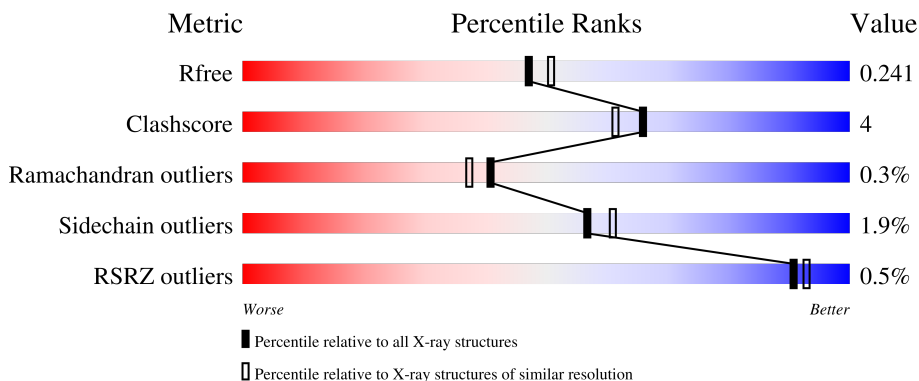
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



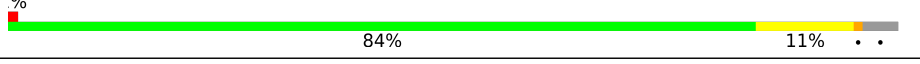

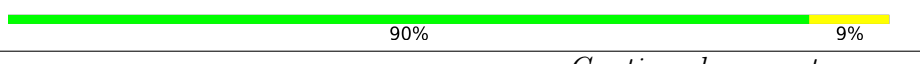
The reported resolution of this entry is 2.15 Å.

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.



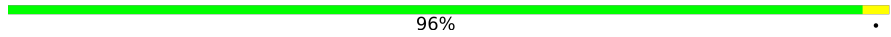
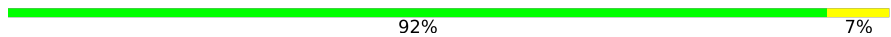
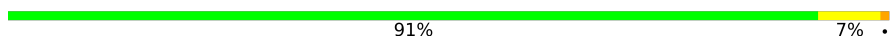
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	1013	 85% 11% .
1	C	1013	 88% 8% .
1	K	1013	 % 84% 11% . .
1	R	1013	 84% 12% .
2	B	215	 90% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	215	 96%
2	L	215	 92% 7%
2	S	215	 91% 7%

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	PEG	C	1101	-	-	X	-
3	PEG	D	401	-	-	X	-
3	PEG	L	301	-	-	X	-
3	PEG	L	302	-	-	X	-
3	PEG	S	302	-	-	X	-
8	GOL	S	301	-	-	X	-

2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 38677 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 Formate dehydrogenase, alpha subunit, selenocysteine-containing.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	974	7628	4863	1329	1395	40	1	0	0	0
1	C	974	7628	4863	1329	1395	40	1	0	0	0
1	K	973	7621	4858	1328	1394	40	1	0	0	0
1	R	973	7621	4858	1328	1394	40	1	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	872	ALA	CYS	engineered mutation	UNP Q72EJ1
A	1006	TRP	-	expression tag	UNP Q72EJ1
A	1007	SER	-	expression tag	UNP Q72EJ1
A	1008	HIS	-	expression tag	UNP Q72EJ1
A	1009	PRO	-	expression tag	UNP Q72EJ1
A	1010	GLN	-	expression tag	UNP Q72EJ1
A	1011	PHE	-	expression tag	UNP Q72EJ1
A	1012	GLU	-	expression tag	UNP Q72EJ1
A	1013	LYS	-	expression tag	UNP Q72EJ1
C	872	ALA	CYS	engineered mutation	UNP Q72EJ1
C	1006	TRP	-	expression tag	UNP Q72EJ1
C	1007	SER	-	expression tag	UNP Q72EJ1
C	1008	HIS	-	expression tag	UNP Q72EJ1
C	1009	PRO	-	expression tag	UNP Q72EJ1
C	1010	GLN	-	expression tag	UNP Q72EJ1
C	1011	PHE	-	expression tag	UNP Q72EJ1
C	1012	GLU	-	expression tag	UNP Q72EJ1
C	1013	LYS	-	expression tag	UNP Q72EJ1
K	872	ALA	CYS	engineered mutation	UNP Q72EJ1
K	1006	TRP	-	expression tag	UNP Q72EJ1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	1007	SER	-	expression tag	UNP Q72EJ1
K	1008	HIS	-	expression tag	UNP Q72EJ1
K	1009	PRO	-	expression tag	UNP Q72EJ1
K	1010	GLN	-	expression tag	UNP Q72EJ1
K	1011	PHE	-	expression tag	UNP Q72EJ1
K	1012	GLU	-	expression tag	UNP Q72EJ1
K	1013	LYS	-	expression tag	UNP Q72EJ1
R	872	ALA	CYS	engineered mutation	UNP Q72EJ1
R	1006	TRP	-	expression tag	UNP Q72EJ1
R	1007	SER	-	expression tag	UNP Q72EJ1
R	1008	HIS	-	expression tag	UNP Q72EJ1
R	1009	PRO	-	expression tag	UNP Q72EJ1
R	1010	GLN	-	expression tag	UNP Q72EJ1
R	1011	PHE	-	expression tag	UNP Q72EJ1
R	1012	GLU	-	expression tag	UNP Q72EJ1
R	1013	LYS	-	expression tag	UNP Q72EJ1

- Molecule 2 is a protein called Formate dehydrogenase, beta subunit, putative.

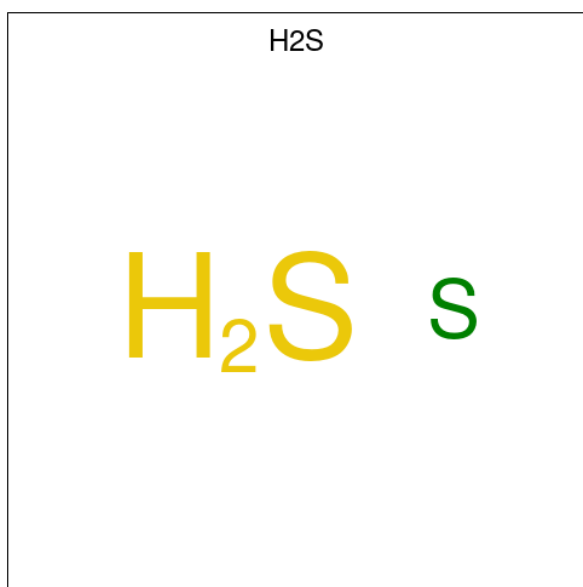
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	214	Total 1664	C 1041	N 291	O 316	S 16	0	0	0
2	D	214	Total 1664	C 1041	N 291	O 316	S 16	0	0	0
2	L	214	Total 1664	C 1041	N 291	O 316	S 16	0	0	0
2	S	214	Total 1664	C 1041	N 291	O 316	S 16	0	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 7 4 3	0	0
3	C	1	Total C O 7 4 3	0	0
3	D	1	Total C O 7 4 3	0	0
3	L	1	Total C O 7 4 3	0	0
3	L	1	Total C O 7 4 3	0	0
3	S	1	Total C O 7 4 3	0	0

- Molecule 4 is HYDROSULFURIC ACID (three-letter code: H2S) (formula: H₂S) (labeled as "Ligand of Interest" by depositor).

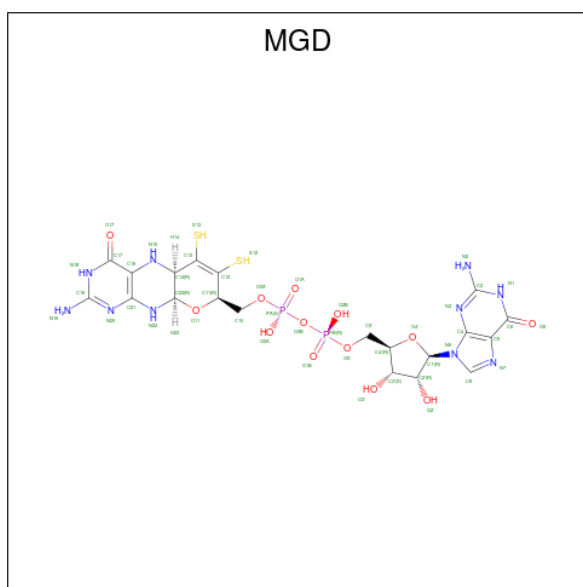


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total S 1 1	0	0
4	C	1	Total S 1 1	0	0
4	K	1	Total S 1 1	0	0
4	R	1	Total S 1 1	0	0

- Molecule 5 is TUNGSTEN ION (three-letter code: W) (formula: W) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total W 1 1	0	0
5	C	1	Total W 1 1	0	0
5	K	1	Total W 1 1	0	0
5	R	1	Total W 1 1	0	0

- Molecule 6 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: C₂₀H₂₆N₁₀O₁₃P₂S₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
6	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	K	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	K	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	R	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
6	R	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		

- Molecule 7 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄) (labeled as "Ligand of Interest" by depositor).



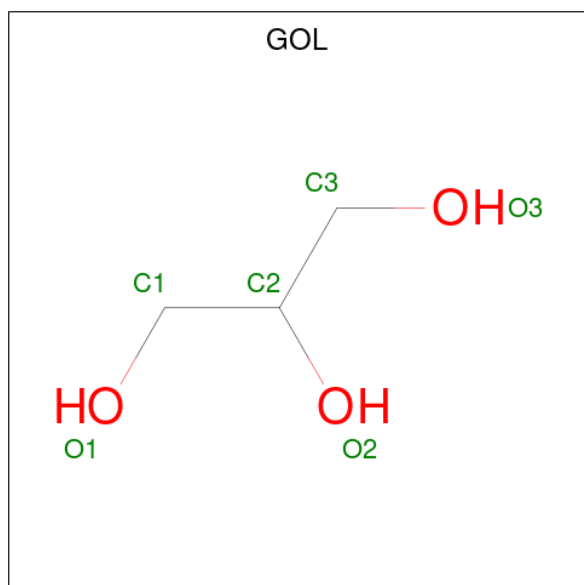
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	Fe	S	0	0
			8	4	4		
7	B	1	Total	Fe	S	0	0
			8	4	4		
7	B	1	Total	Fe	S	0	0
			8	4	4		
7	B	1	Total	Fe	S	0	0
			8	4	4		
7	C	1	Total	Fe	S	0	0
			8	4	4		
7	D	1	Total	Fe	S	0	0
			8	4	4		
7	D	1	Total	Fe	S	0	0
			8	4	4		
7	D	1	Total	Fe	S	0	0
			8	4	4		
7	K	1	Total	Fe	S	0	0
			8	4	4		
7	L	1	Total	Fe	S	0	0
			8	4	4		
7	L	1	Total	Fe	S	0	0
			8	4	4		
7	L	1	Total	Fe	S	0	0
			8	4	4		
7	R	1	Total	Fe	S	0	0
			8	4	4		
7	S	1	Total	Fe	S	0	0
			8	4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	S	1	Total	Fe	S	0	0
			8	4	4		
7	S	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	S	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	194	Total	O	0	0
			194	194		
9	B	99	Total	O	0	0
			99	99		
9	C	171	Total	O	0	0
			171	171		
9	D	90	Total	O	0	0
			90	90		
9	K	114	Total	O	0	0
			114	114		
9	L	56	Total	O	0	0
			56	56		

Continued on next page...

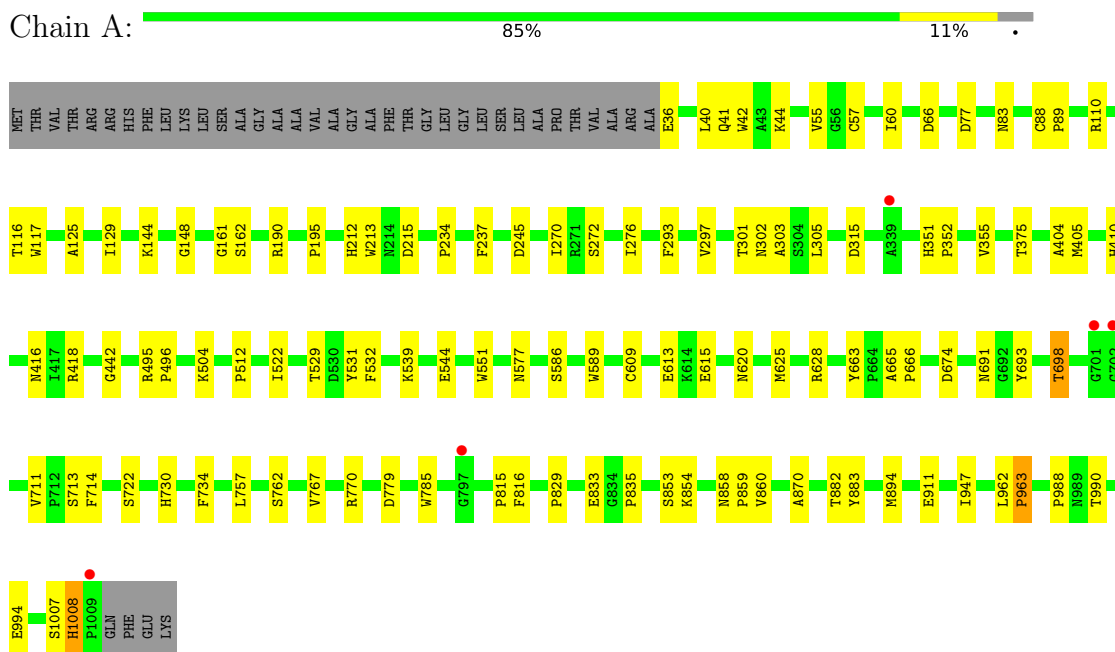
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	R	154	Total 154	O 154	0	0
9	S	85	Total 85	O 85	0	0

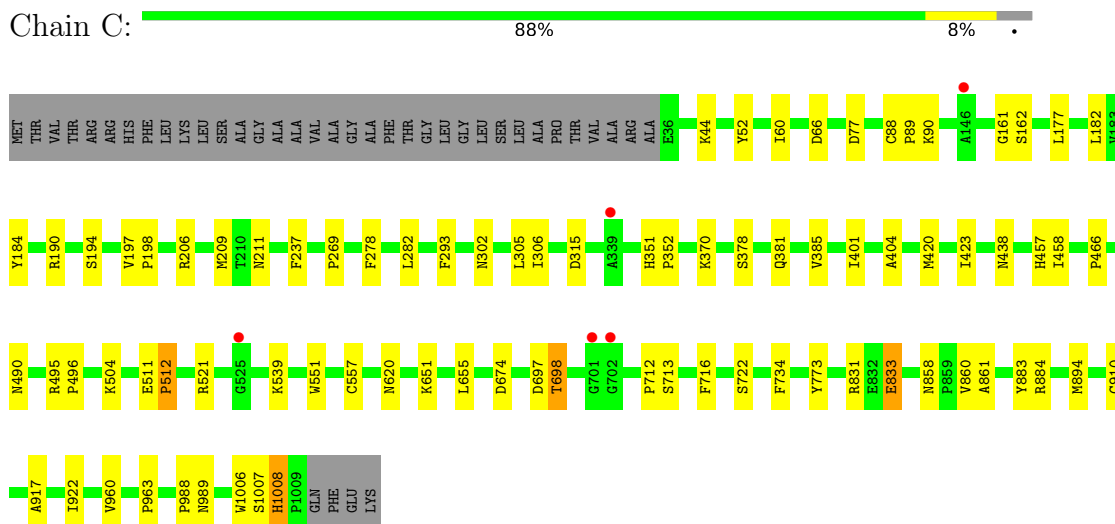
3 Residue-property plots

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.

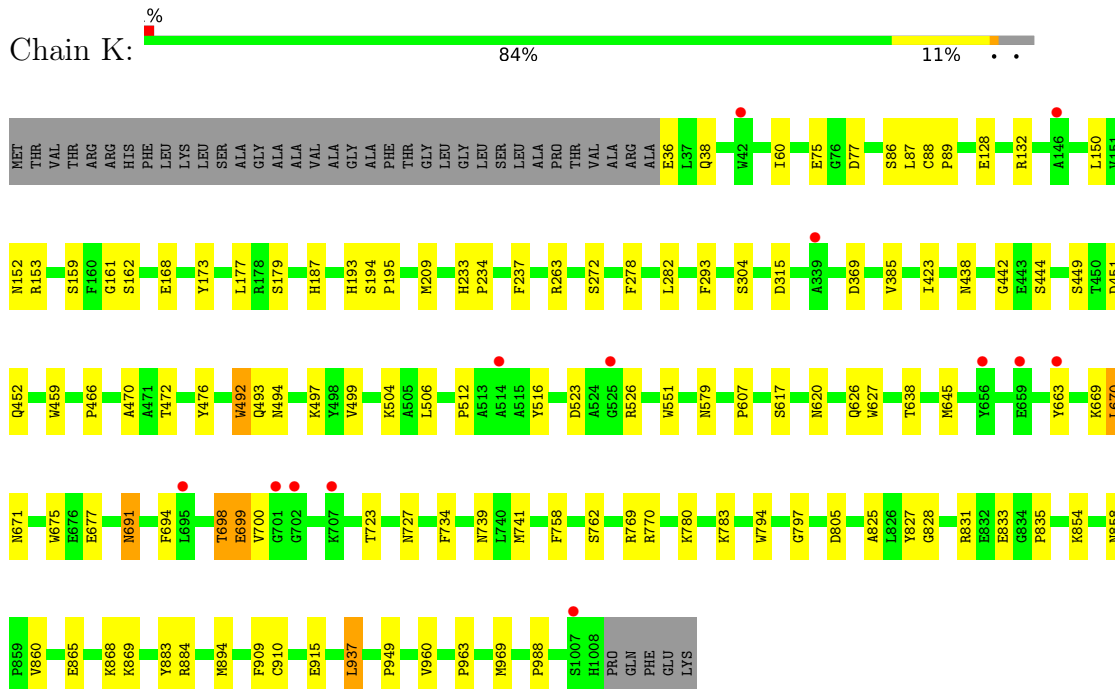
- Molecule 1: Formate dehydrogenase, alpha subunit, selenocysteine-containing



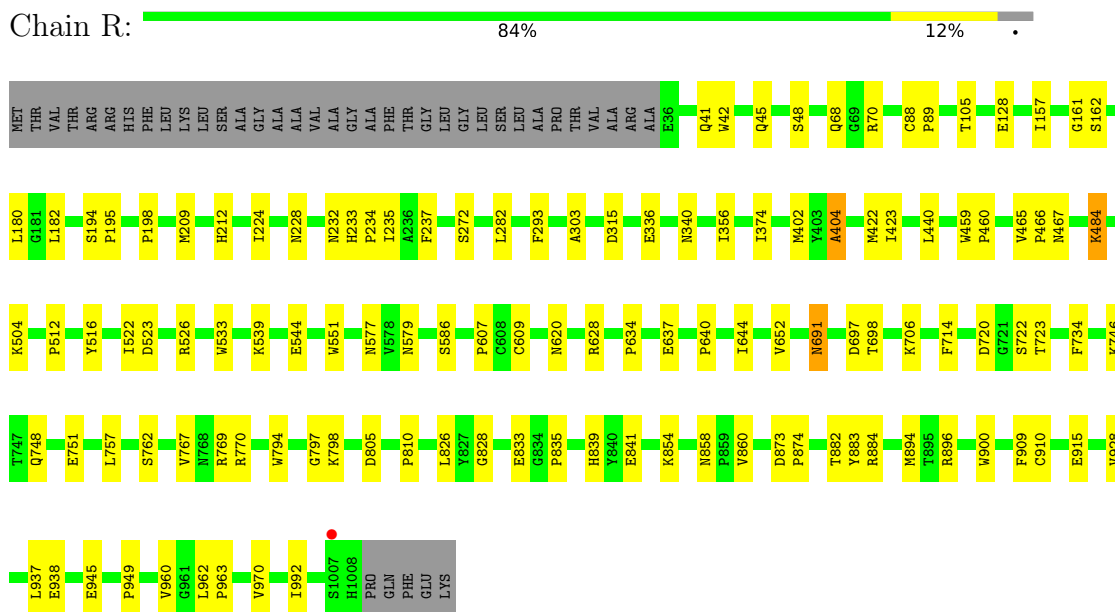
- Molecule 1: Formate dehydrogenase, alpha subunit, selenocysteine-containing



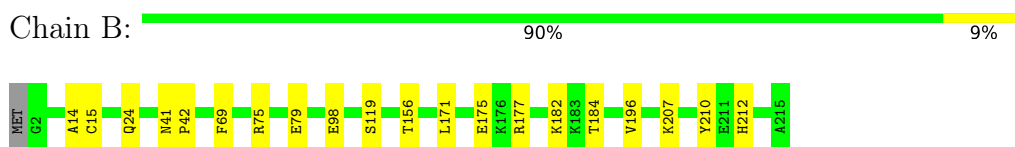
- Molecule 1: Formate dehydrogenase, alpha subunit, selenocysteine-containing



- Molecule 1: Formate dehydrogenase, alpha subunit, selenocysteine-containing



- Molecule 2: Formate dehydrogenase, beta subunit, putative



- Molecule 2: Formate dehydrogenase, beta subunit, putative

Chain D:  96%




- Molecule 2: Formate dehydrogenase, beta subunit, putative

Chain L:  92%



- Molecule 2: Formate dehydrogenase, beta subunit, putative

Chain S:  91%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.44Å 261.06Å 126.21Å 90.00° 90.58° 90.00°	Depositor
Resolution (Å)	48.83 – 2.15 48.78 – 2.15	Depositor EDS
% Data completeness (in resolution range)	96.8 (48.83-2.15) 96.8 (48.78-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.85 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.184 , 0.240 0.189 , 0.241	Depositor DCC
R_{free} test set	12240 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	27.5	Xtrriage
Anisotropy	0.106	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 13.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.247 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	38677	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.43% of the height of the origin peak. No significant pseudotranslation is detected.*

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: PEG, SF4, MGD, W, GOL, SEC, H2S

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	0.74	0/7834	0.89	5/10629 (0.0%)
1	C	0.74	0/7834	0.88	2/10629 (0.0%)
1	K	0.75	0/7826	0.89	2/10617 (0.0%)
1	R	0.74	1/7826 (0.0%)	0.89	1/10617 (0.0%)
2	B	0.78	1/1699 (0.1%)	0.91	0/2302
2	D	0.76	0/1699	0.93	1/2302 (0.0%)
2	L	0.78	0/1699	0.95	2/2302 (0.1%)
2	S	0.81	0/1699	0.94	2/2302 (0.1%)
All	All	0.75	2/38116 (0.0%)	0.90	15/51700 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	C	0	1
1	R	0	1
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	79	GLU	CD-OE2	5.34	1.31	1.25
1	R	945	GLU	CD-OE1	5.29	1.31	1.25

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	S	143	ASP	CB-CG-OD1	7.13	124.72	118.30
2	L	143	ASP	CB-CG-OD1	6.39	124.05	118.30
1	A	190	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	A	190	ARG	NE-CZ-NH2	-5.93	117.33	120.30
1	C	833	GLU	CB-CA-C	-5.77	98.86	110.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1007	SER	Peptide
1	A	40	LEU	Peptide
1	C	1007	SER	Peptide
1	R	723	THR	Peptide

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	7628	0	7454	65	0
1	C	7628	0	7454	44	0
1	K	7621	0	7447	64	0
1	R	7621	0	7447	74	0
2	B	1664	0	1633	8	0
2	D	1664	0	1633	5	0
2	L	1664	0	1633	18	0
2	S	1664	0	1633	23	0
3	A	7	0	10	0	0
3	C	7	0	10	4	0
3	D	7	0	9	4	0
3	L	14	0	18	13	0
3	S	7	0	9	13	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	K	1	0	0	0	0
4	R	1	0	0	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	K	1	0	0	0	0
5	R	1	0	0	0	0
6	A	94	0	44	4	0
6	C	94	0	44	2	0
6	K	94	0	44	4	0
6	R	94	0	44	5	0
7	A	8	0	0	1	0
7	B	24	0	0	0	0
7	C	8	0	0	0	0
7	D	24	0	0	0	0
7	K	8	0	0	0	0
7	L	24	0	0	0	0
7	R	8	0	0	0	0
7	S	24	0	0	0	0
8	S	6	0	8	4	0
9	A	194	0	0	3	0
9	B	99	0	0	1	0
9	C	171	0	0	3	0
9	D	90	0	0	0	0
9	K	114	0	0	2	0
9	L	56	0	0	2	0
9	R	154	0	0	2	0
9	S	85	0	0	0	0
All	All	38677	0	36574	299	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 299 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)
2:L:125:ILE:HG23	3:L:301:PEG:H21	1.42	0.97
2:L:137:LYS:HA	3:L:301:PEG:H32	1.47	0.96
1:R:224:ILE:HG12	1:R:402:MET:HE2	1.49	0.91
2:S:137:LYS:HA	3:S:302:PEG:H31	1.50	0.91
1:R:224:ILE:HG12	1:R:402:MET:CE	2.00	0.90

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	971/1013 (96%)	922 (95%)	45 (5%)	4 (0%)	34	29
1	C	971/1013 (96%)	934 (96%)	34 (4%)	3 (0%)	41	37
1	K	970/1013 (96%)	906 (93%)	58 (6%)	6 (1%)	25	18
1	R	970/1013 (96%)	926 (96%)	41 (4%)	3 (0%)	41	37
2	B	212/215 (99%)	205 (97%)	7 (3%)	0	100	100
2	D	212/215 (99%)	201 (95%)	11 (5%)	0	100	100
2	L	212/215 (99%)	199 (94%)	13 (6%)	0	100	100
2	S	212/215 (99%)	200 (94%)	12 (6%)	0	100	100
All	All	4730/4912 (96%)	4493 (95%)	221 (5%)	16 (0%)	41	37

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1008	HIS
1	C	1008	HIS
1	A	404	ALA
1	C	404	ALA
1	C	490	ASN

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	791/818 (97%)	780 (99%)	11 (1%)	67	72

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	791/818 (97%)	778 (98%)	13 (2%)	62	67
1	K	790/818 (97%)	769 (97%)	21 (3%)	44	46
1	R	790/818 (97%)	774 (98%)	16 (2%)	55	59
2	B	185/186 (100%)	179 (97%)	6 (3%)	39	38
2	D	185/186 (100%)	183 (99%)	2 (1%)	73	78
2	L	185/186 (100%)	181 (98%)	4 (2%)	52	55
2	S	185/186 (100%)	182 (98%)	3 (2%)	62	67
All	All	3902/4016 (97%)	3826 (98%)	76 (2%)	57	61

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	R	45	GLN
1	R	854	LYS
1	R	237	PHE
1	R	465	VAL
2	S	167	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	748	GLN
1	R	325	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 39 ligands modelled in this entry, 4 are modelled with single atom and 4 are monoatomic - leaving 31 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
3	PEG	D	401	-	6,6,6	0.75	0	5,5,5	0.73	0
7	SF4	L	304	2	0,12,12	-	-	-		
7	SF4	K	1105	1	0,12,12	-	-	-		
7	SF4	R	1105	1	0,12,12	-	-	-		
3	PEG	L	301	-	6,6,6	0.74	0	5,5,5	0.30	0
7	SF4	C	1106	1	0,12,12	-	-	-		
7	SF4	A	1106	1	0,12,12	-	-	-		
6	MGD	A	1105	5	41,52,52	1.01	4 (9%)	40,81,81	1.31	5 (12%)
7	SF4	D	402	2	0,12,12	-	-	-		
7	SF4	B	302	2	0,12,12	-	-	-		
7	SF4	S	303	2	0,12,12	-	-	-		
6	MGD	C	1104	5	41,52,52	0.90	2 (4%)	40,81,81	1.26	4 (10%)
8	GOL	S	301	-	5,5,5	0.54	0	5,5,5	1.13	0
6	MGD	C	1105	5	41,52,52	1.69	3 (7%)	40,81,81	1.49	3 (7%)
7	SF4	S	304	2	0,12,12	-	-	-		
7	SF4	D	403	2	0,12,12	-	-	-		
6	MGD	K	1104	5	41,52,52	1.04	3 (7%)	40,81,81	1.30	2 (5%)
7	SF4	B	301	2	0,12,12	-	-	-		
7	SF4	D	404	2	0,12,12	-	-	-		
7	SF4	B	303	2	0,12,12	-	-	-		
3	PEG	C	1101	-	6,6,6	0.40	0	5,5,5	0.26	0
7	SF4	L	305	2	0,12,12	-	-	-		
3	PEG	L	302	-	6,6,6	1.16	1 (16%)	5,5,5	0.33	0
6	MGD	R	1104	5	41,52,52	0.98	2 (4%)	40,81,81	1.42	6 (15%)
7	SF4	L	303	2	0,12,12	-	-	-		
3	PEG	A	1101	-	6,6,6	0.21	0	5,5,5	0.14	0
7	SF4	S	305	2	0,12,12	-	-	-		
6	MGD	A	1104	5	41,52,52	1.03	5 (12%)	40,81,81	1.39	5 (12%)
6	MGD	R	1103	5	41,52,52	0.86	3 (7%)	40,81,81	1.11	3 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
6	MGD	K	1103	5	41,52,52	1.27	3 (7%)	40,81,81	1.28	6 (15%)
3	PEG	S	302	-	6,6,6	0.81	0	5,5,5	0.46	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
3	PEG	D	401	-	-	2/4/4/4	-
7	SF4	L	304	2	-	-	0/6/5/5
7	SF4	K	1105	1	-	-	0/6/5/5
7	SF4	R	1105	1	-	-	0/6/5/5
3	PEG	L	301	-	-	4/4/4/4	-
7	SF4	C	1106	1	-	-	0/6/5/5
7	SF4	A	1106	1	-	-	0/6/5/5
6	MGD	A	1105	5	-	5/18/66/66	0/6/6/6
7	SF4	D	402	2	-	-	0/6/5/5
7	SF4	B	302	2	-	-	0/6/5/5
7	SF4	S	303	2	-	-	0/6/5/5
6	MGD	C	1104	5	-	2/18/66/66	0/6/6/6
8	GOL	S	301	-	-	4/4/4/4	-
6	MGD	C	1105	5	-	6/18/66/66	0/6/6/6
7	SF4	S	304	2	-	-	0/6/5/5
7	SF4	D	403	2	-	-	0/6/5/5
6	MGD	K	1104	5	-	6/18/66/66	0/6/6/6
7	SF4	B	301	2	-	-	0/6/5/5
7	SF4	D	404	2	-	-	0/6/5/5
7	SF4	B	303	2	-	-	0/6/5/5
3	PEG	C	1101	-	-	2/4/4/4	-
7	SF4	L	305	2	-	-	0/6/5/5
3	PEG	L	302	-	-	3/4/4/4	-
6	MGD	R	1104	5	-	6/18/66/66	0/6/6/6
7	SF4	L	303	2	-	-	0/6/5/5
3	PEG	A	1101	-	-	4/4/4/4	-
7	SF4	S	305	2	-	-	0/6/5/5
6	MGD	A	1104	5	-	3/18/66/66	0/6/6/6
6	MGD	R	1103	5	-	1/18/66/66	0/6/6/6
6	MGD	K	1103	5	-	2/18/66/66	0/6/6/6

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PEG	S	302	-	-	2/4/4/4	-

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1105	MGD	C23-C14	9.12	1.60	1.53
6	K	1103	MGD	C23-C14	6.13	1.58	1.53
6	K	1104	MGD	C23-C14	3.62	1.56	1.53
6	A	1105	MGD	C23-C14	3.40	1.56	1.53
6	A	1104	MGD	C5-C6	-2.86	1.41	1.47

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	1105	MGD	O11-C23-C14	5.56	112.67	108.96
6	K	1104	MGD	O4'-C1'-C2'	-4.48	100.38	106.93
6	A	1104	MGD	O11-C23-C14	4.45	111.93	108.96
6	R	1104	MGD	O4'-C1'-C2'	-3.96	101.14	106.93
6	R	1104	MGD	O11-C23-C14	3.83	111.52	108.96

There are no chirality outliers.

5 of 52 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1105	MGD	C5'-O5'-PB-O1B
6	A	1105	MGD	C5'-O5'-PB-O2B
6	A	1105	MGD	C4'-C5'-O5'-PB
6	C	1105	MGD	C5'-O5'-PB-O2B
6	C	1105	MGD	C4'-C5'-O5'-PB

There are no ring outliers.

14 monomers are involved in 54 short contacts:

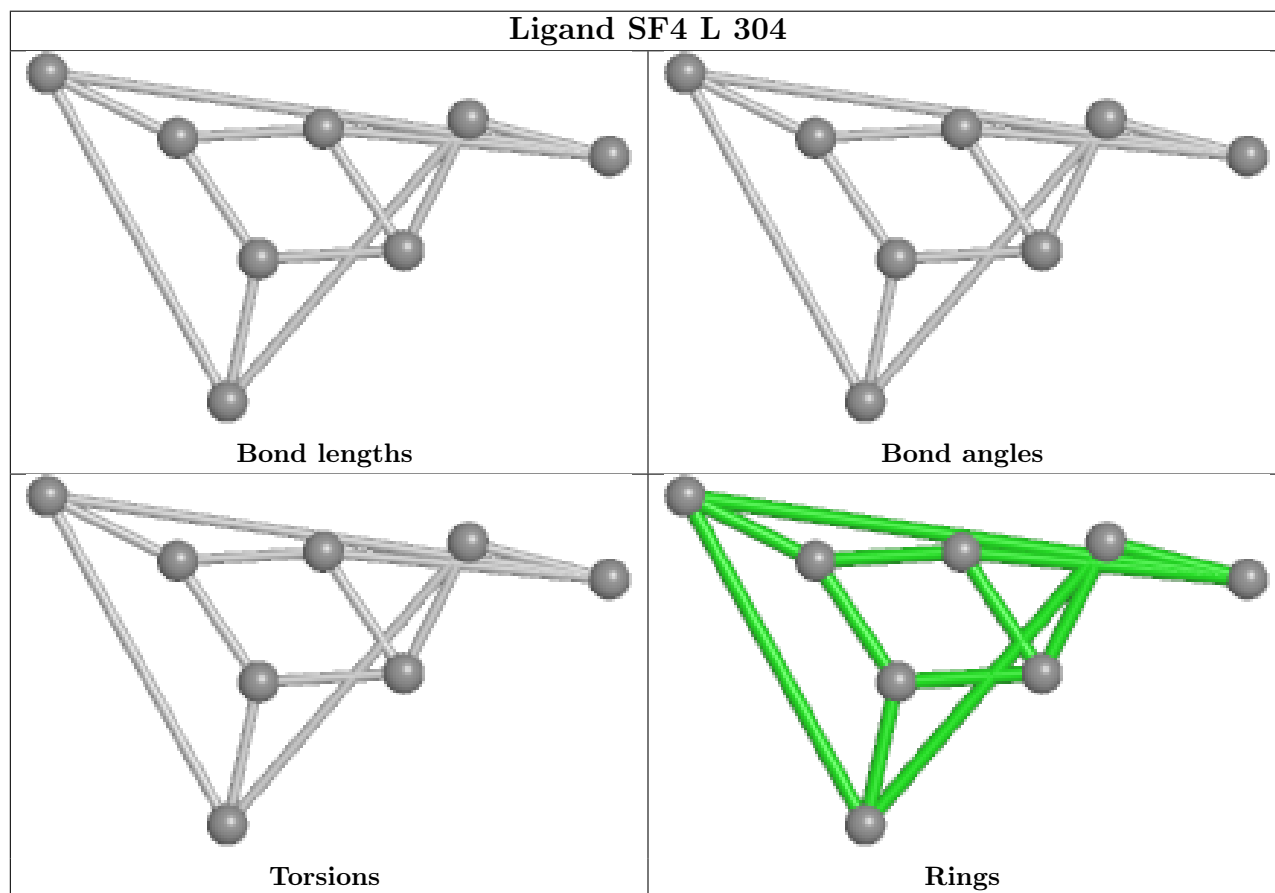
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	401	PEG	4	0
3	L	301	PEG	9	0
7	A	1106	SF4	1	0
6	A	1105	MGD	1	0
8	S	301	GOL	4	0
6	C	1105	MGD	2	0
6	K	1104	MGD	1	0

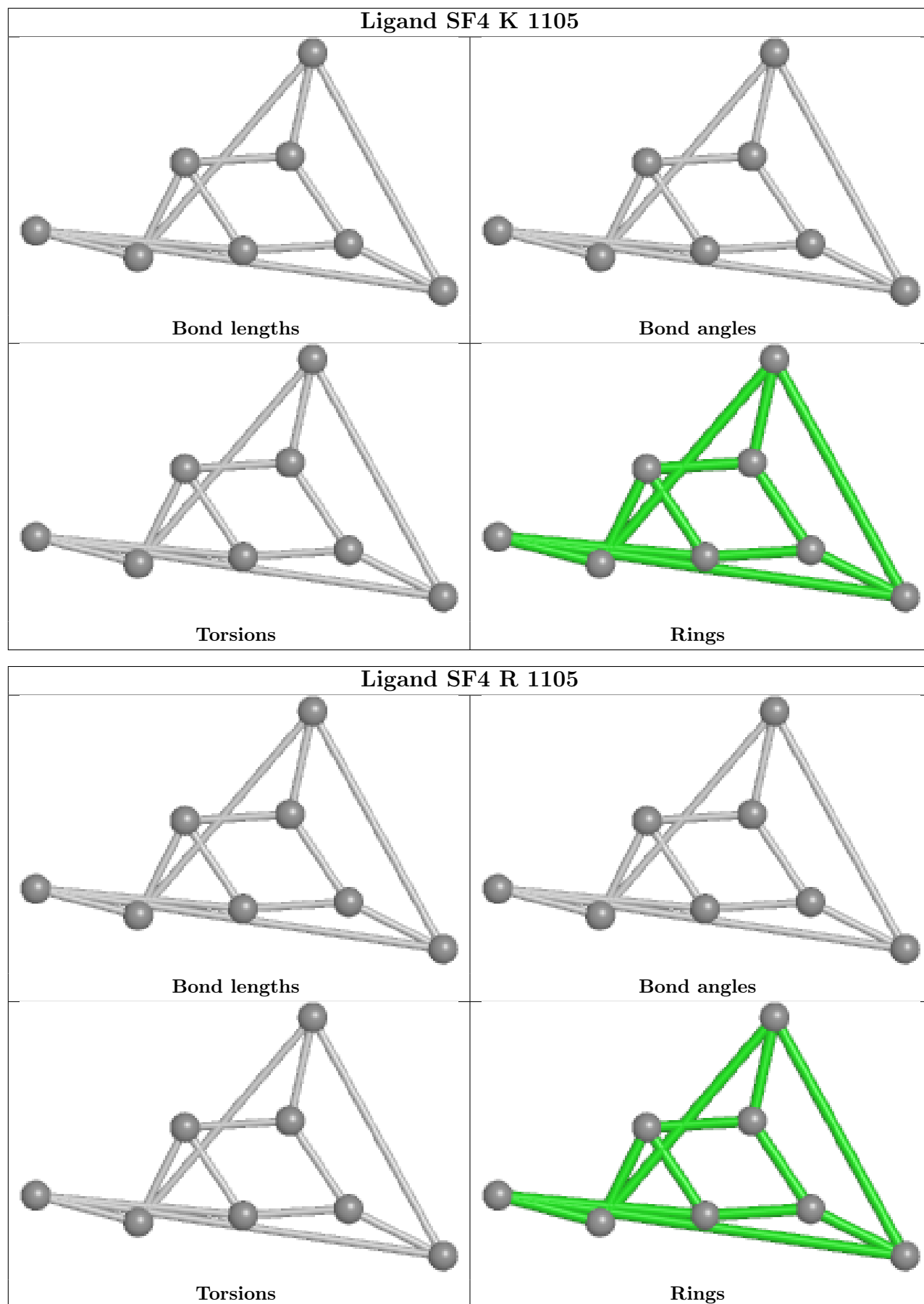
Continued on next page...

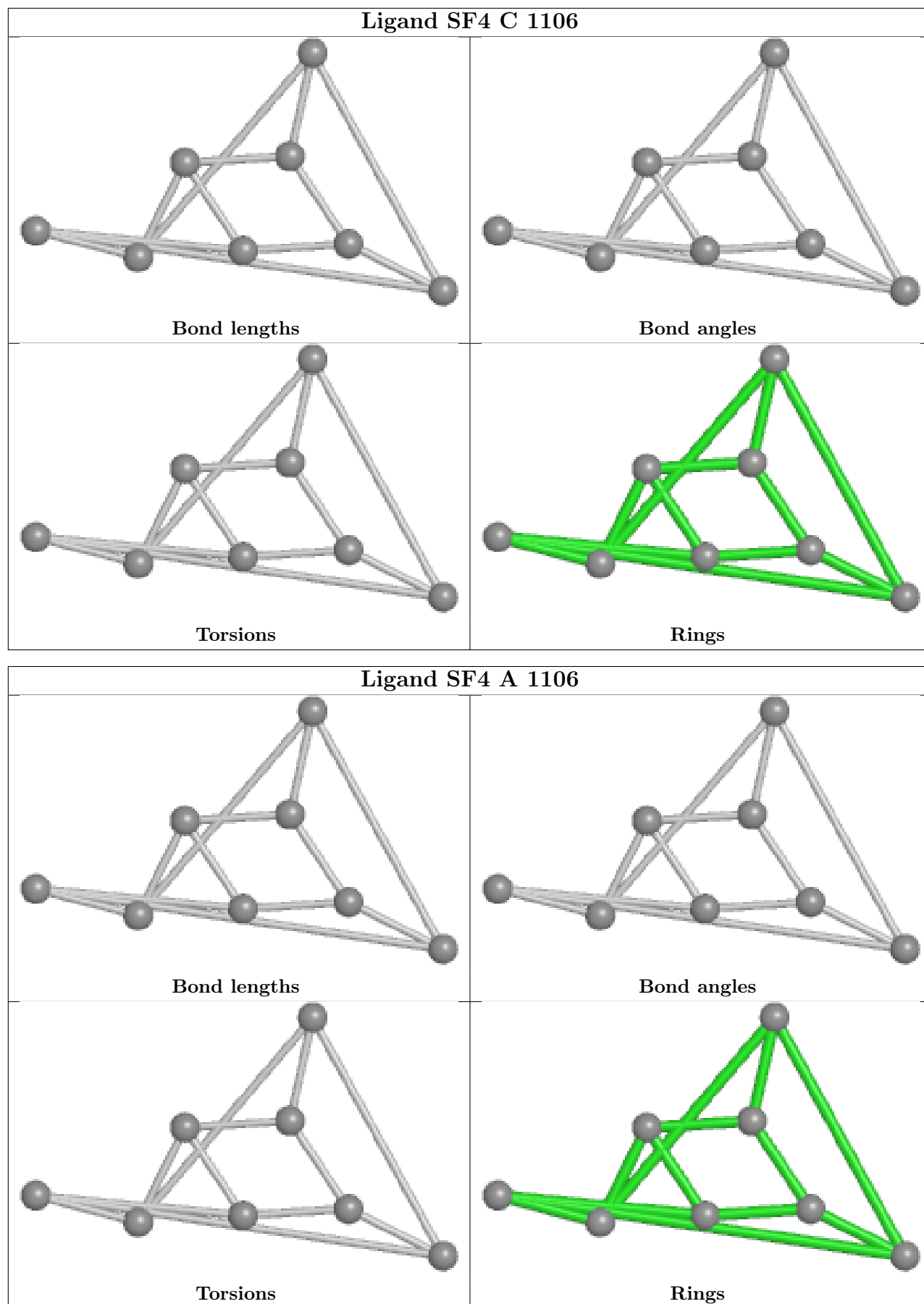
Continued from previous page...

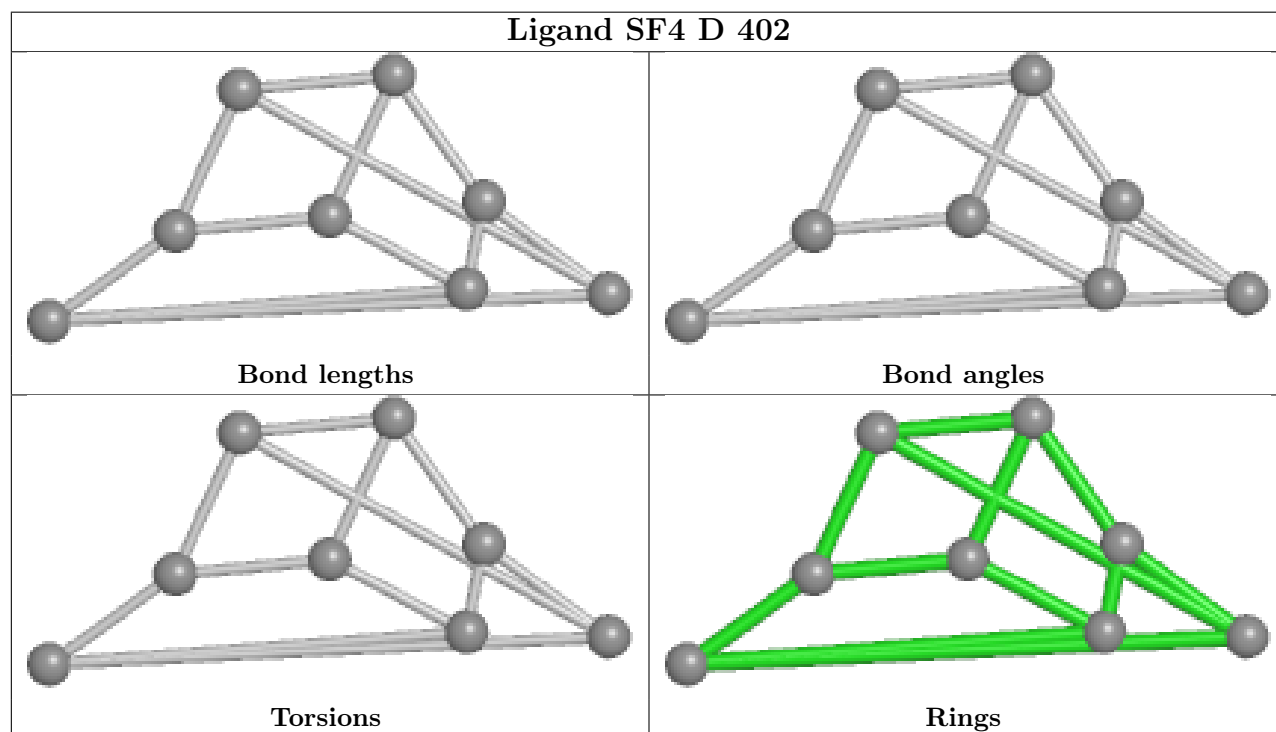
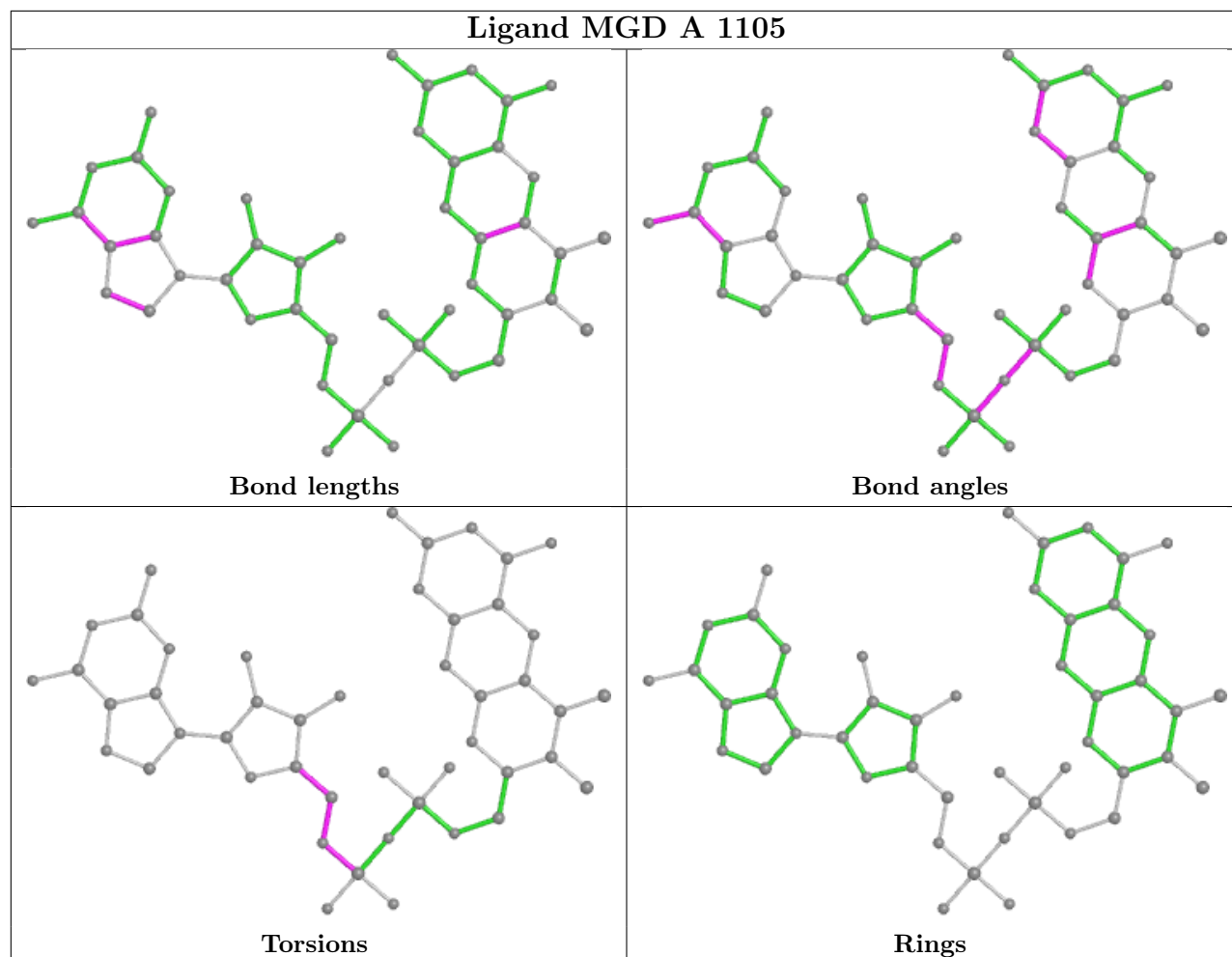
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1101	PEG	4	0
3	L	302	PEG	4	0
6	R	1104	MGD	2	0
6	A	1104	MGD	3	0
6	R	1103	MGD	3	0
6	K	1103	MGD	3	0
3	S	302	PEG	13	0

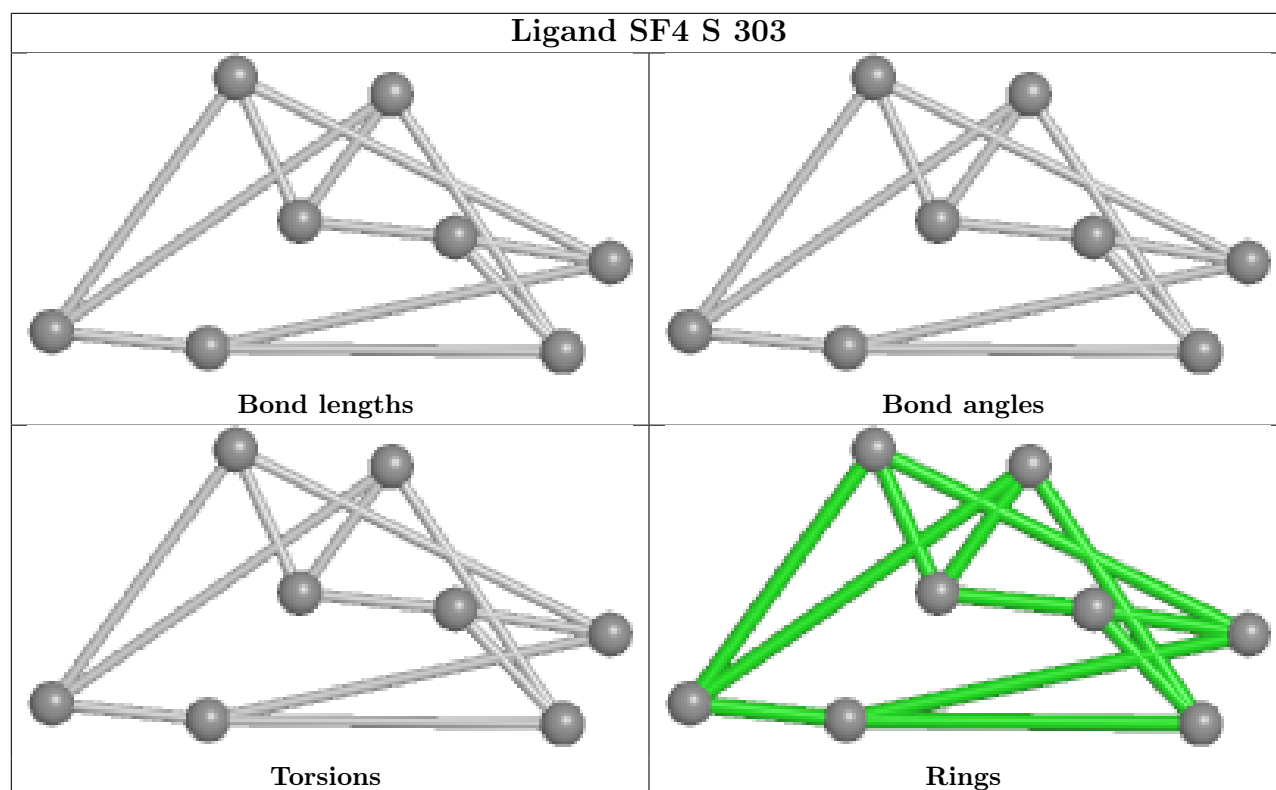
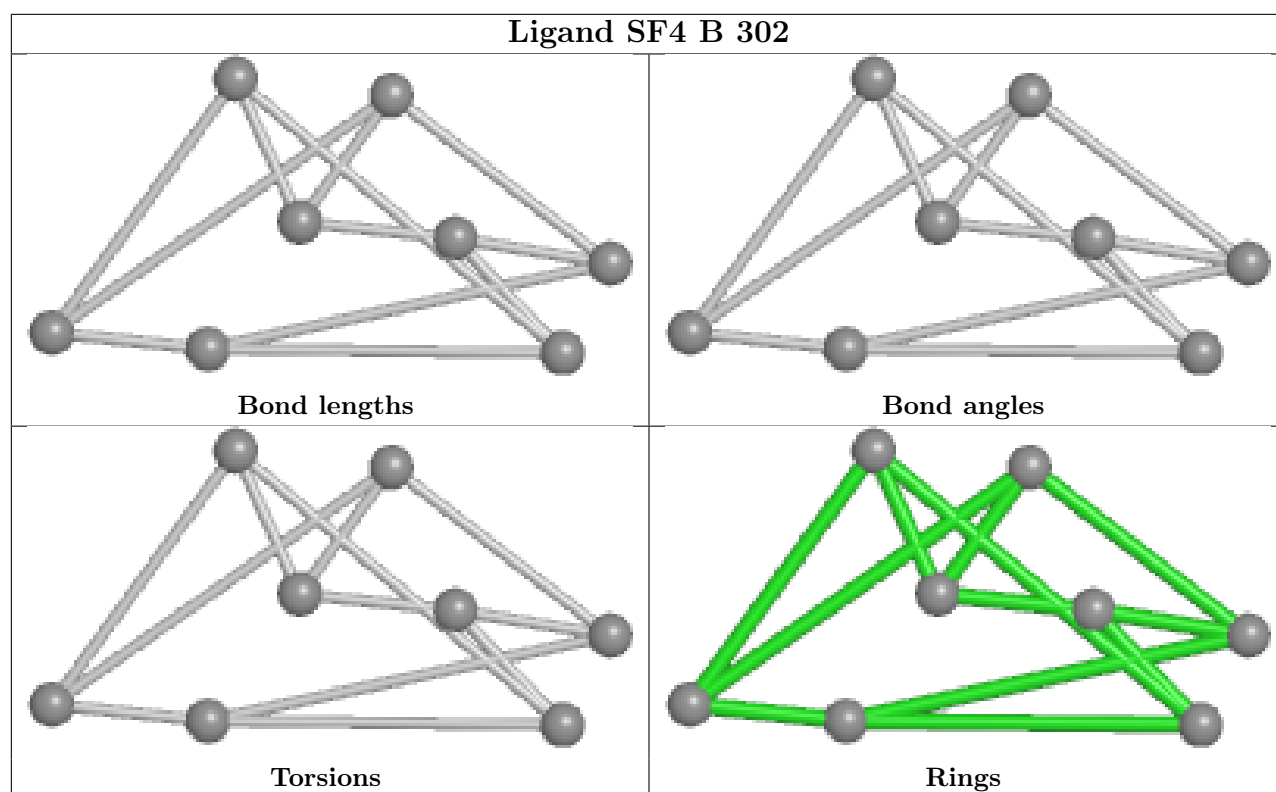
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

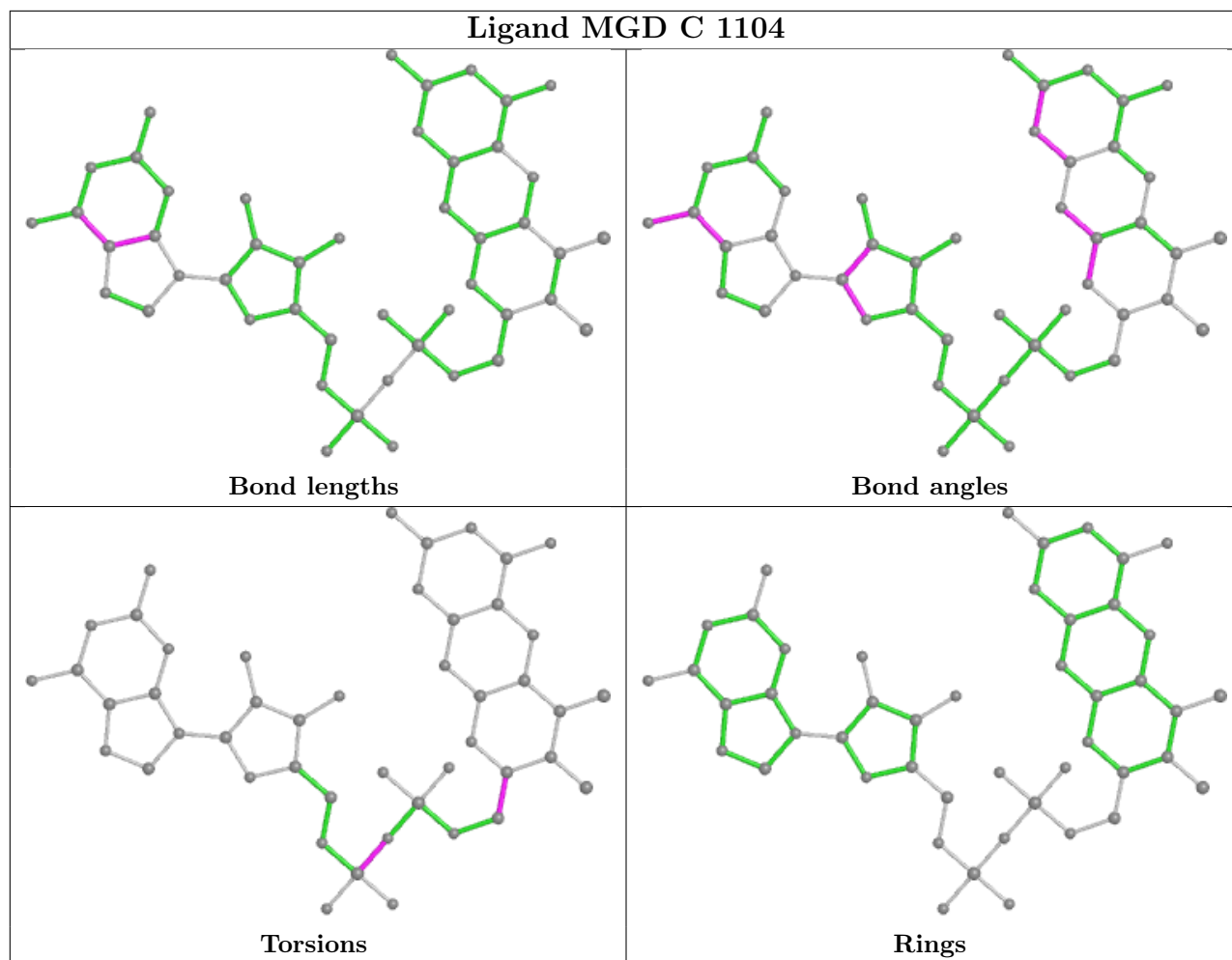


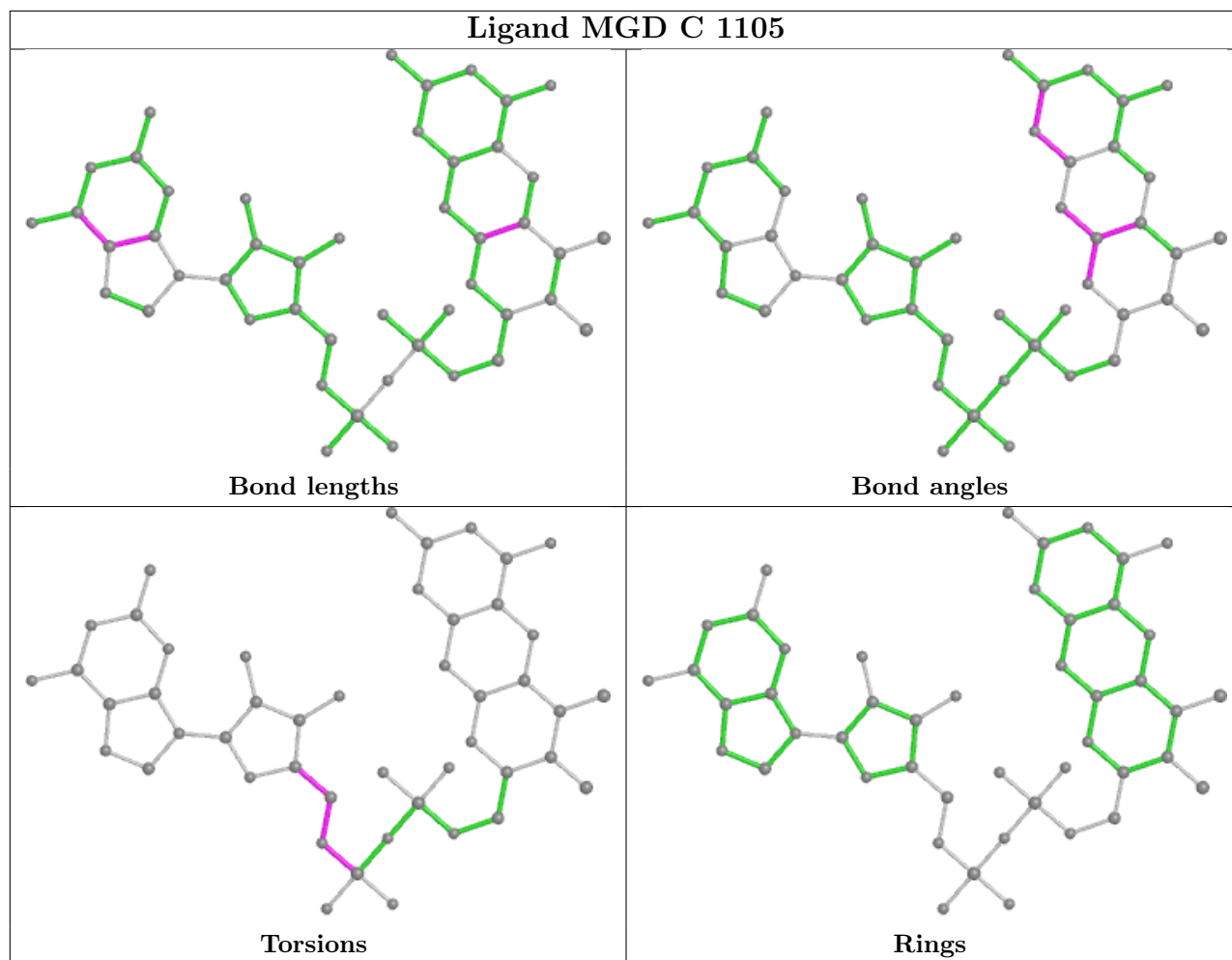


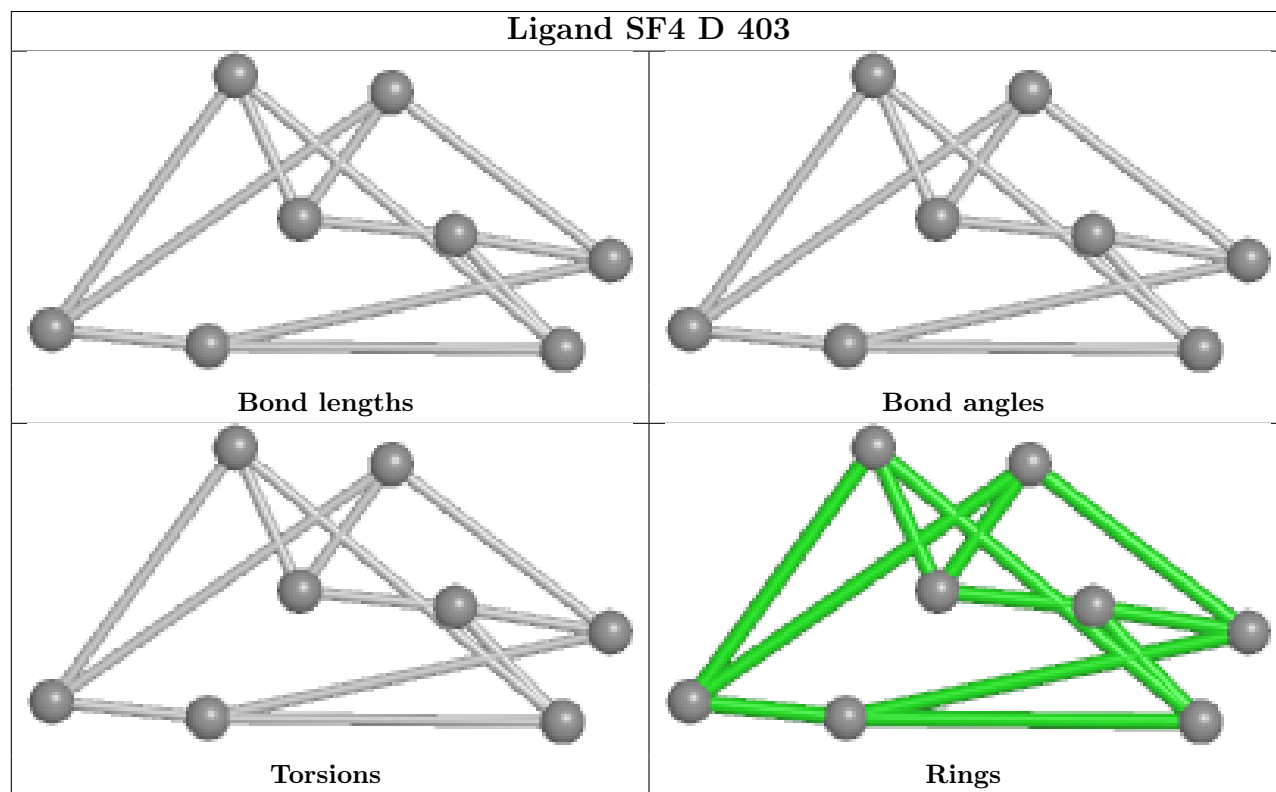
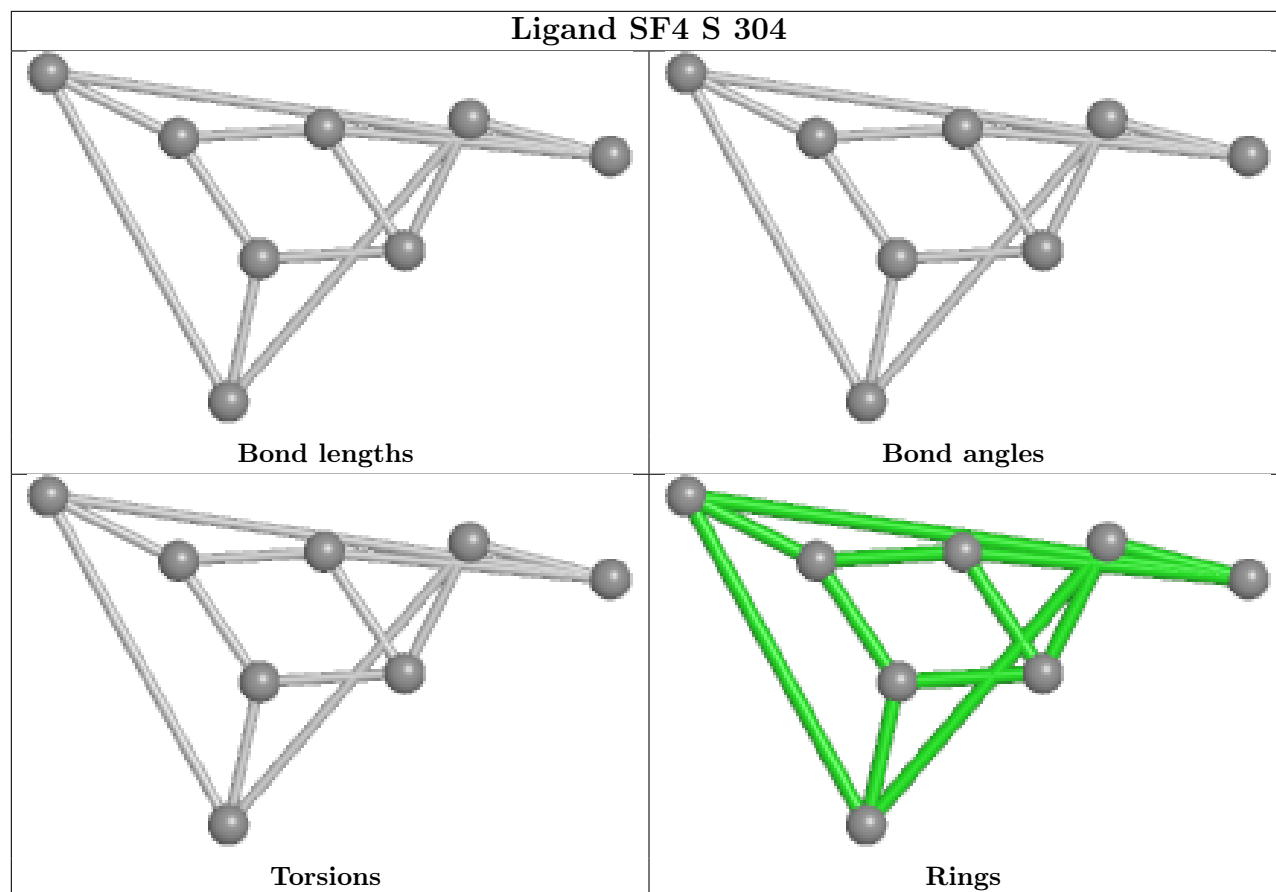


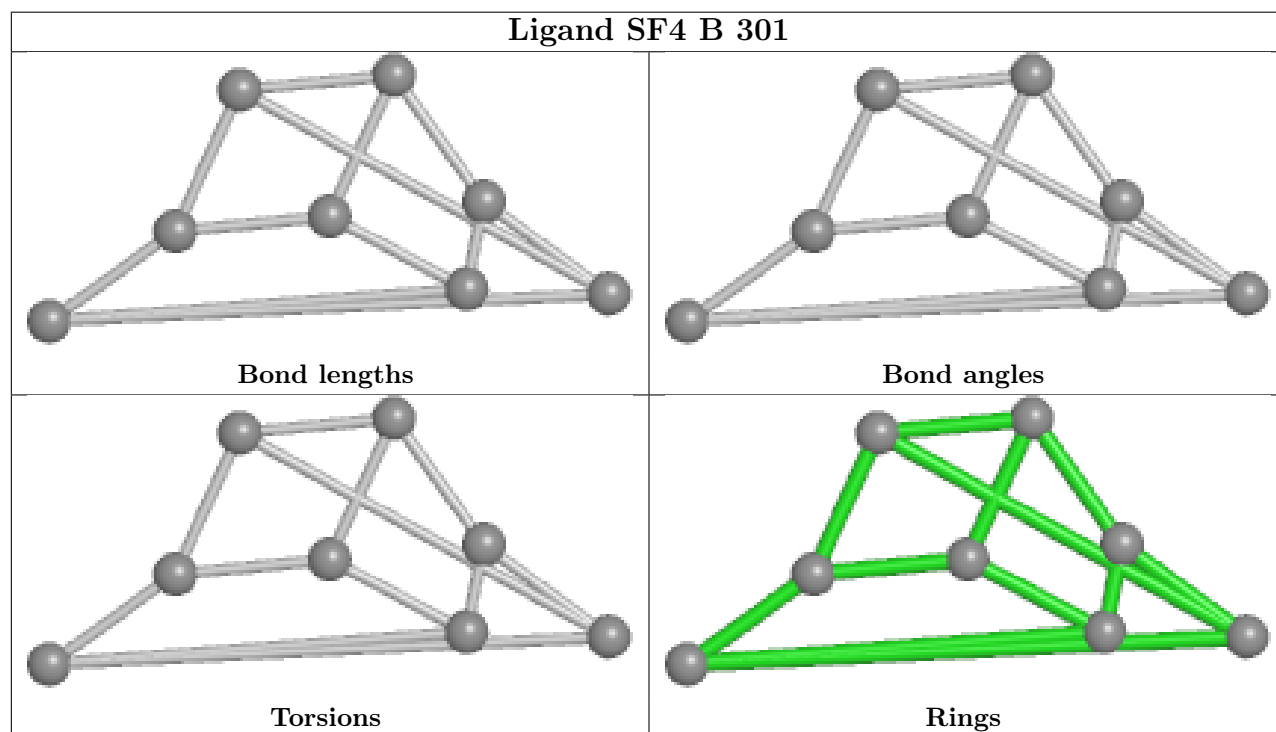
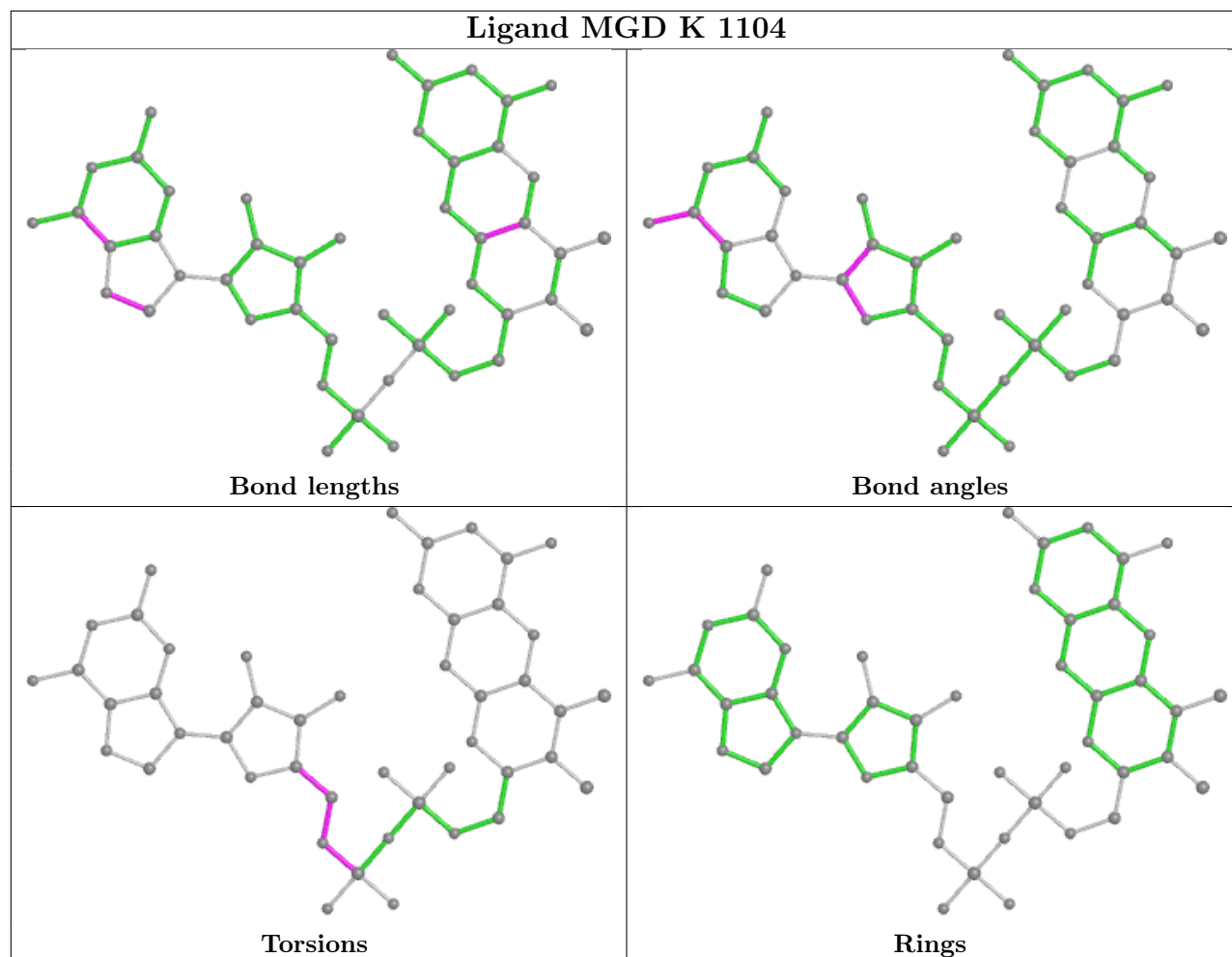


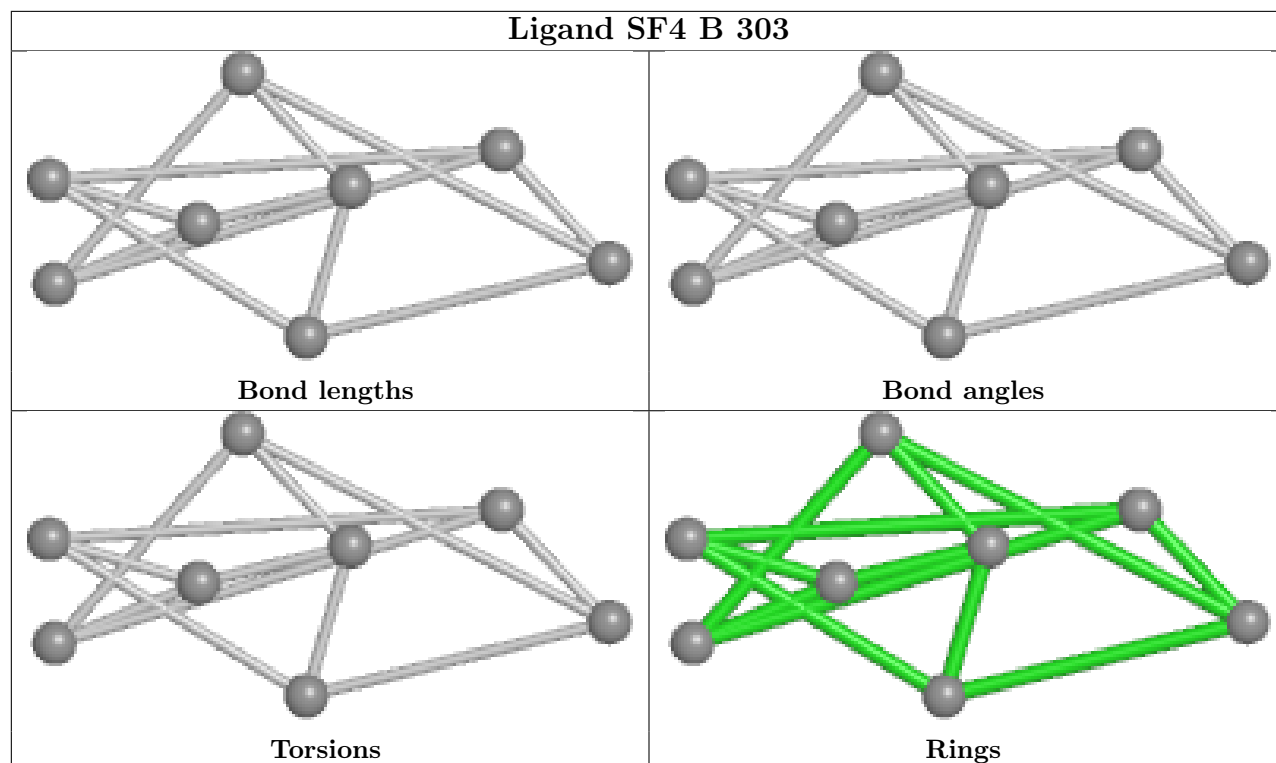
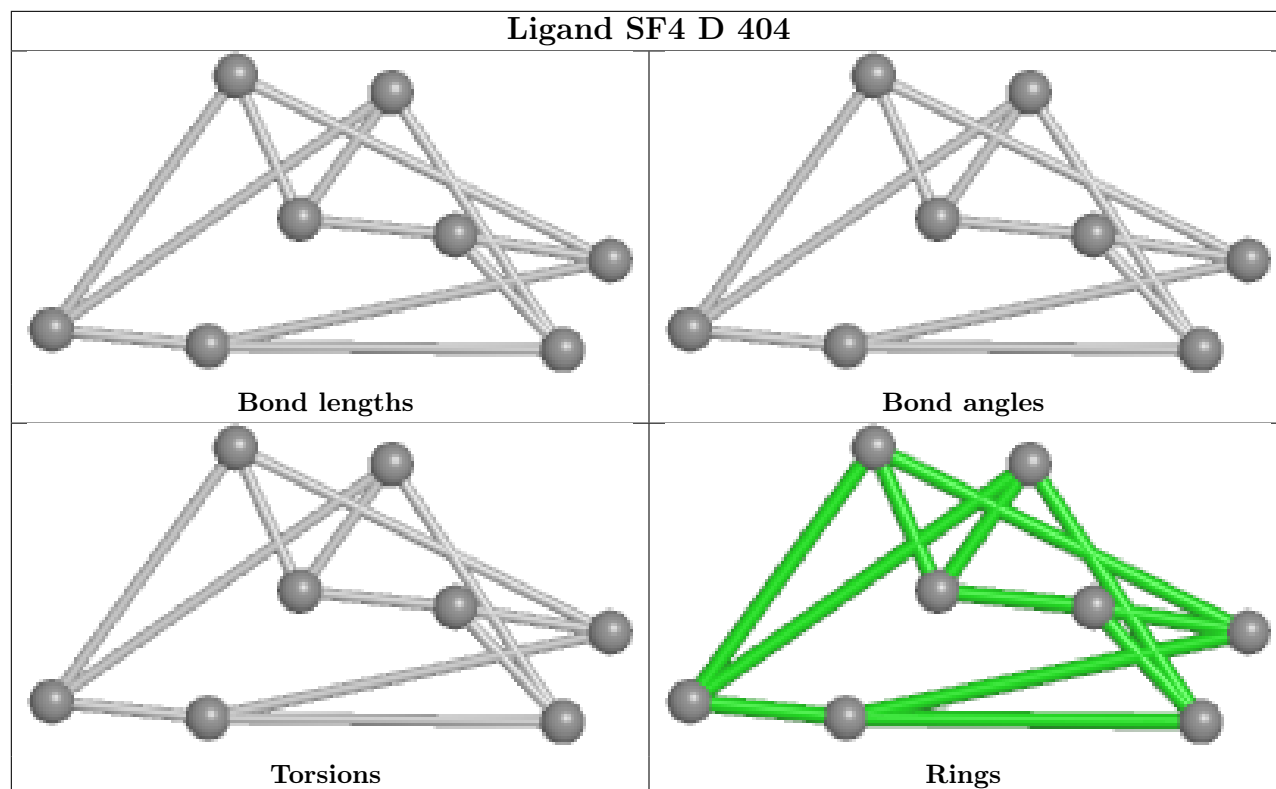


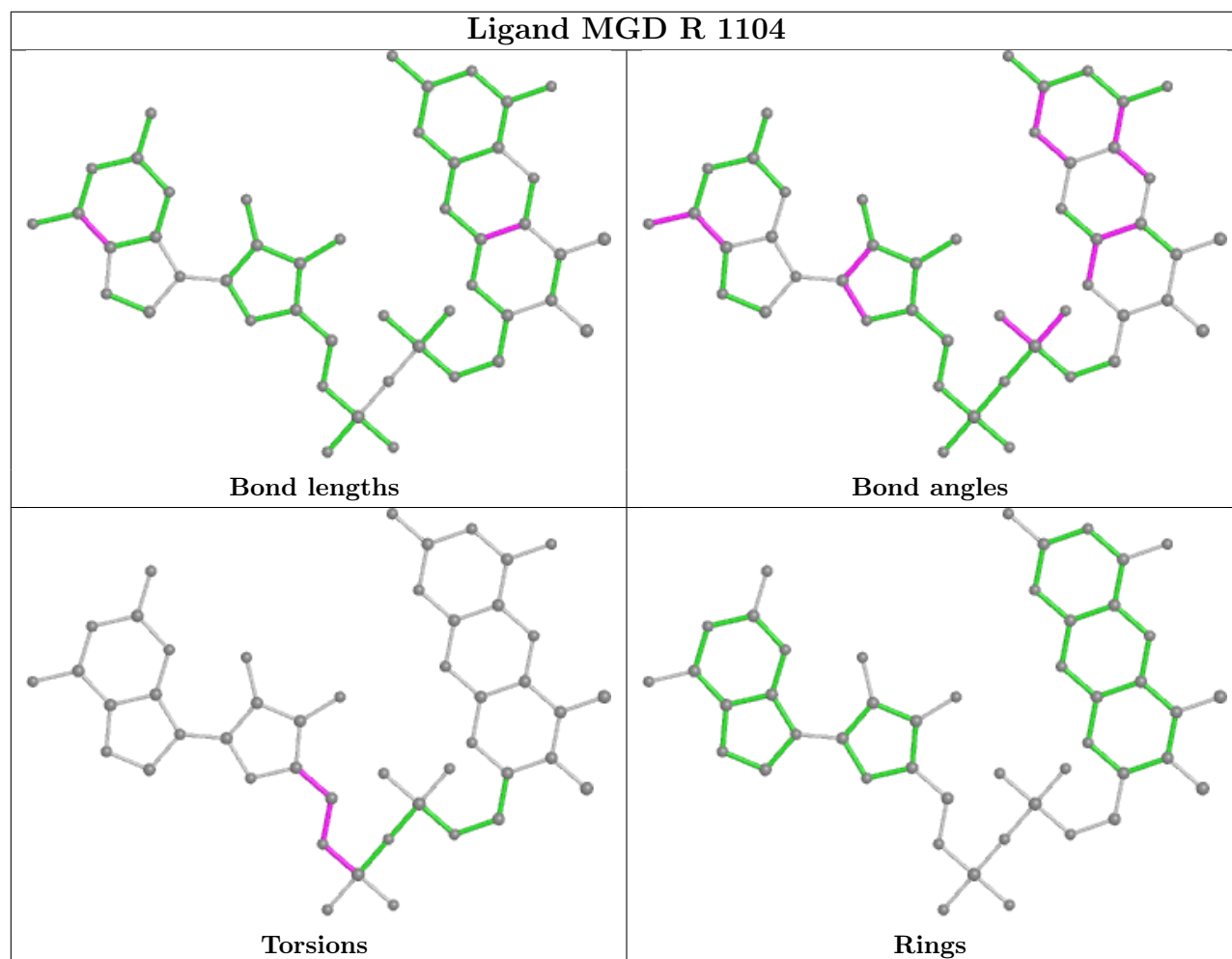
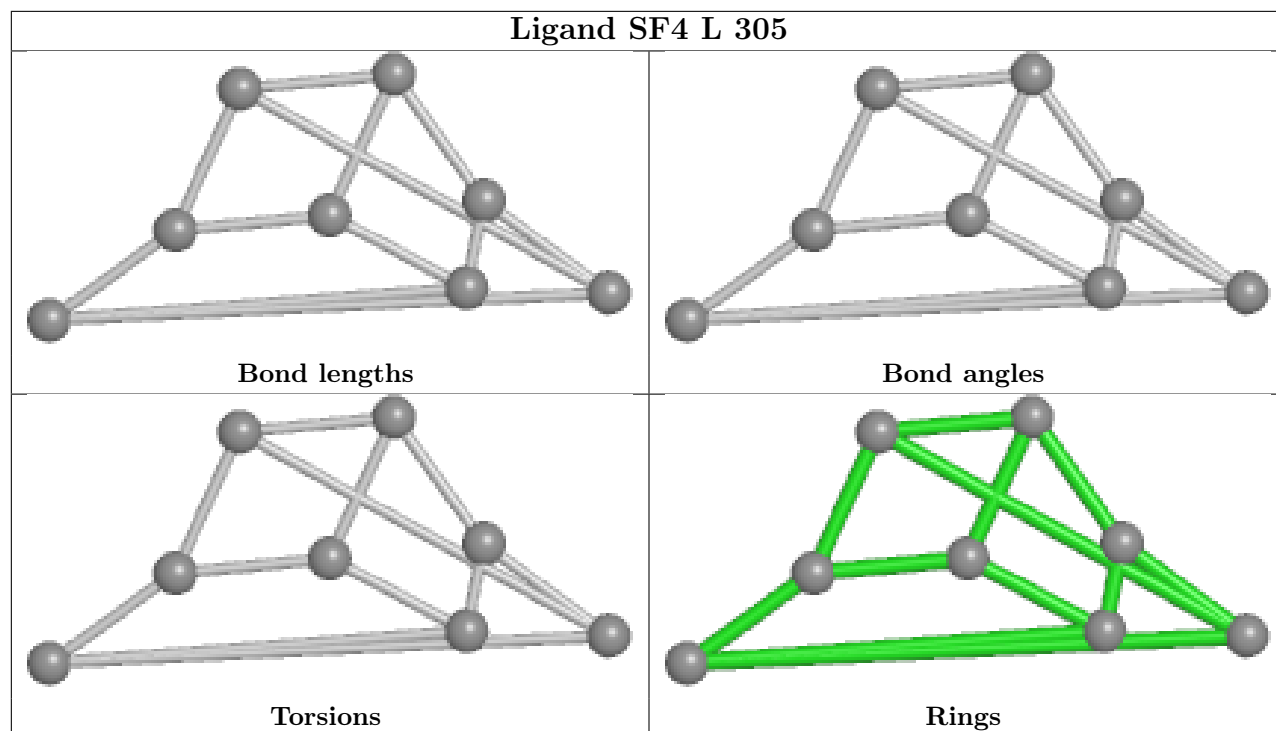


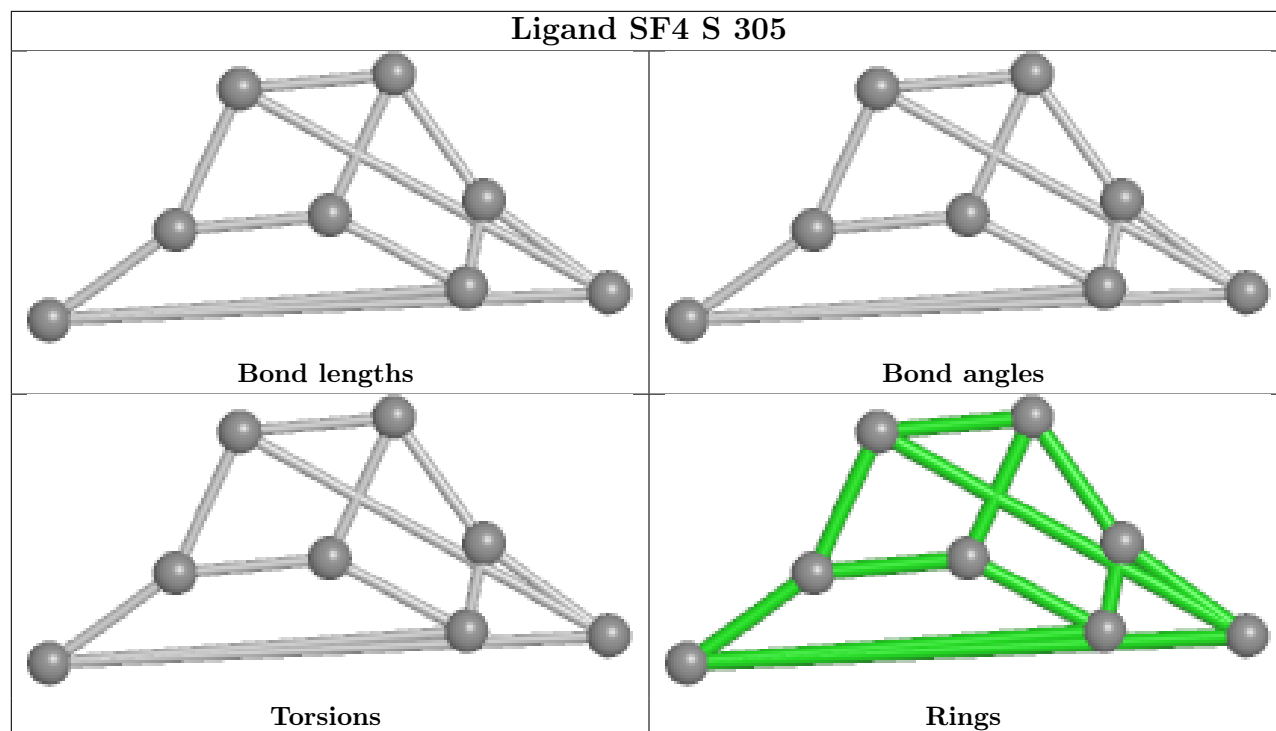
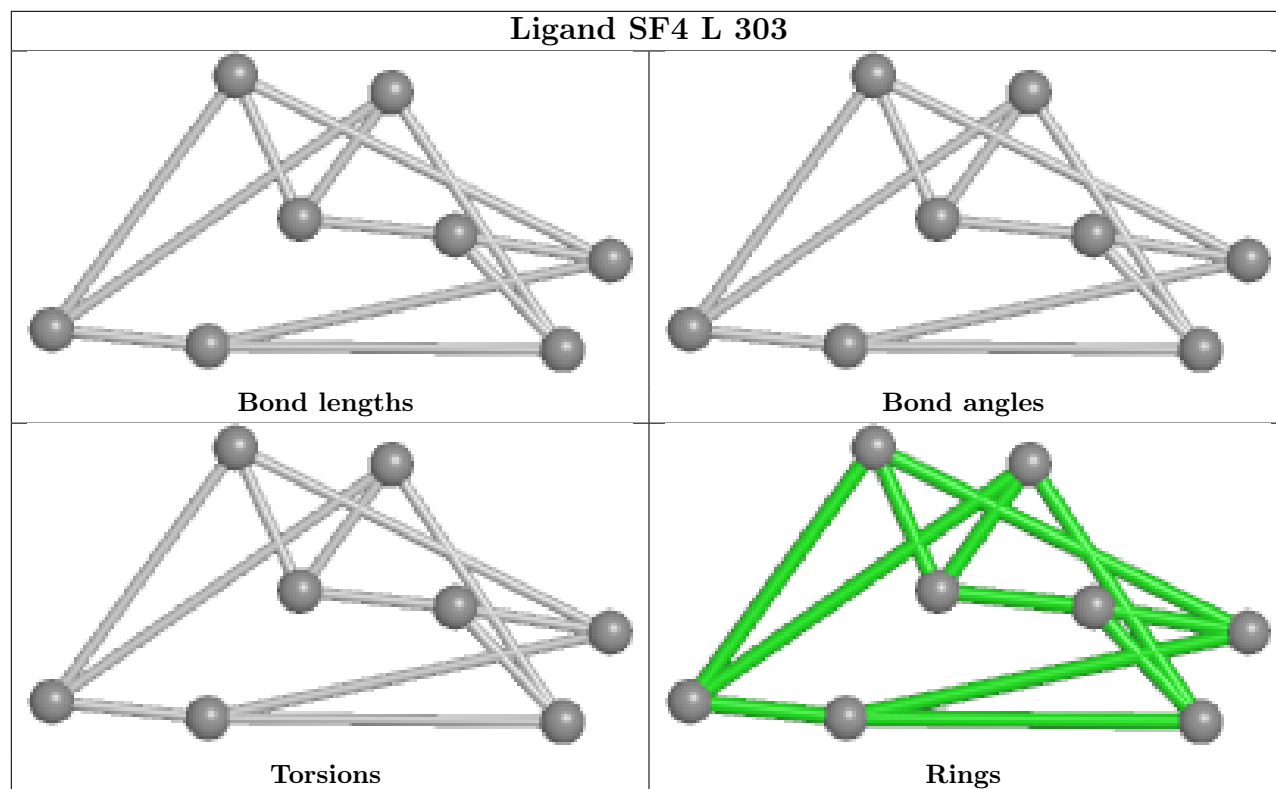


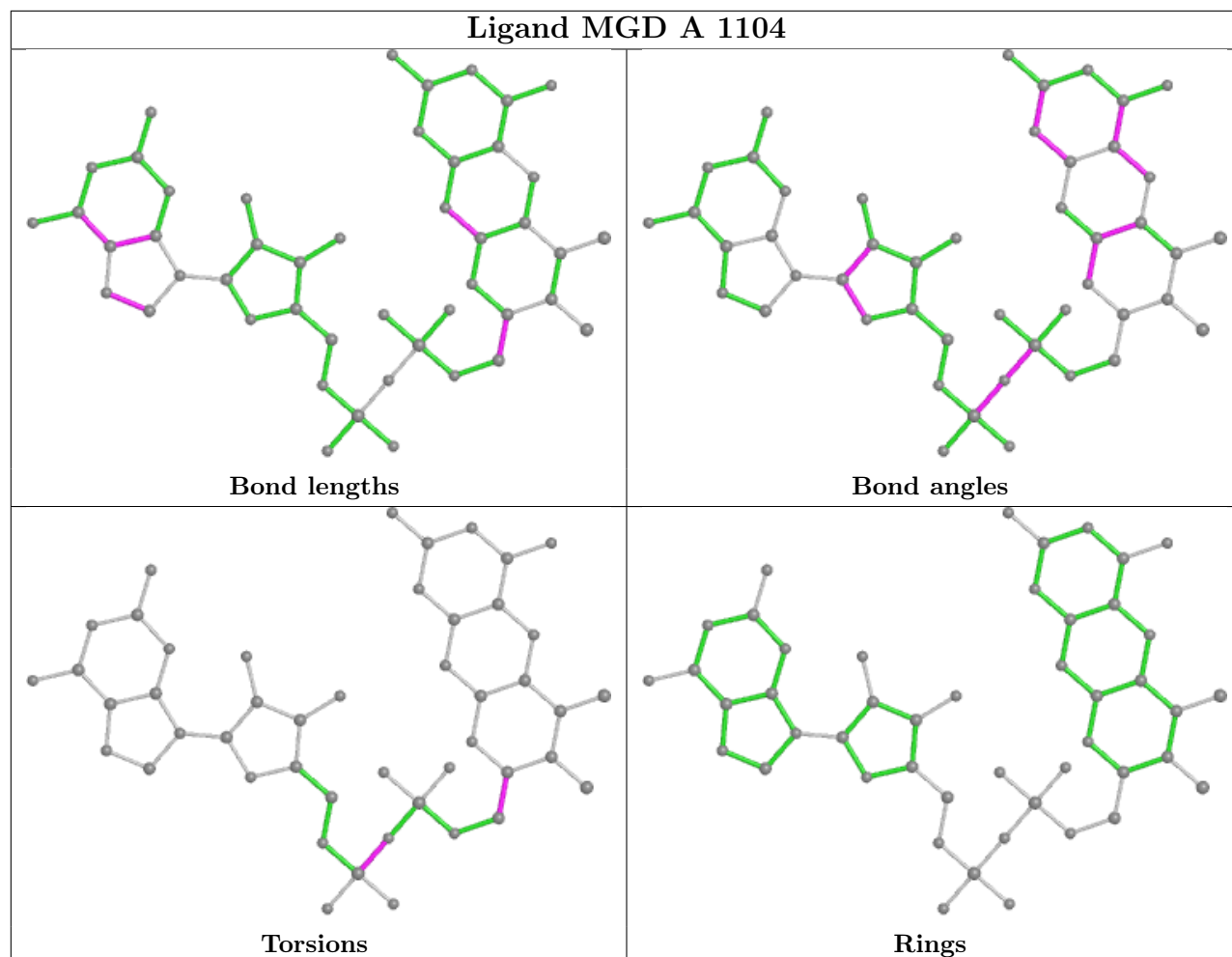


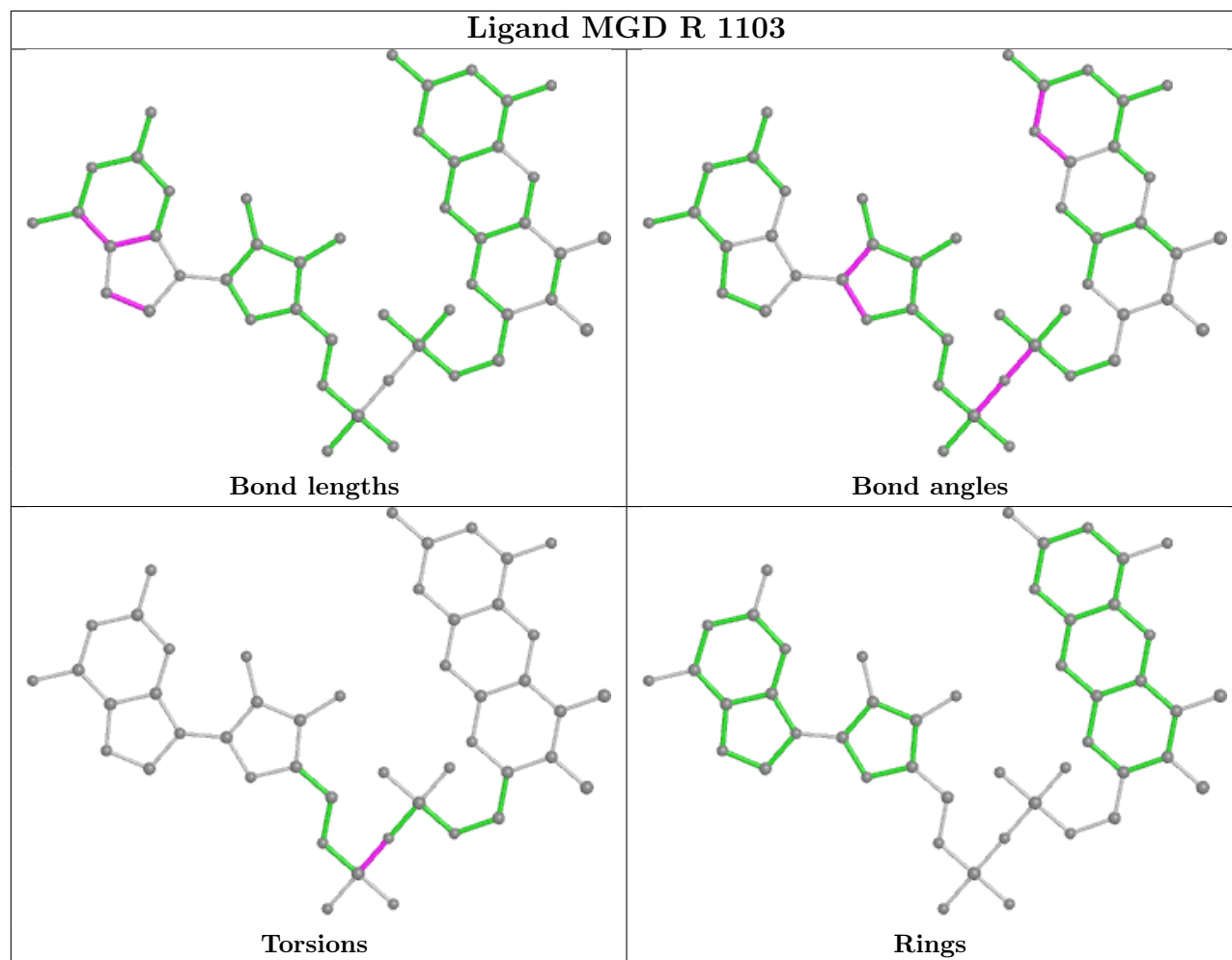


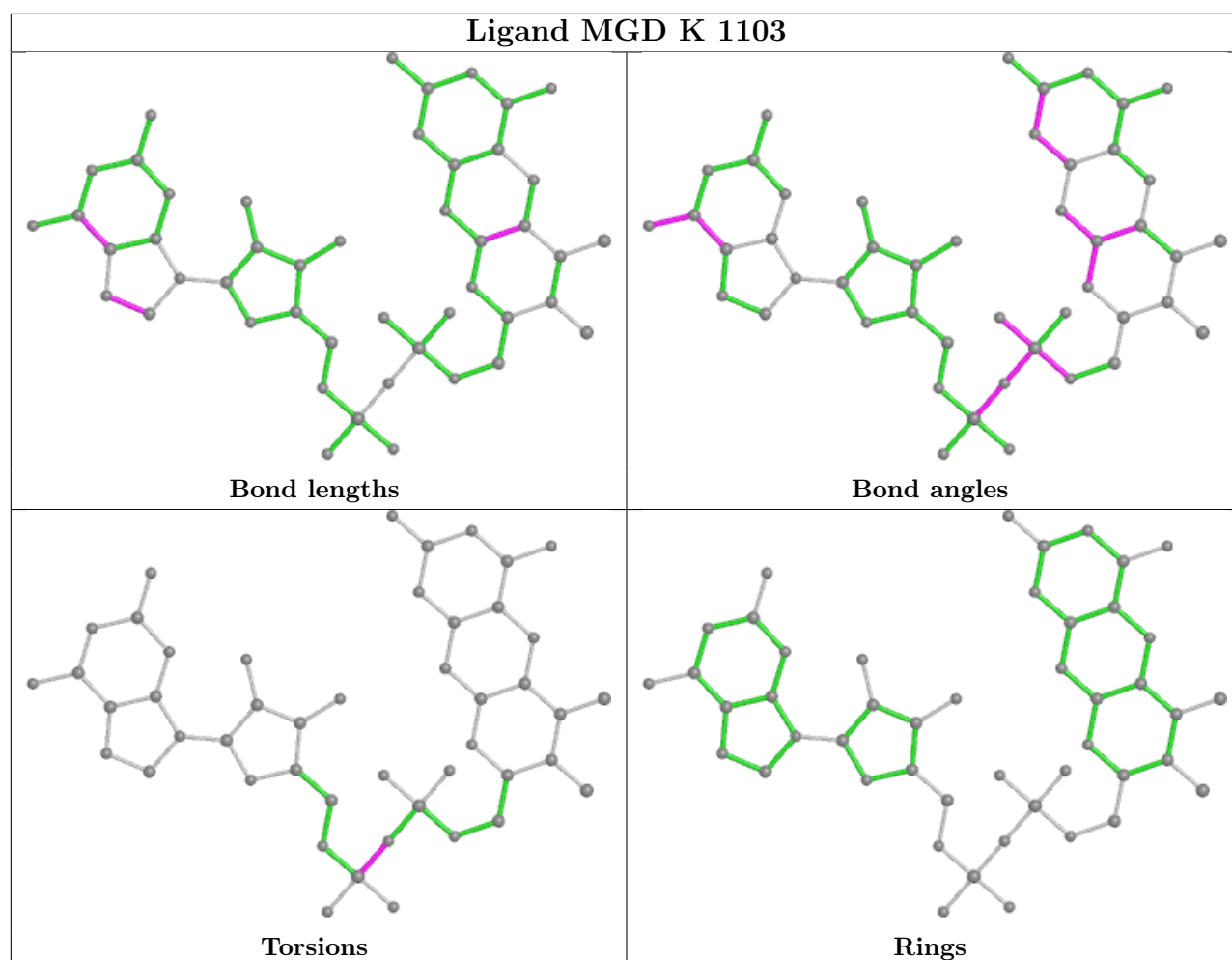












5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	973/1013 (96%)	-0.53	5 (0%) 91 93	15, 33, 64, 94	0
1	C	973/1013 (96%)	-0.50	5 (0%) 91 93	17, 36, 65, 97	0
1	K	972/1013 (95%)	-0.30	13 (1%) 77 82	17, 41, 71, 102	0
1	R	972/1013 (95%)	-0.47	1 (0%) 95 96	17, 38, 69, 100	0
2	B	214/215 (99%)	-0.72	0 100 100	14, 23, 42, 51	0
2	D	214/215 (99%)	-0.70	0 100 100	17, 25, 43, 57	0
2	L	214/215 (99%)	-0.72	0 100 100	15, 23, 43, 68	0
2	S	214/215 (99%)	-0.72	0 100 100	15, 24, 40, 66	0
All	All	4746/4912 (96%)	-0.50	24 (0%) 91 93	14, 34, 65, 102	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	702	GLY	6.7
1	C	339	ALA	5.6
1	A	1009	PRO	4.5
1	A	339	ALA	4.3
1	K	663	TYR	3.5

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.

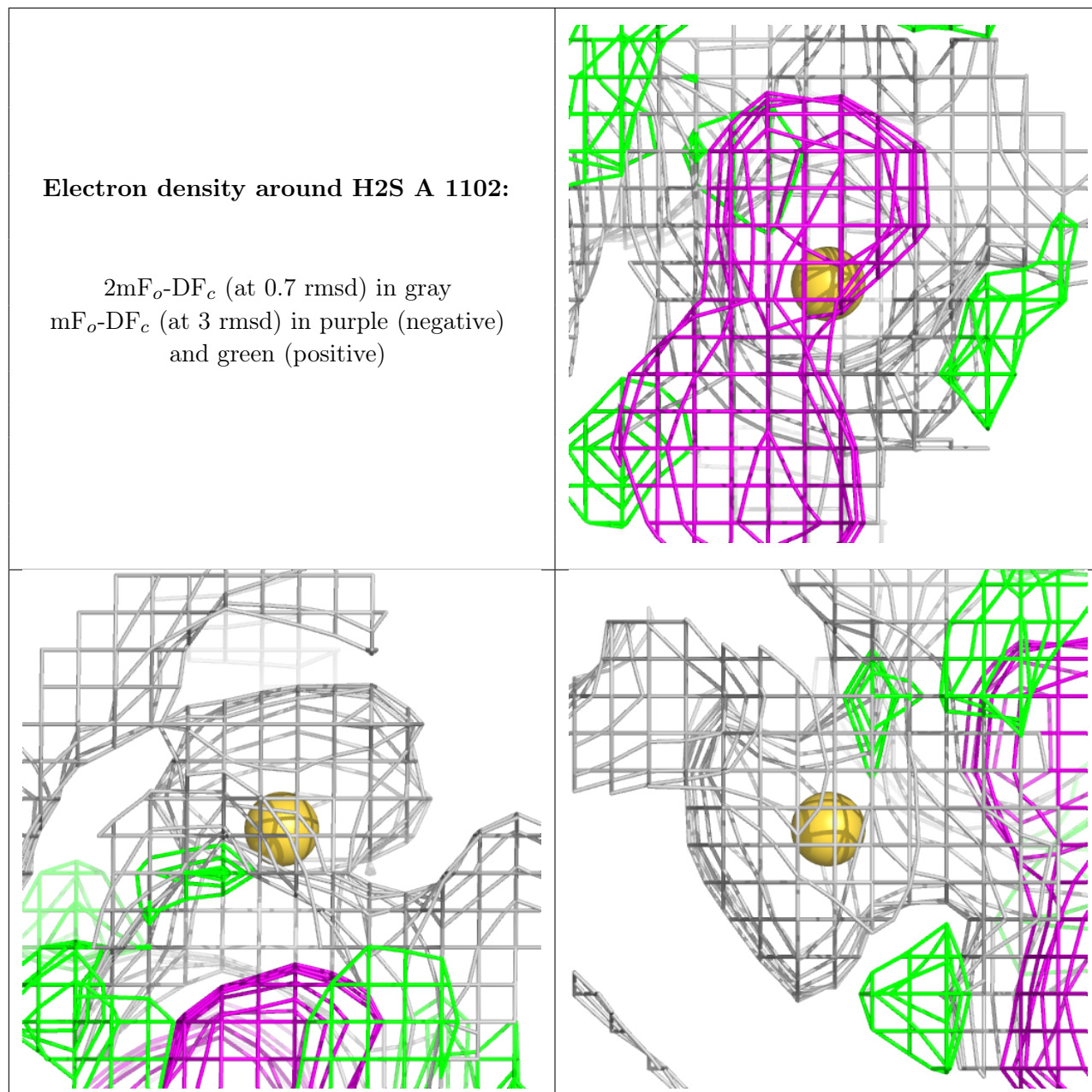
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
8	GOL	S	301	6/6	0.90	0.16	12,23,25,28	0
3	PEG	A	1101	7/7	0.92	0.12	46,53,62,65	0
3	PEG	L	301	7/7	0.93	0.27	19,26,34,35	0
3	PEG	C	1101	7/7	0.94	0.14	32,34,40,41	0
3	PEG	L	302	7/7	0.95	0.17	11,27,38,40	0
3	PEG	S	302	7/7	0.95	0.25	21,24,35,39	0
3	PEG	D	401	7/7	0.95	0.14	14,27,37,38	0
4	H2S	A	1102	1/1	0.96	0.07	38,38,38,38	0
6	MGD	C	1105	47/47	0.98	0.08	17,24,31,33	0
6	MGD	K	1103	47/47	0.98	0.08	19,23,31,35	0
4	H2S	C	1102	1/1	0.98	0.07	23,23,23,23	0
6	MGD	A	1105	47/47	0.99	0.07	15,20,29,31	0
6	MGD	C	1104	47/47	0.99	0.08	16,22,25,26	0
4	H2S	K	1101	1/1	0.99	0.11	20,20,20,20	0
4	H2S	R	1101	1/1	0.99	0.09	31,31,31,31	0
6	MGD	K	1104	47/47	0.99	0.09	16,26,34,40	0
6	MGD	R	1103	47/47	0.99	0.07	18,21,25,27	0
6	MGD	R	1104	47/47	0.99	0.08	16,21,30,33	0
7	SF4	K	1105	8/8	0.99	0.10	18,19,20,21	0
7	SF4	S	304	8/8	0.99	0.10	17,17,20,21	0
6	MGD	A	1104	47/47	0.99	0.08	14,18,22,29	0
5	W	R	1102	1/1	1.00	0.06	22,22,22,22	0
7	SF4	A	1106	8/8	1.00	0.11	16,17,18,20	0
7	SF4	B	301	8/8	1.00	0.11	13,15,16,16	0
7	SF4	B	302	8/8	1.00	0.09	16,17,20,21	0
7	SF4	B	303	8/8	1.00	0.09	15,17,19,19	0
7	SF4	C	1106	8/8	1.00	0.10	20,21,22,23	0
7	SF4	D	402	8/8	1.00	0.12	18,19,21,22	0
7	SF4	D	403	8/8	1.00	0.08	19,19,20,21	0
7	SF4	D	404	8/8	1.00	0.10	19,21,21,22	0
5	W	A	1103	1/1	1.00	0.07	20,20,20,20	0
7	SF4	L	303	8/8	1.00	0.11	16,17,18,18	0
7	SF4	L	304	8/8	1.00	0.09	16,17,17,17	0
7	SF4	L	305	8/8	1.00	0.10	17,17,18,19	0
7	SF4	R	1105	8/8	1.00	0.10	19,22,23,24	0
7	SF4	S	303	8/8	1.00	0.10	14,16,18,18	0
5	W	C	1103	1/1	1.00	0.06	22,22,22,22	0

Continued on next page...

Continued from previous page...

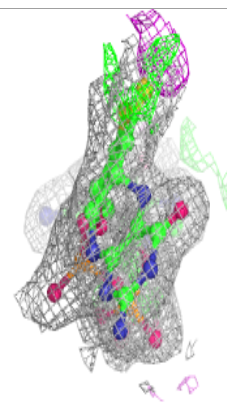
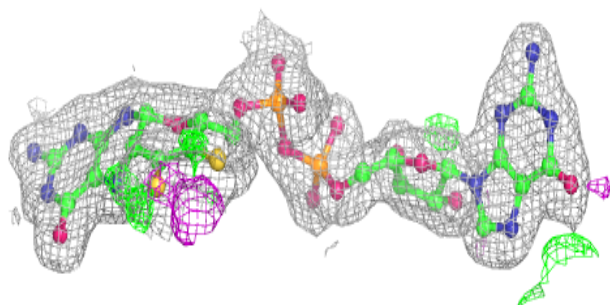
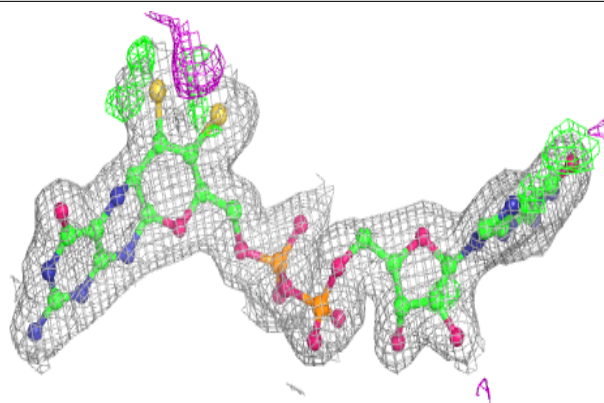
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
7	SF4	S	305	8/8	1.00	0.10	17,18,19,20	0
5	W	K	1102	1/1	1.00	0.06	24,24,24,24	0

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.

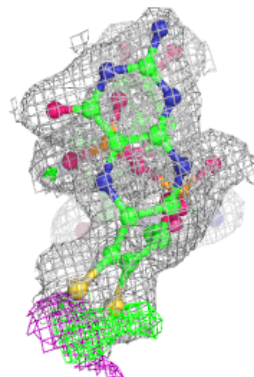
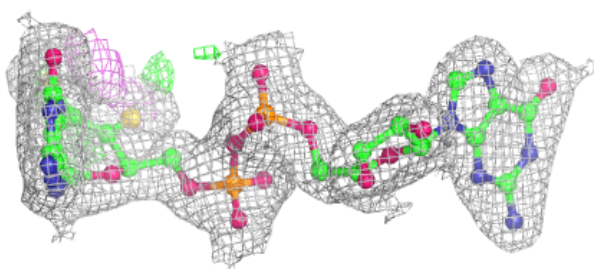
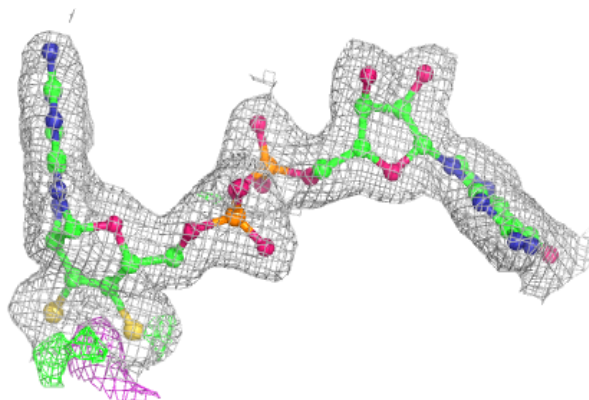


Electron density around MGD C 1105:

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

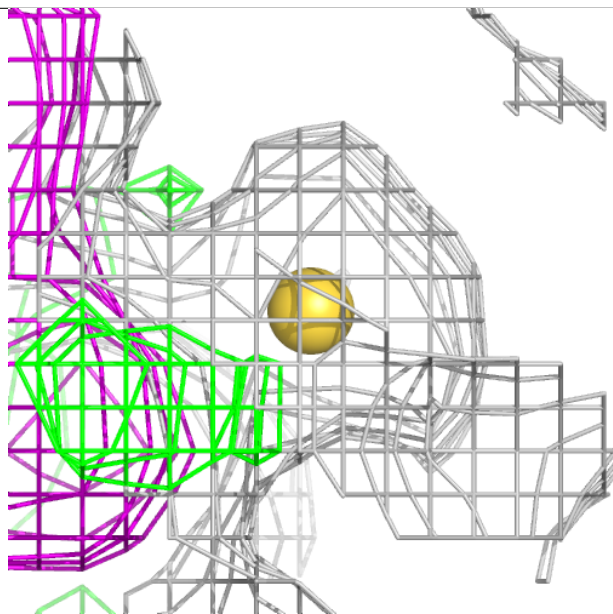
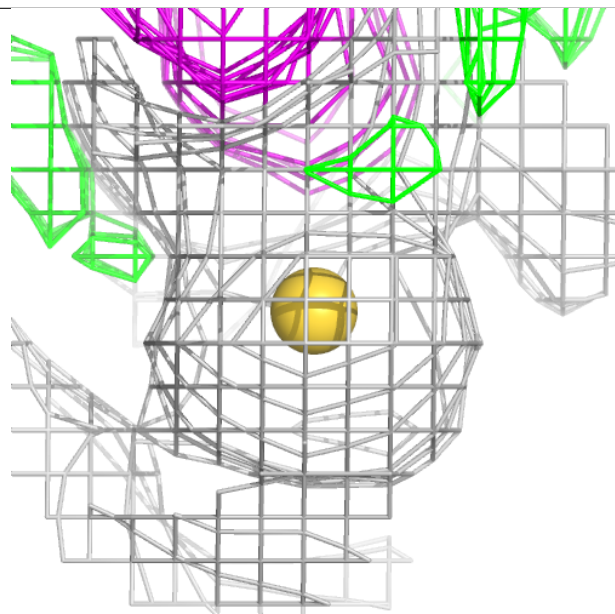
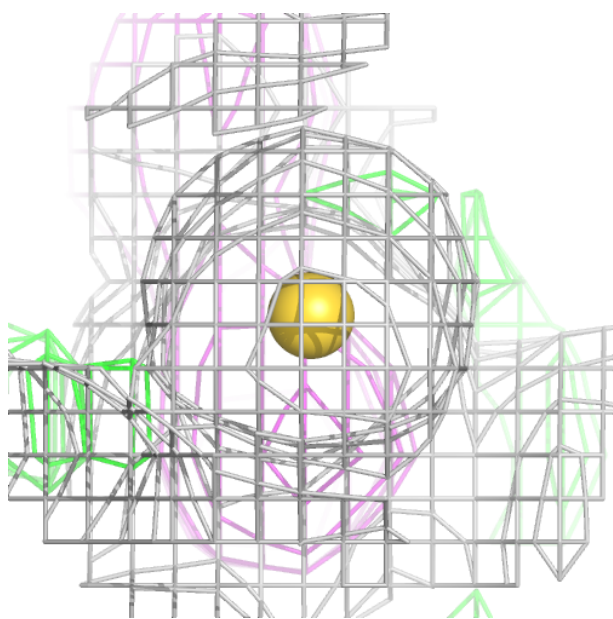
**Electron density around MGD K 1103:**

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



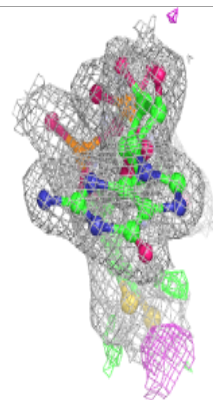
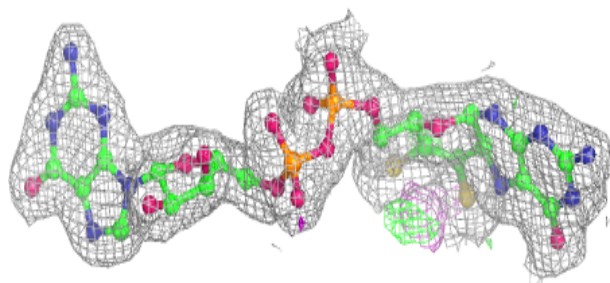
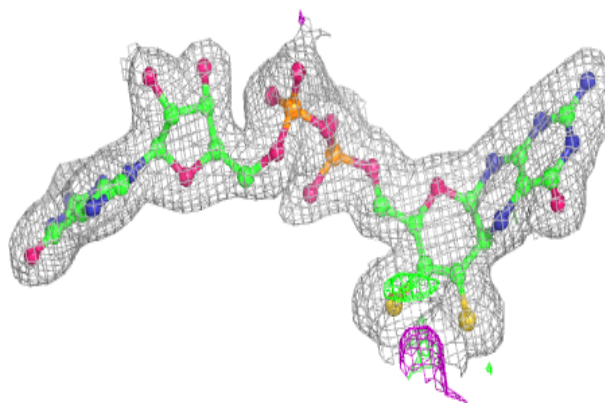
Electron density around H2S C 1102:

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

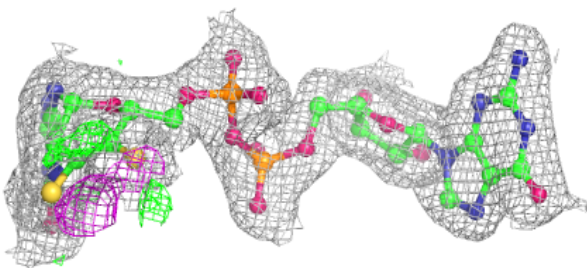
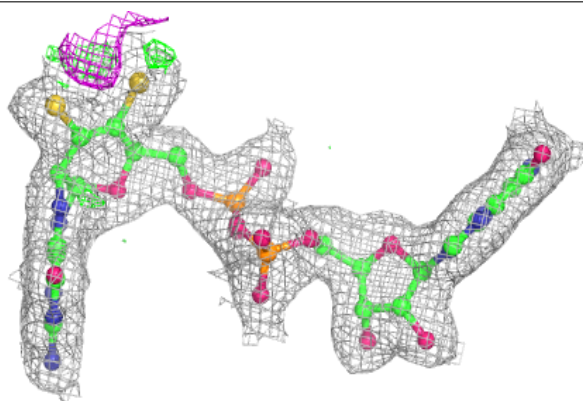


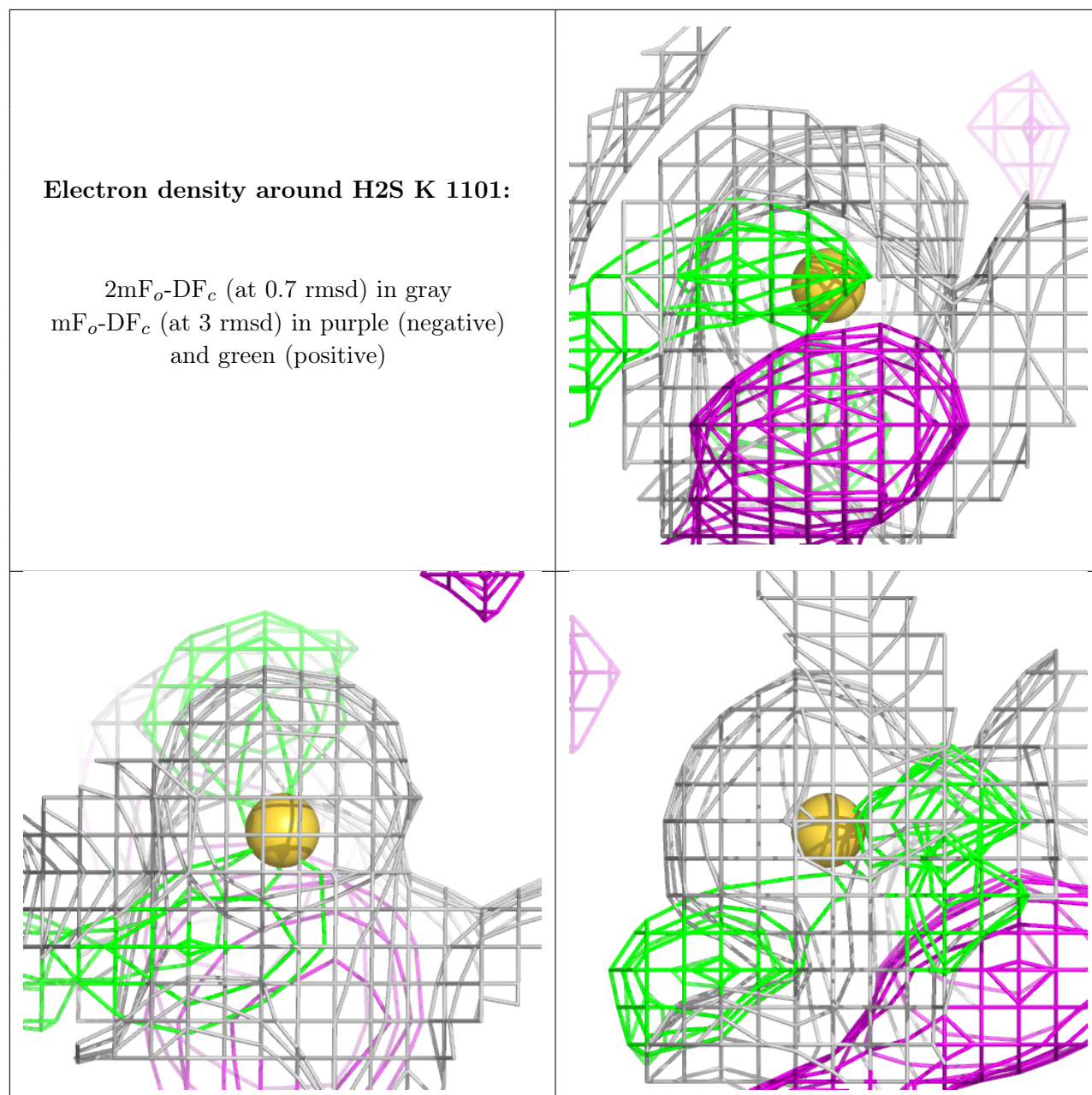
Electron density around MGD A 1105:

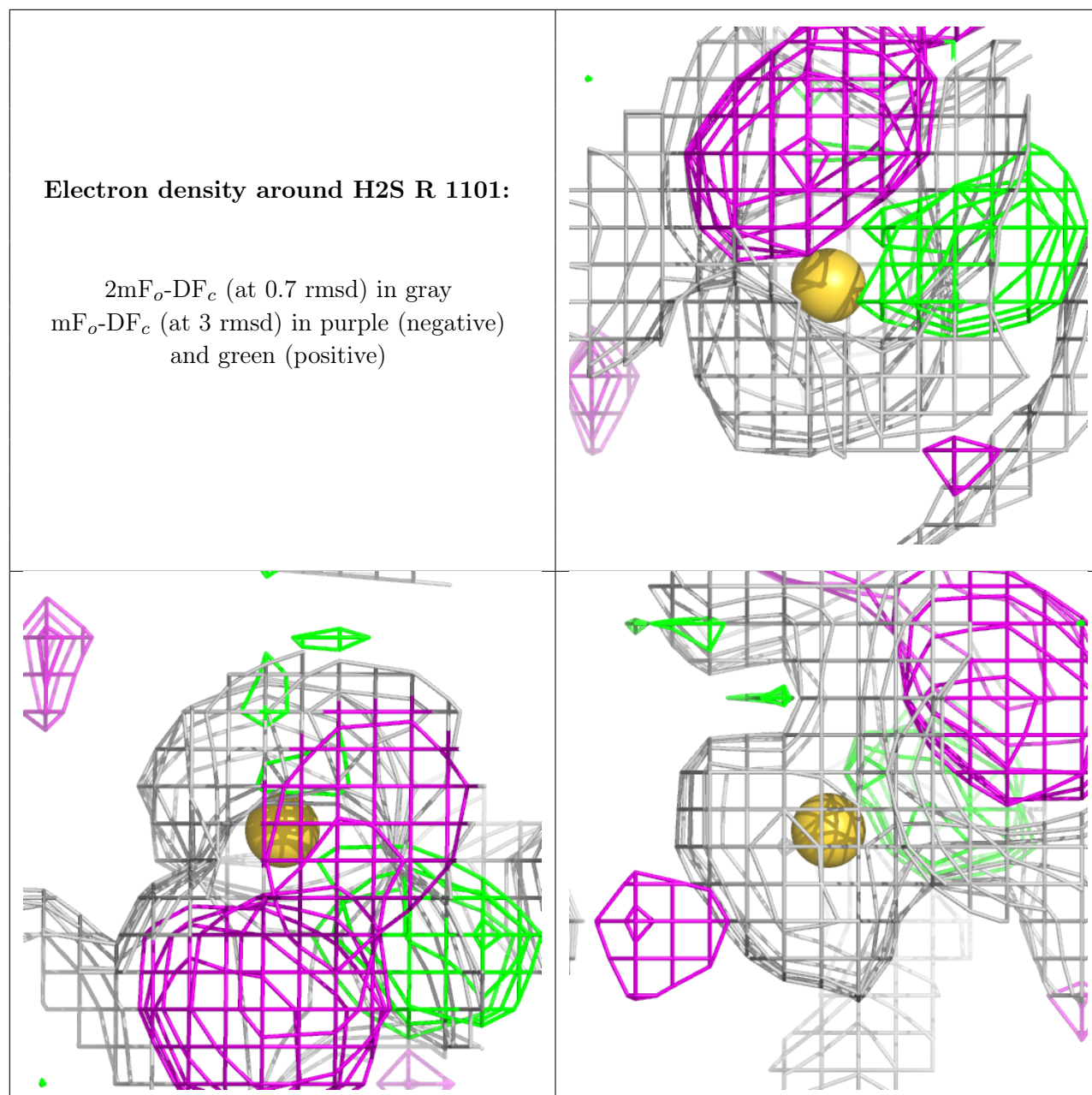
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around MGD C 1104:**

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

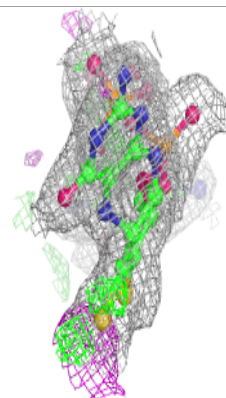
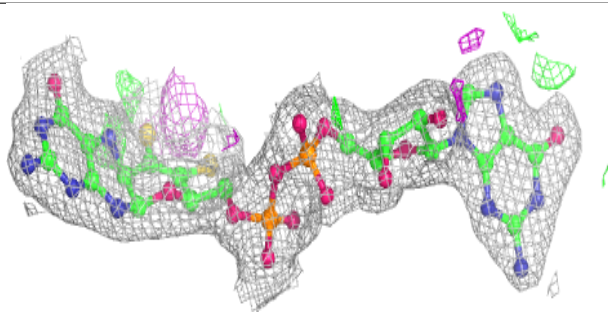
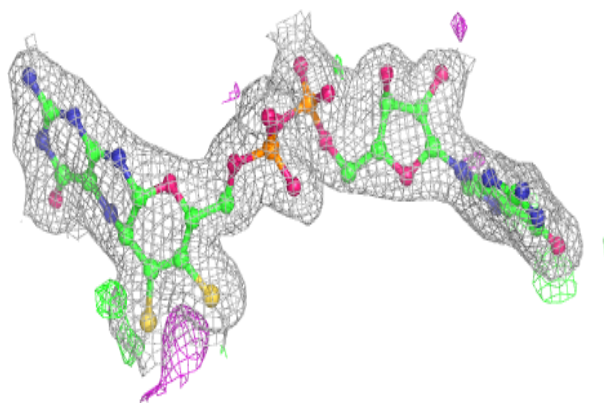




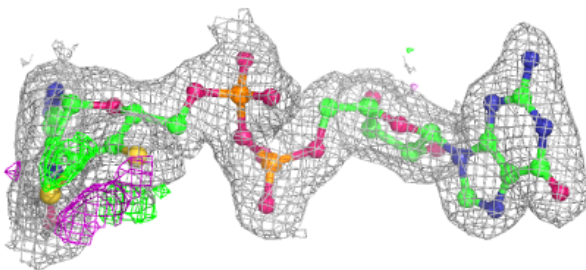
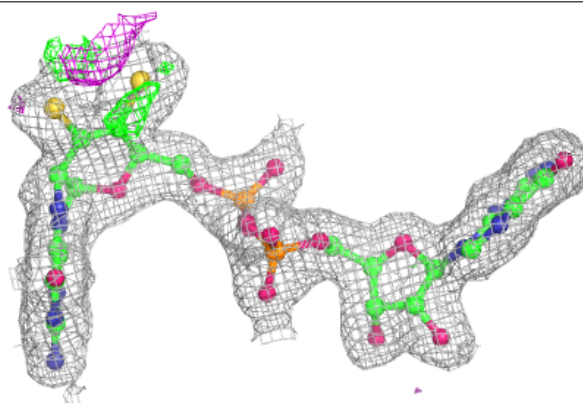


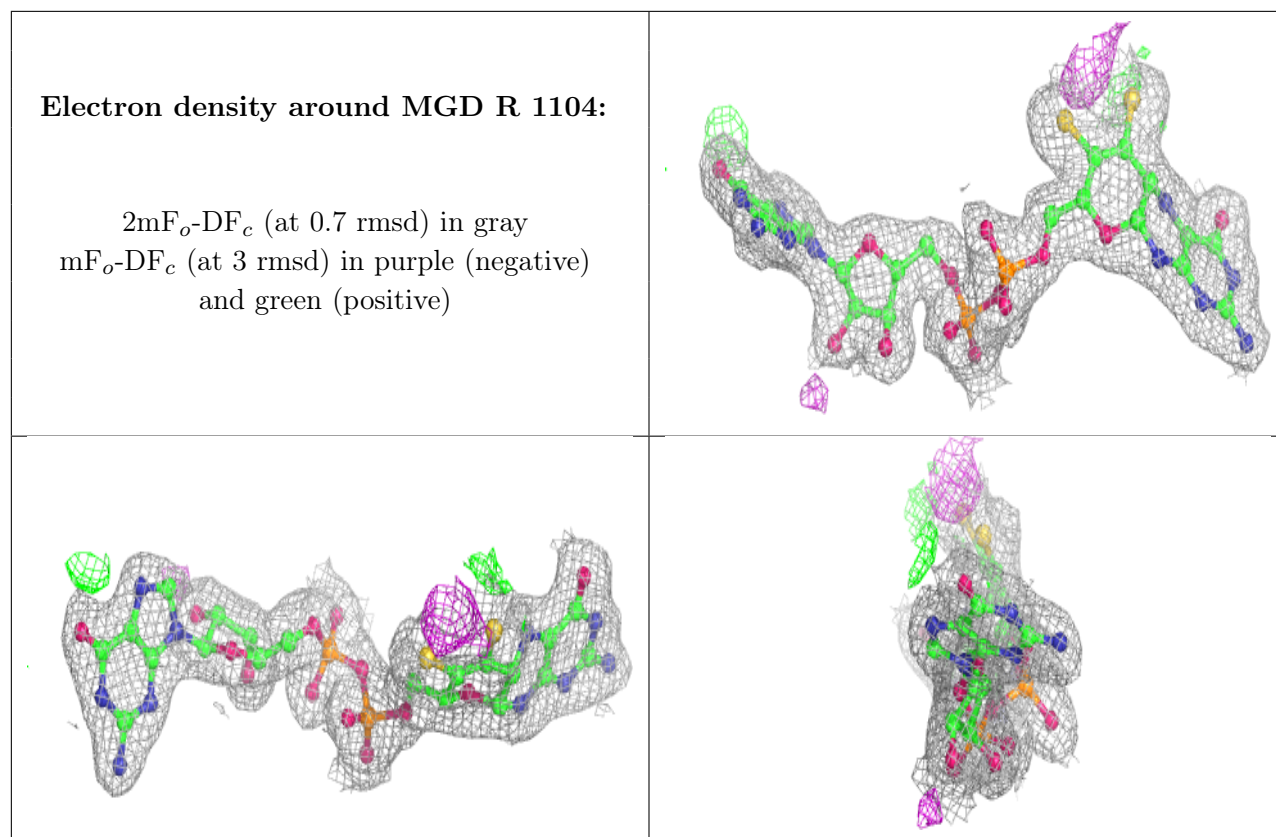
Electron density around MGD K 1104:

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

**Electron density around MGD R 1103:**

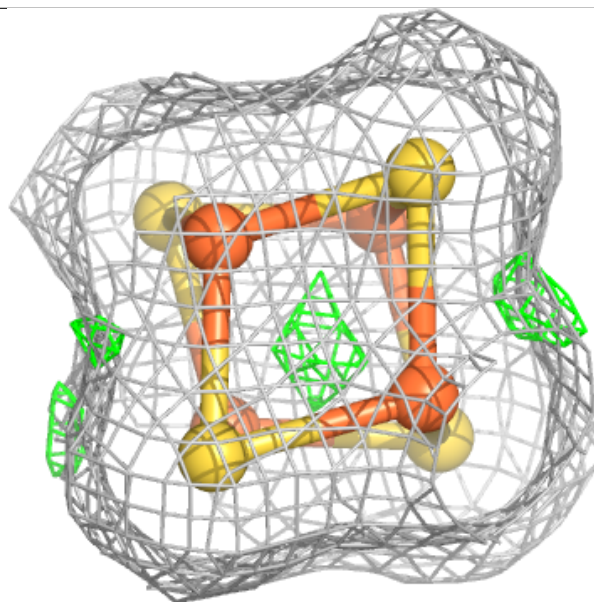
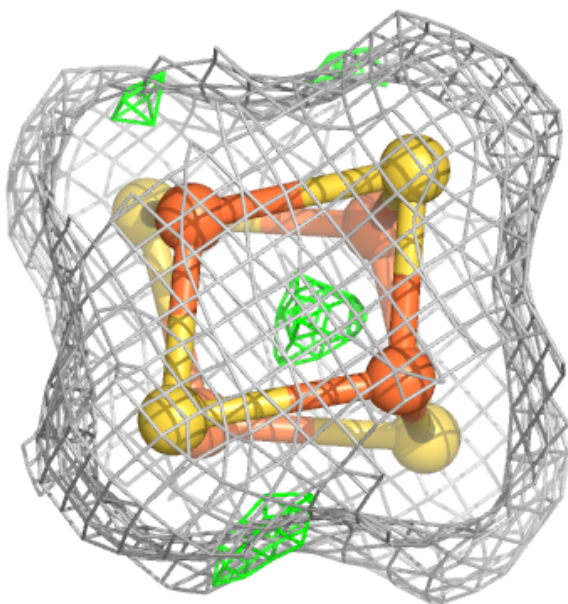
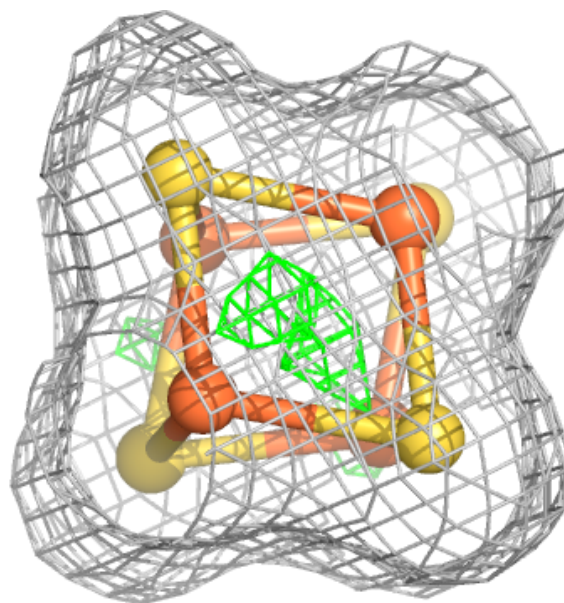
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





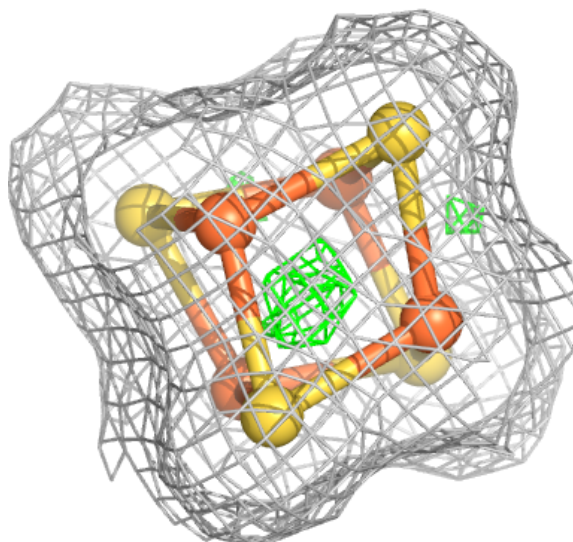
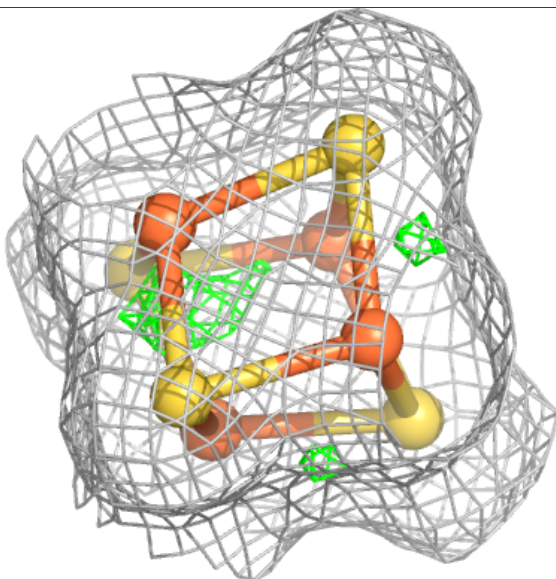
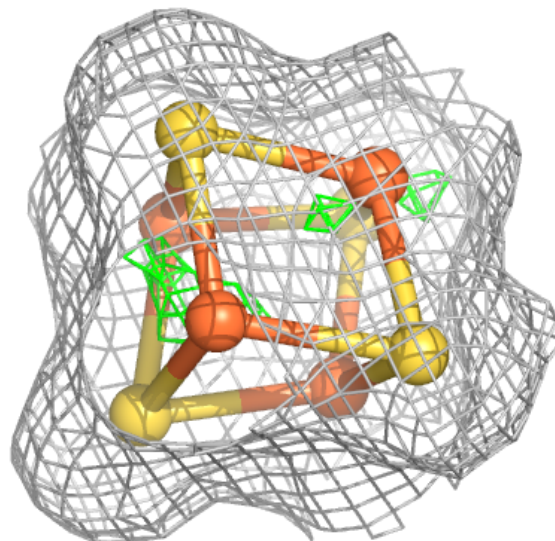
Electron density around SF4 K 1105:

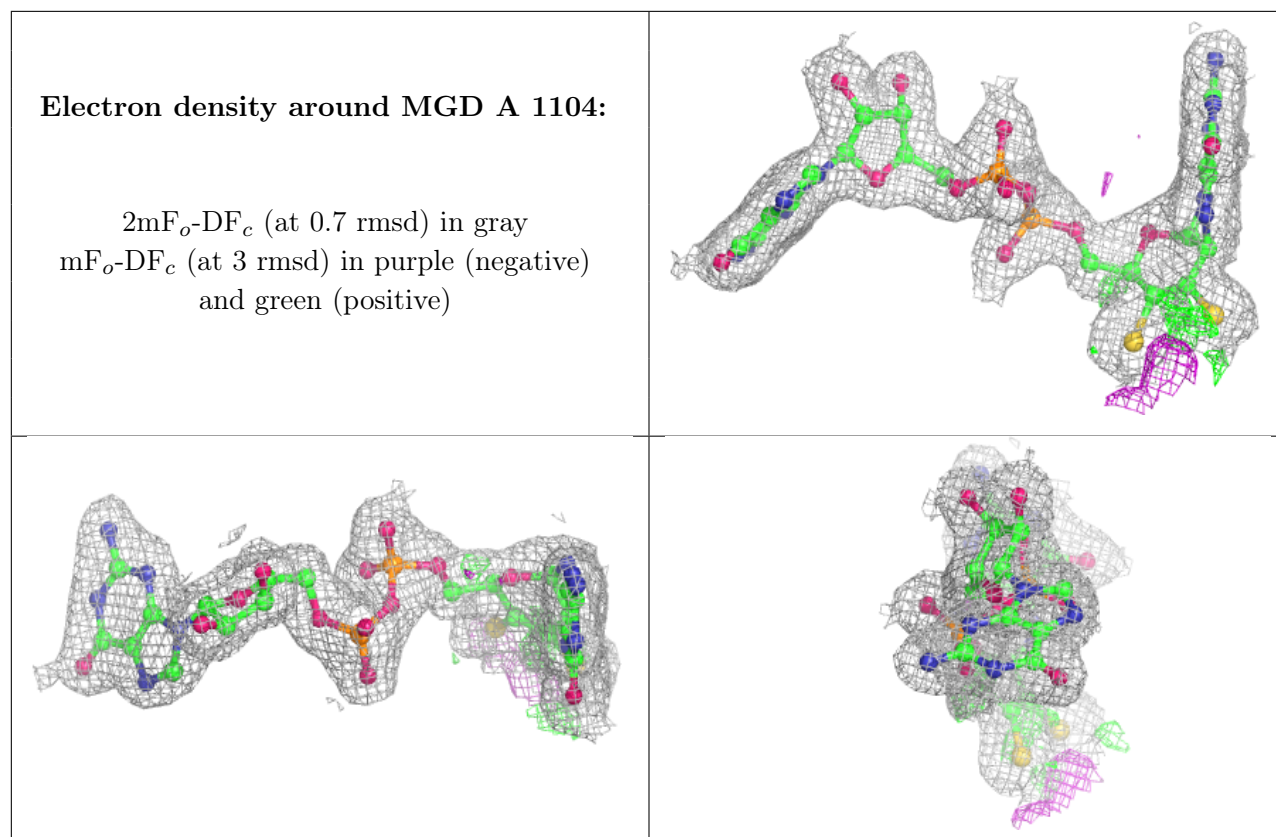
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around SF4 S 304:

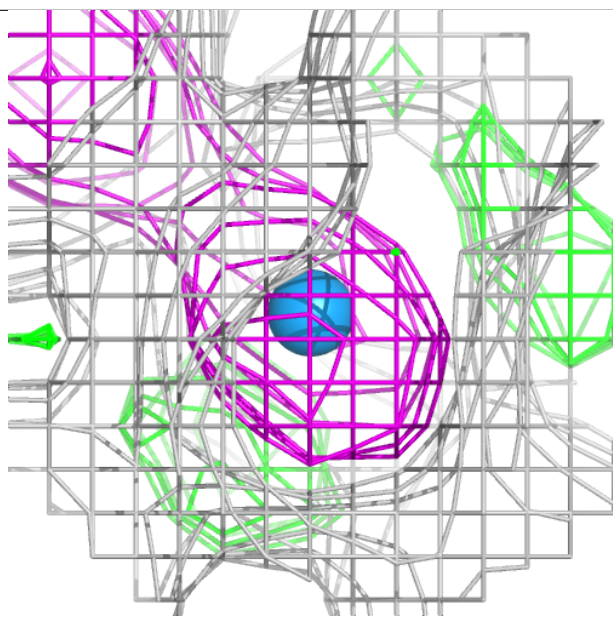
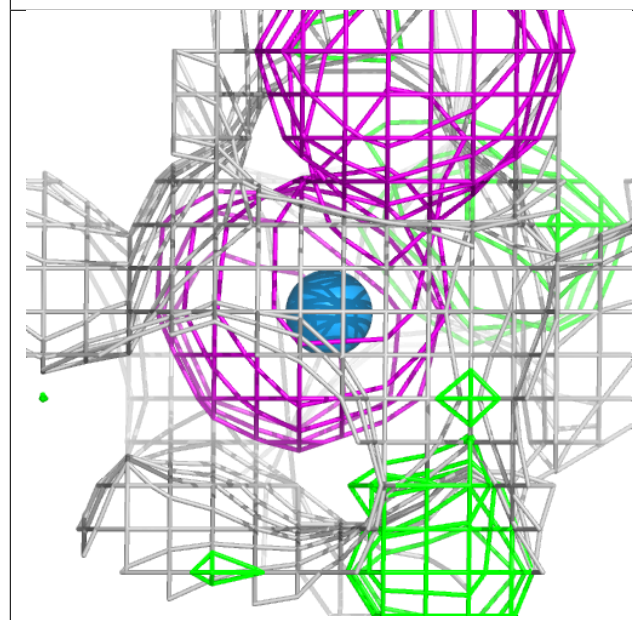
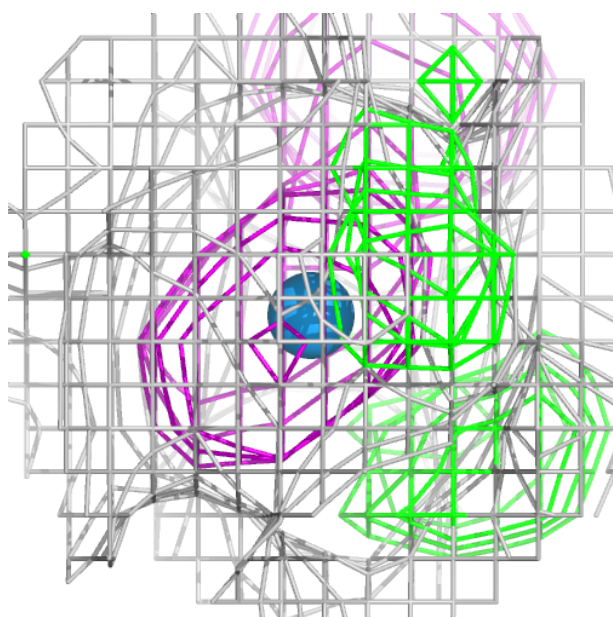
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





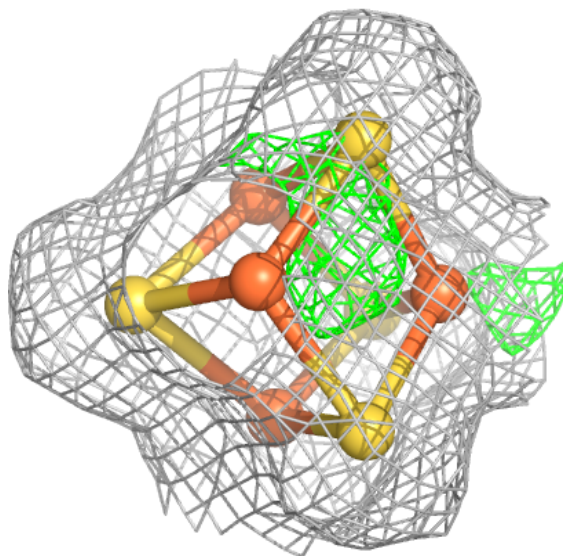
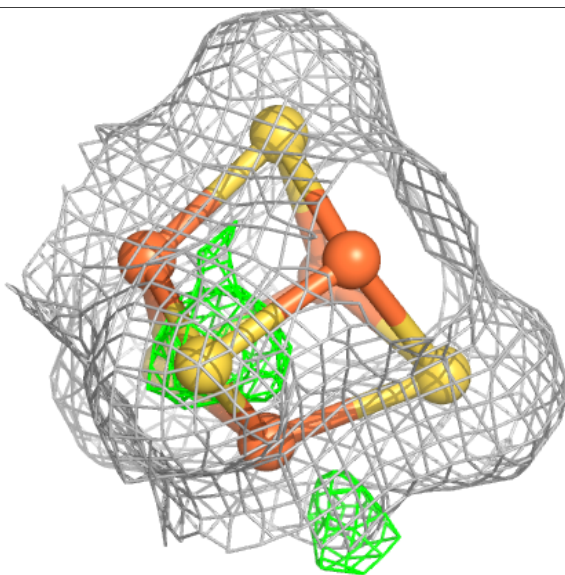
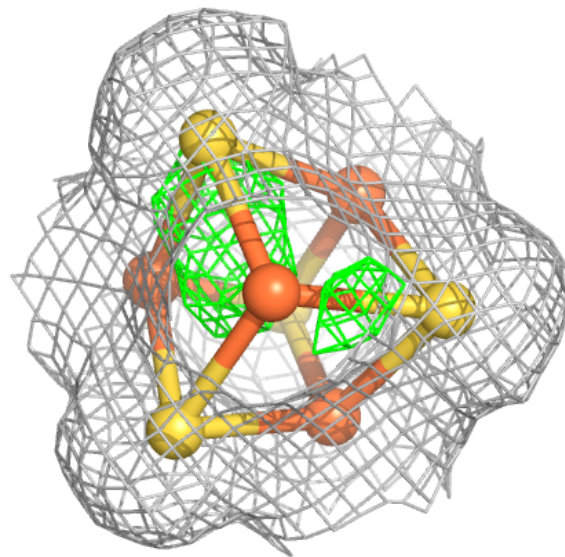
Electron density around W R 1102:

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



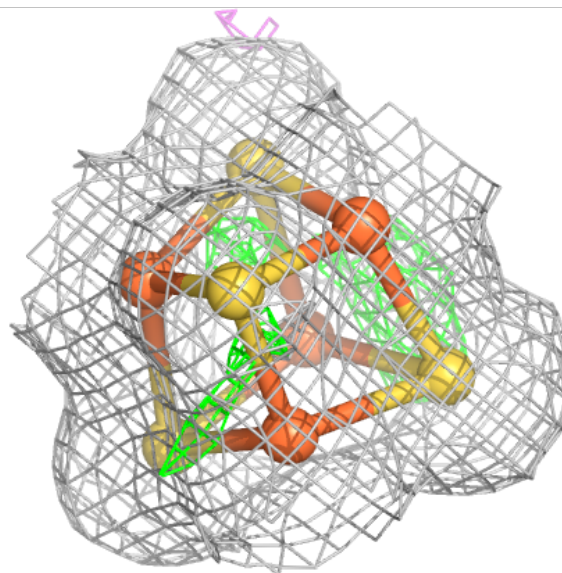
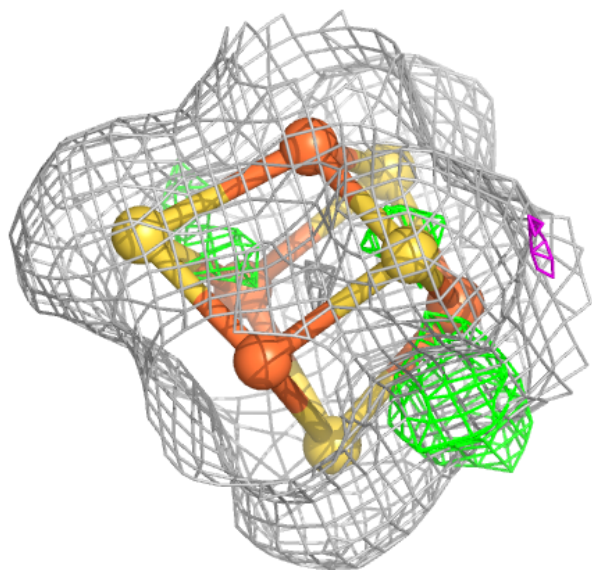
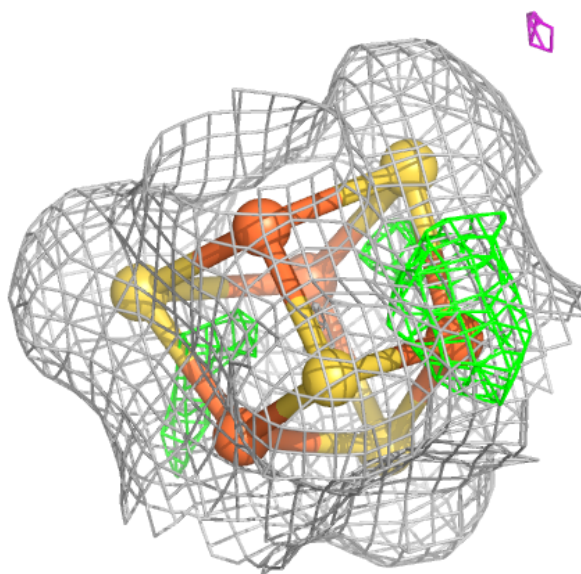
Electron density around SF4 A 1106:

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



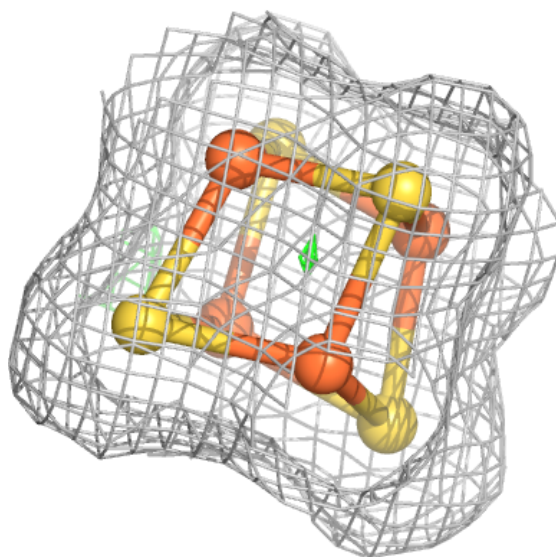
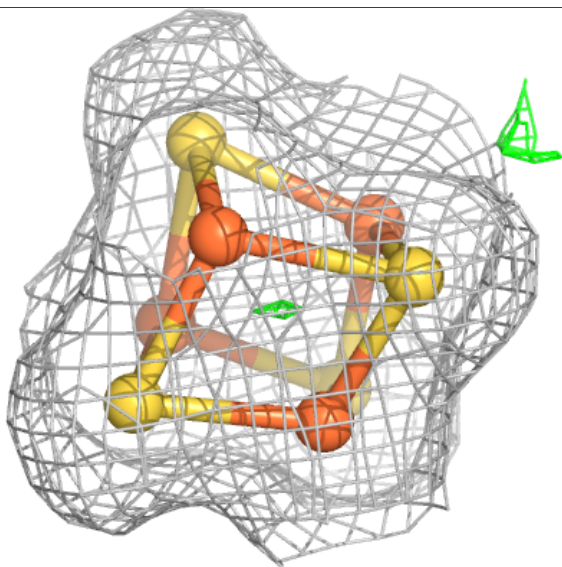
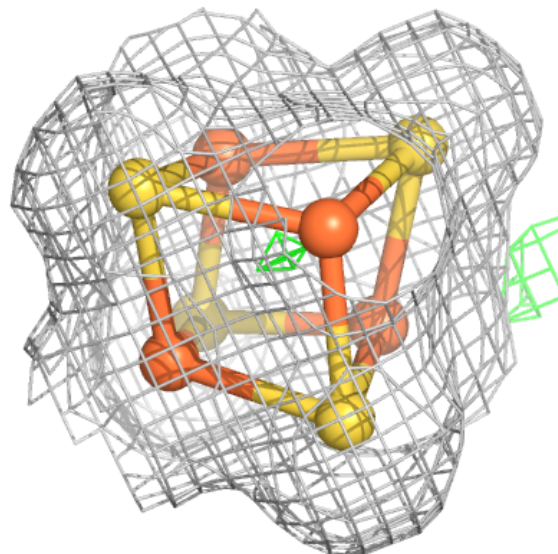
Electron density around SF4 B 301:

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



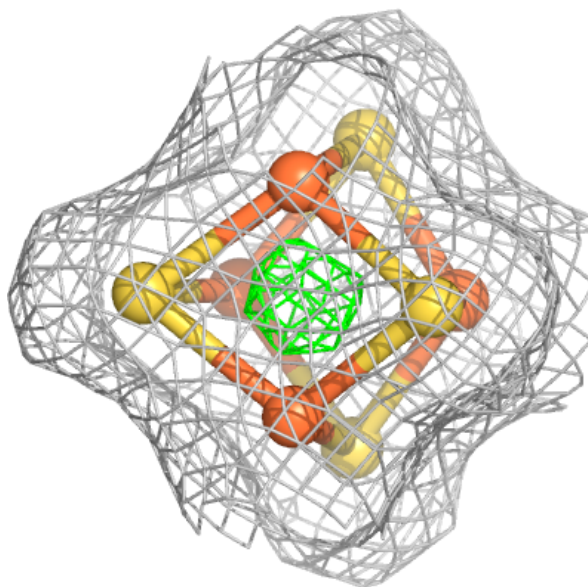
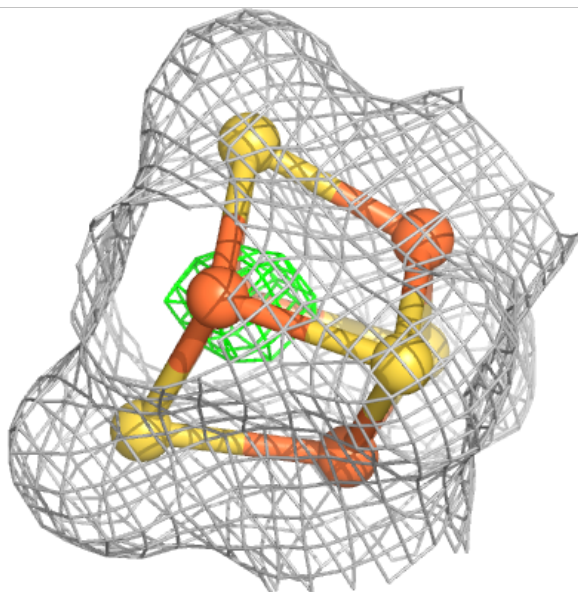
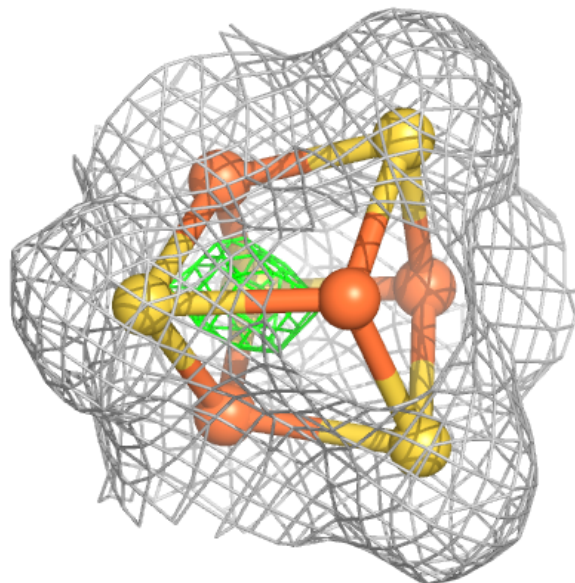
Electron density around SF4 B 302:

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



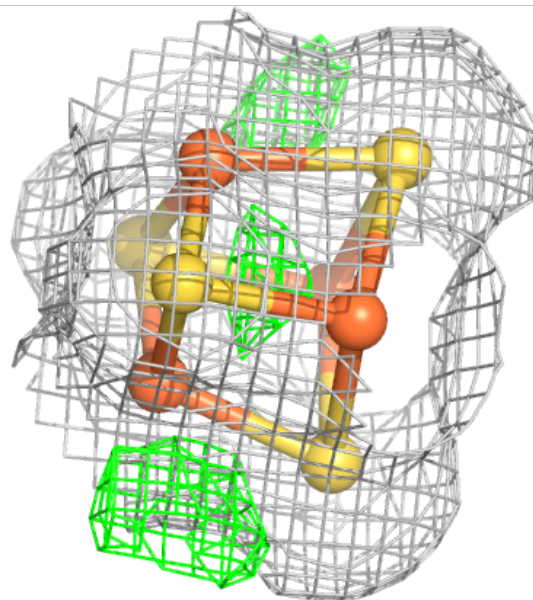
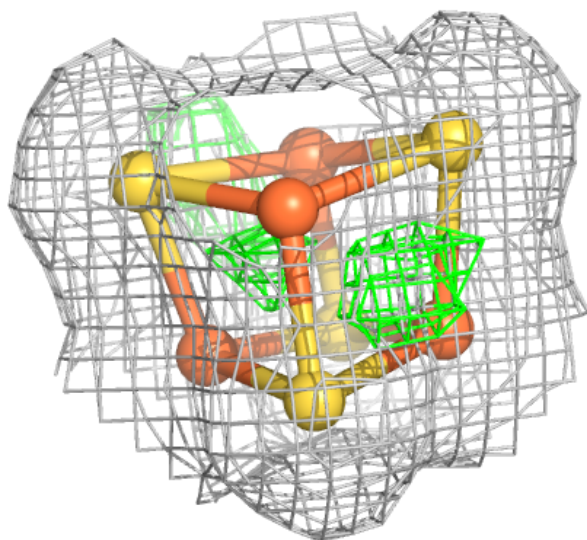
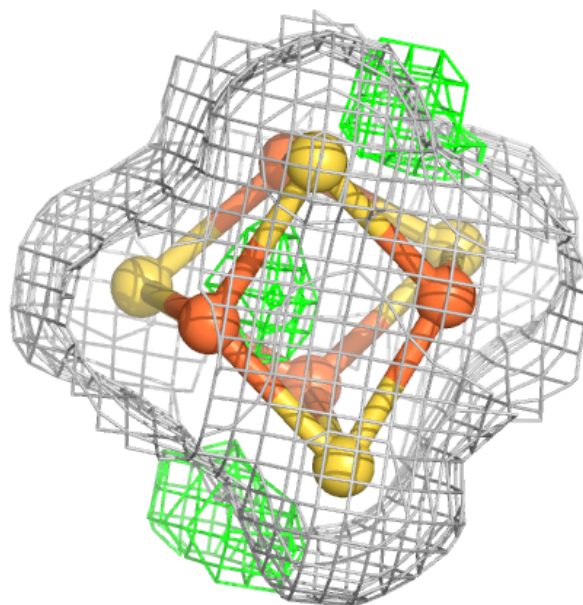
Electron density around SF4 B 303:

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



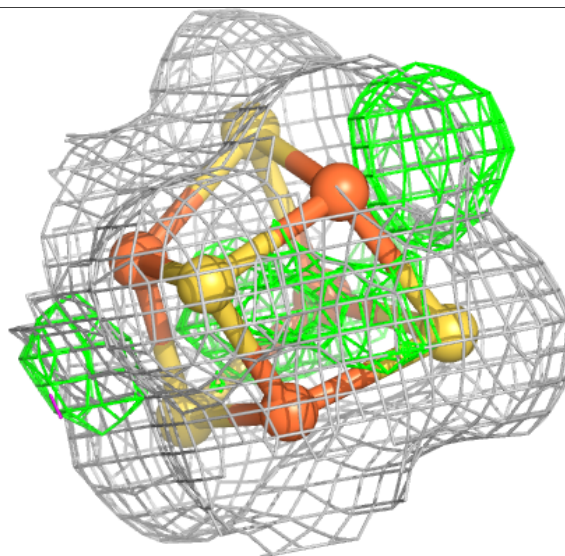
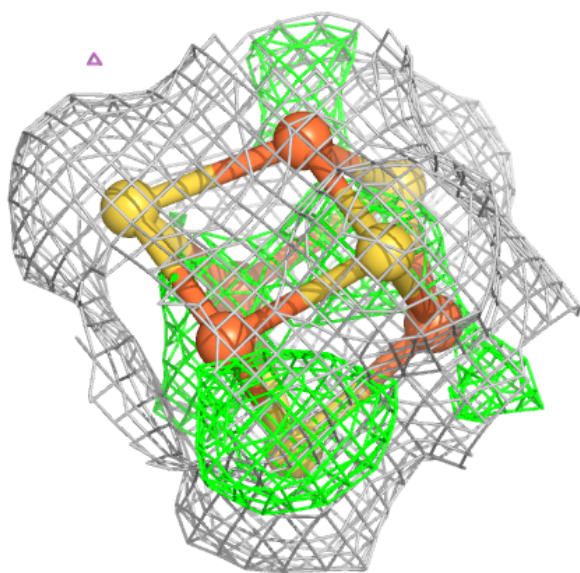
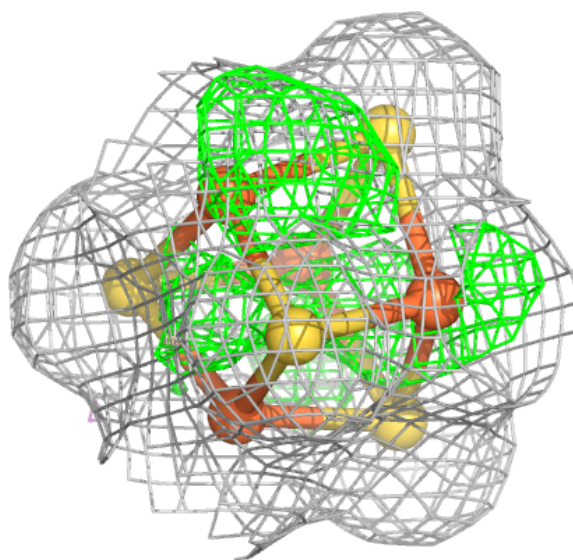
Electron density around SF4 C 1106:

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



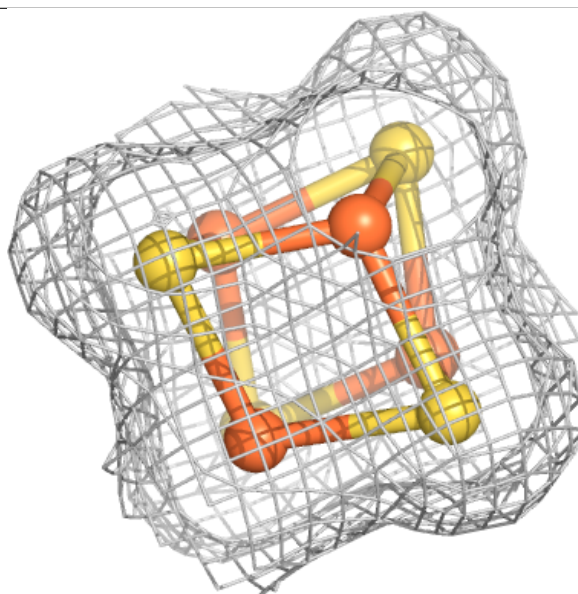
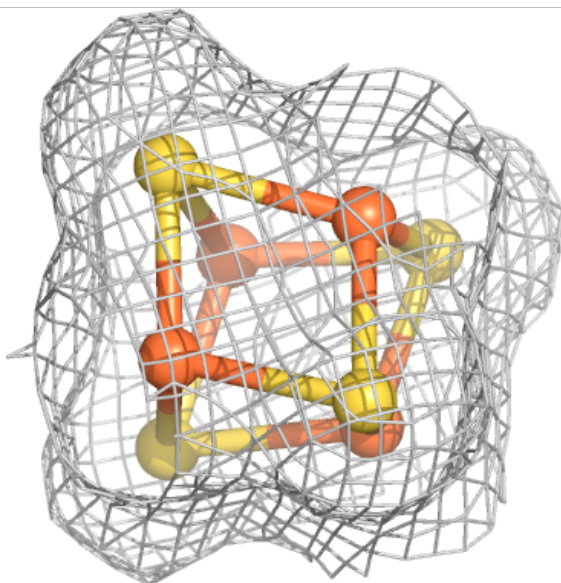
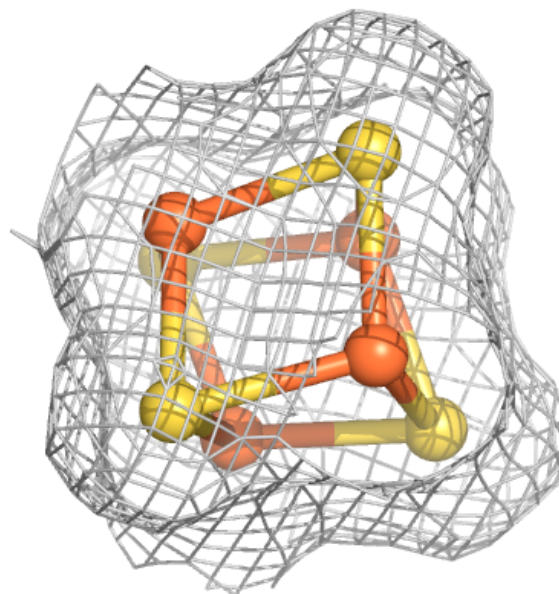
Electron density around SF4 D 402:

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



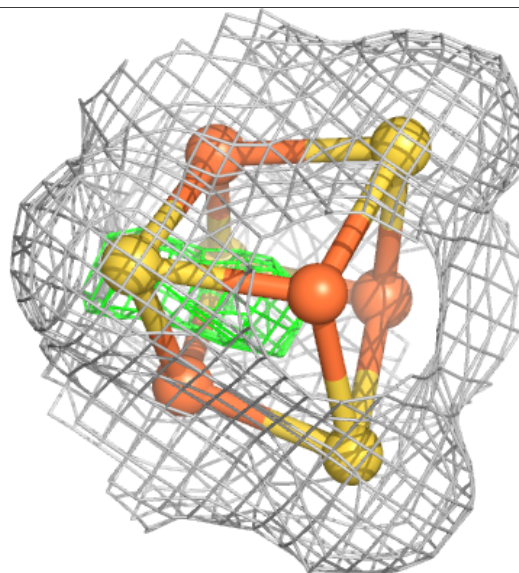
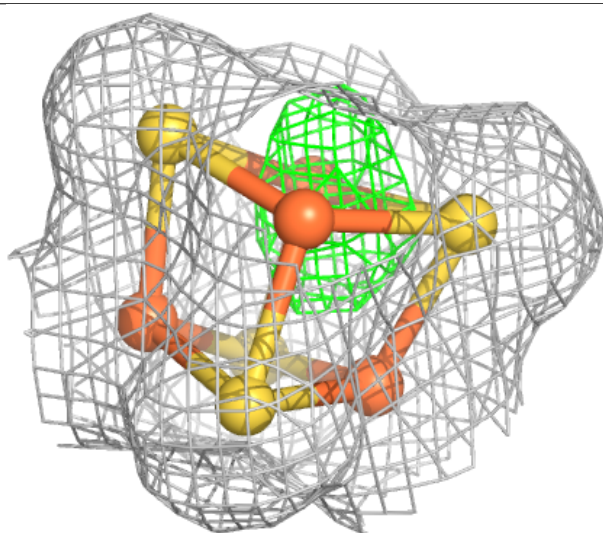
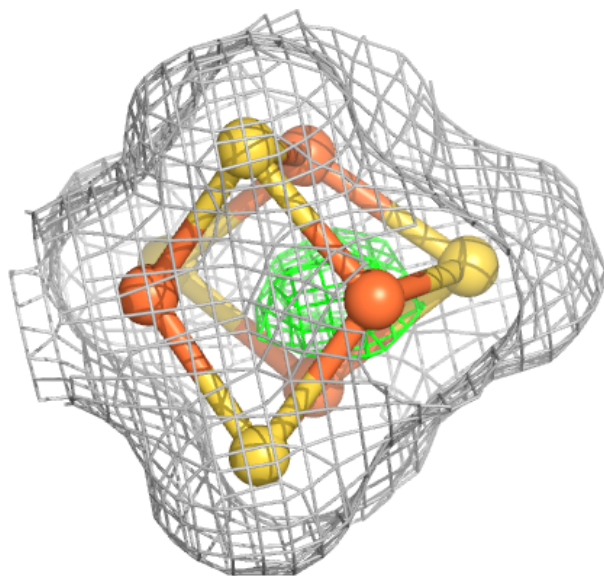
Electron density around SF4 D 403:

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



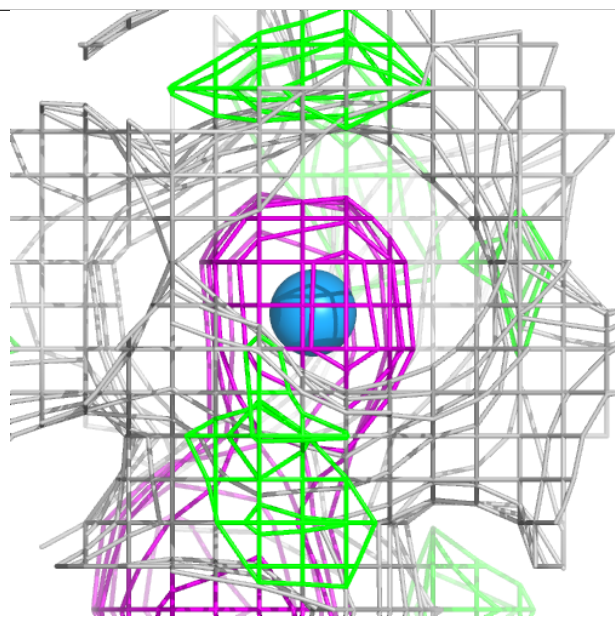
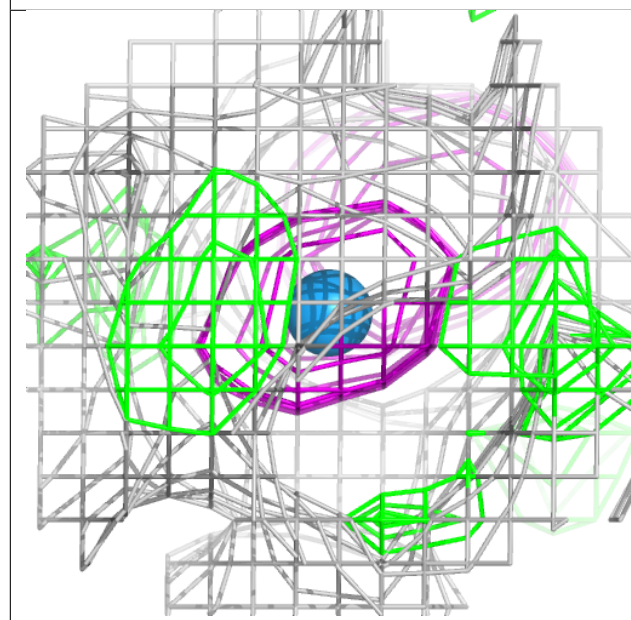
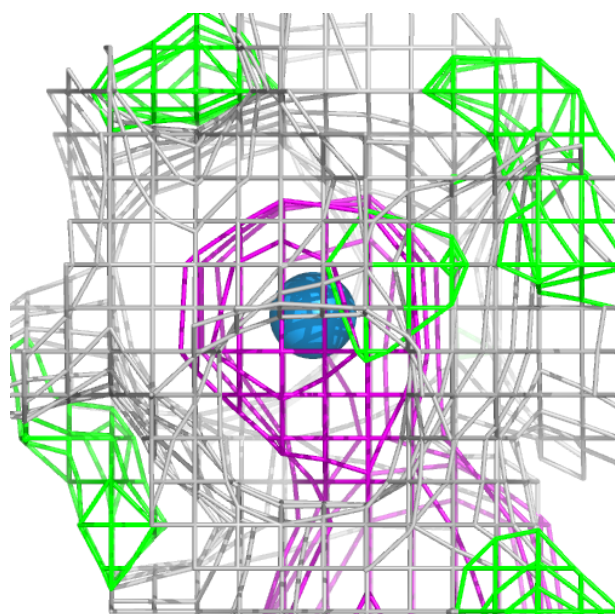
Electron density around SF4 D 404:

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



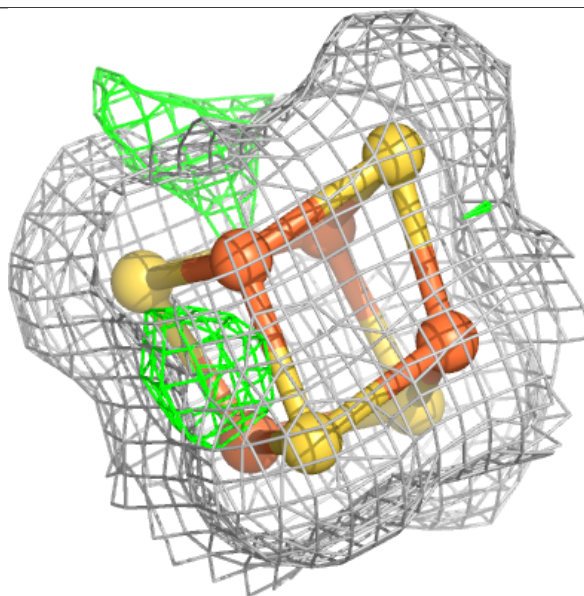
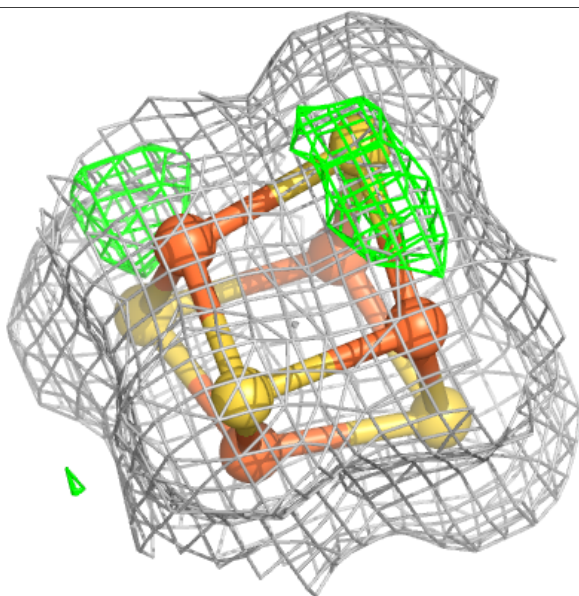
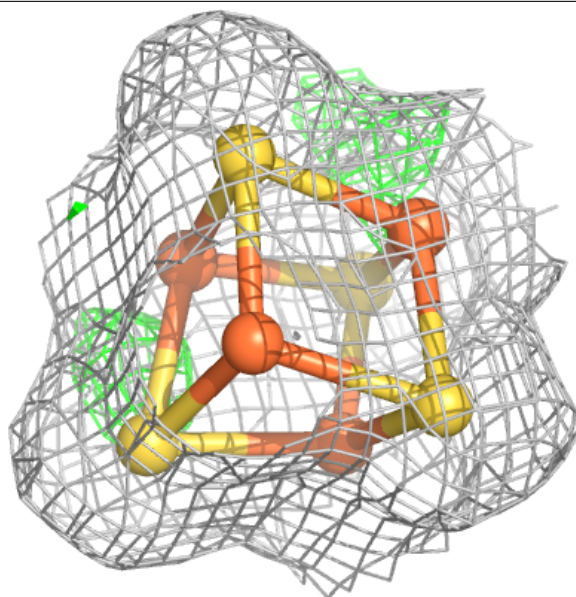
Electron density around W A 1103:

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



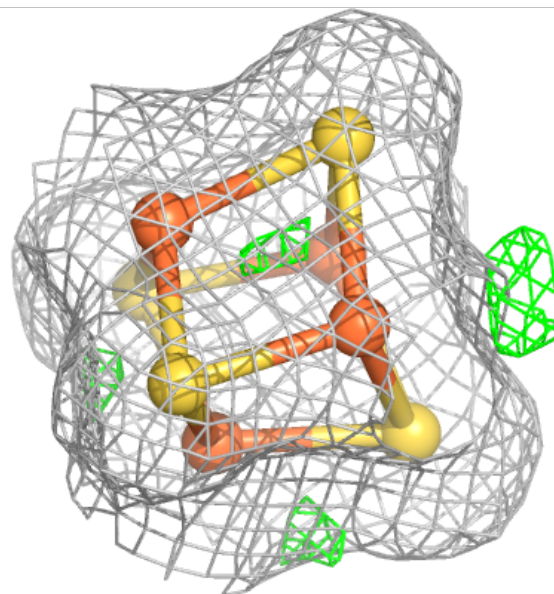
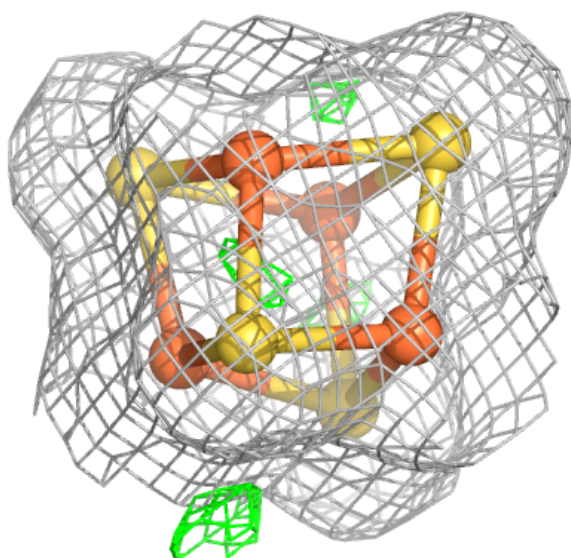
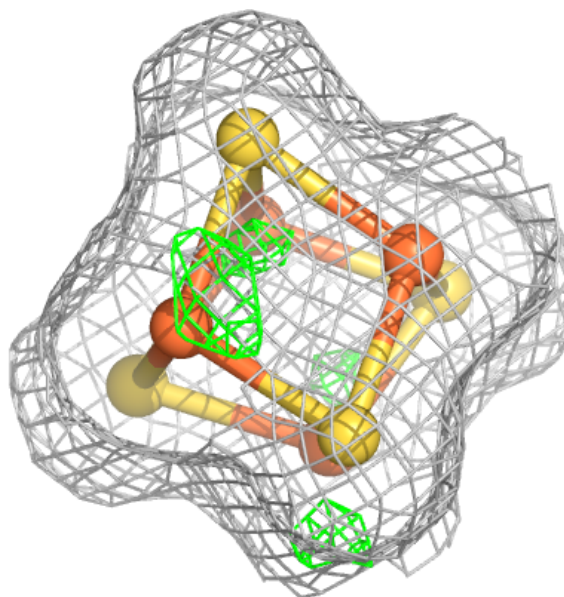
Electron density around SF4 L 303:

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



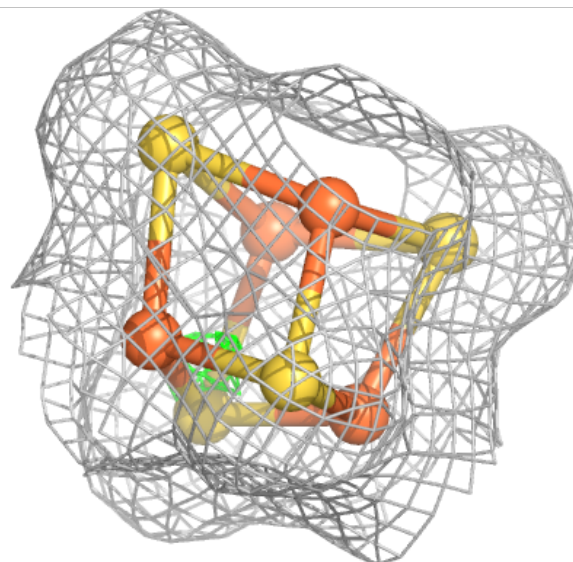
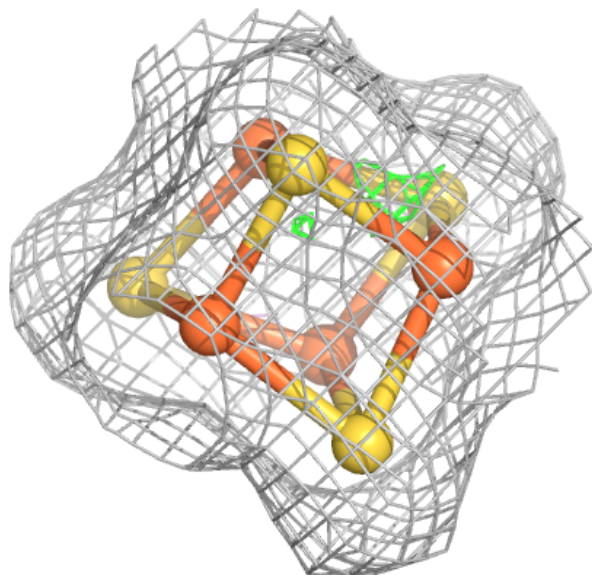
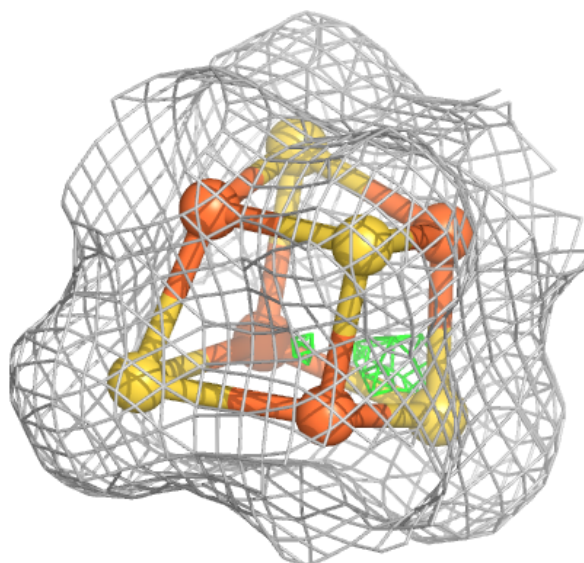
Electron density around SF4 L 304:

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



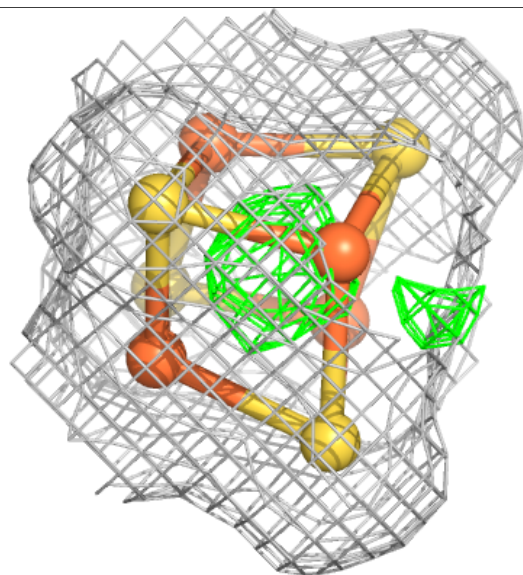
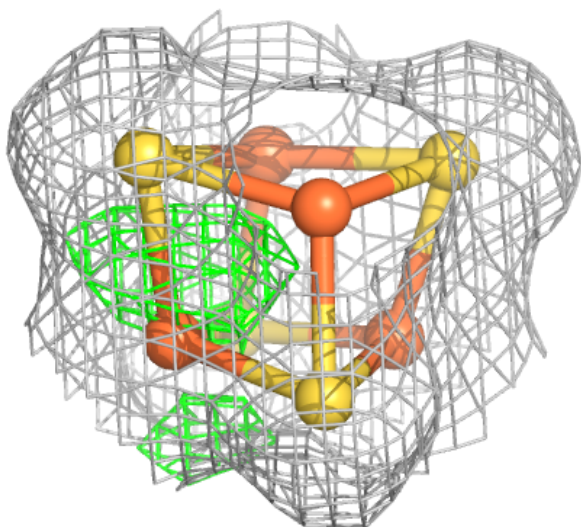
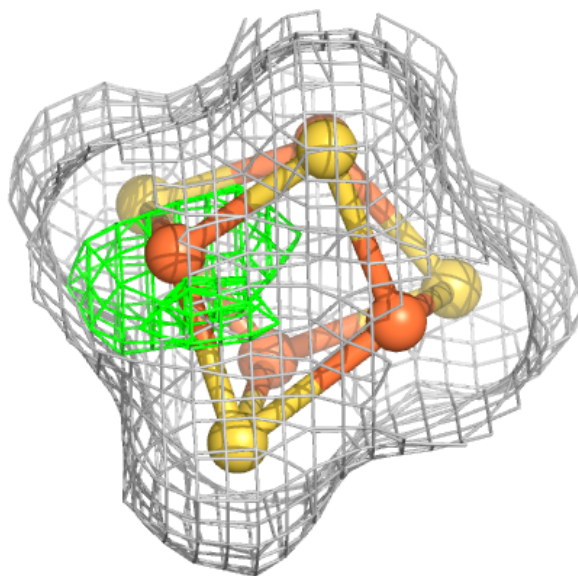
Electron density around SF4 L 305:

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



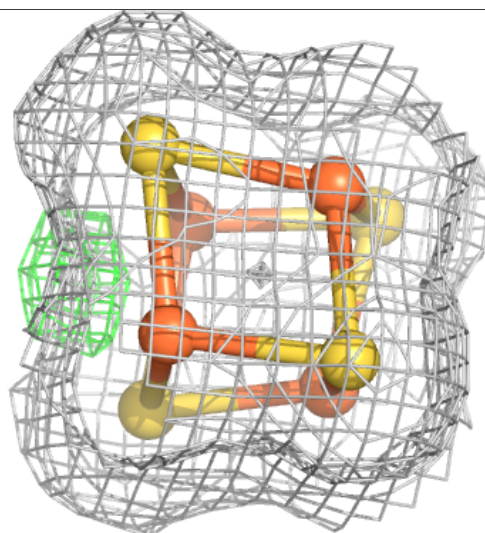
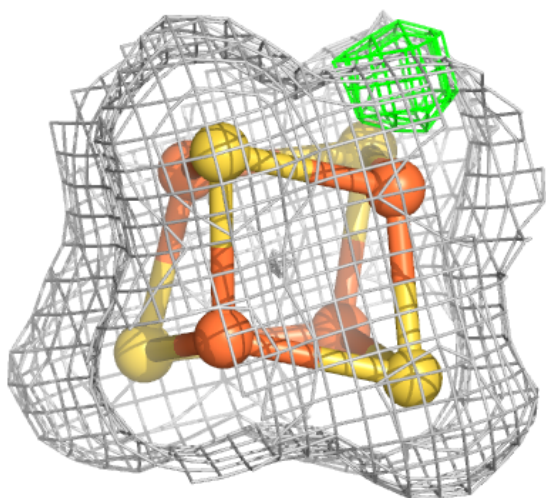
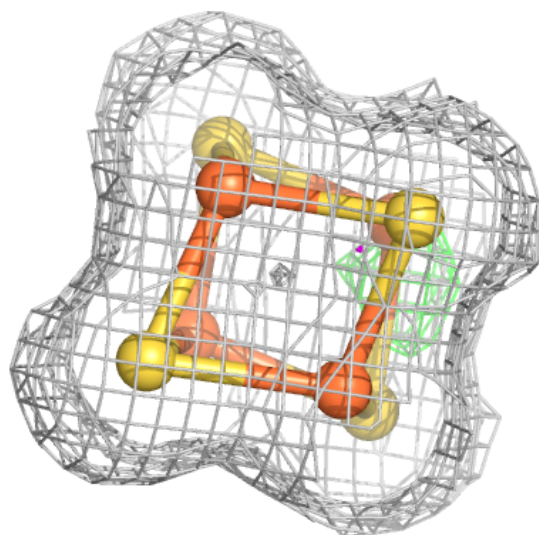
Electron density around SF4 R 1105:

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



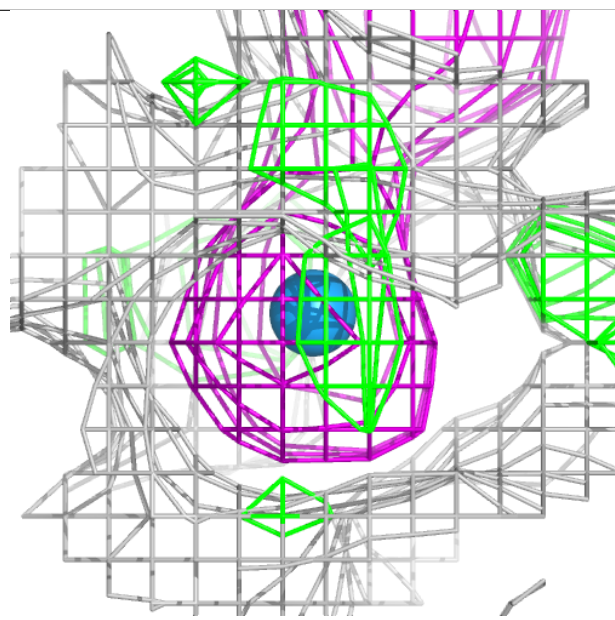
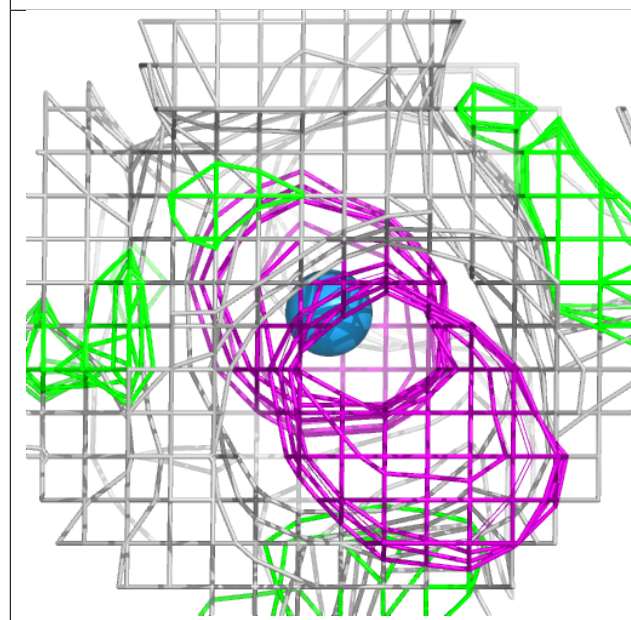
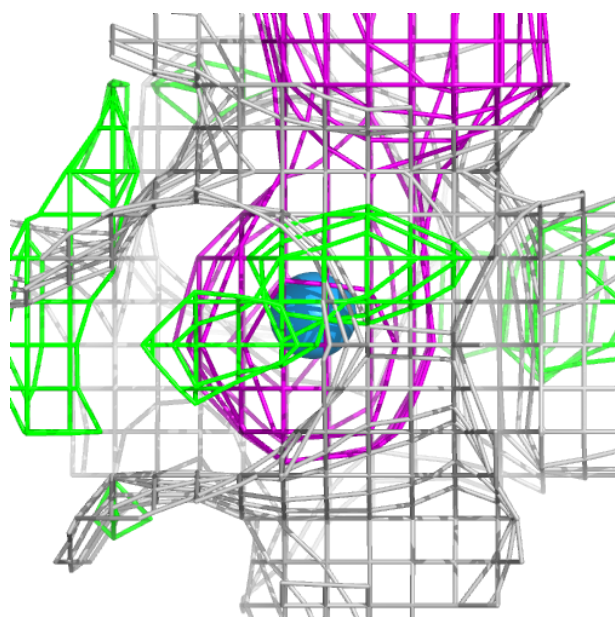
Electron density around SF4 S 303:

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



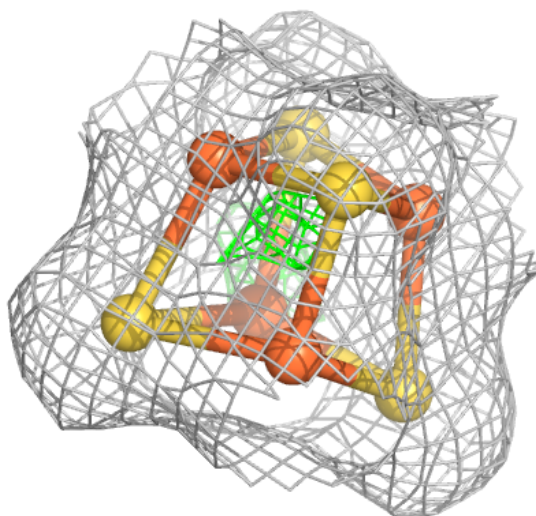
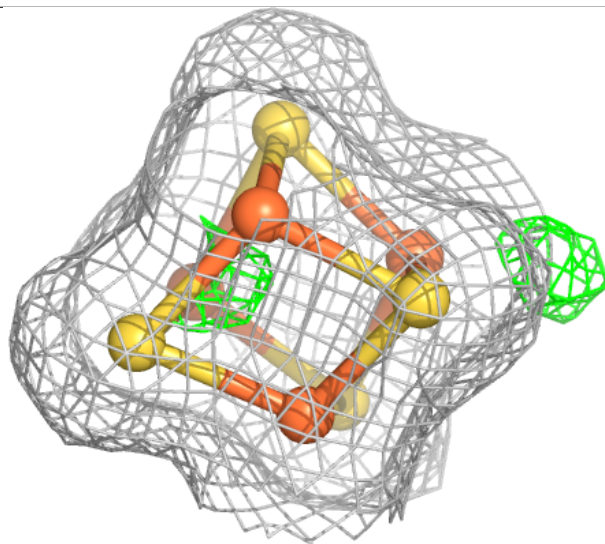
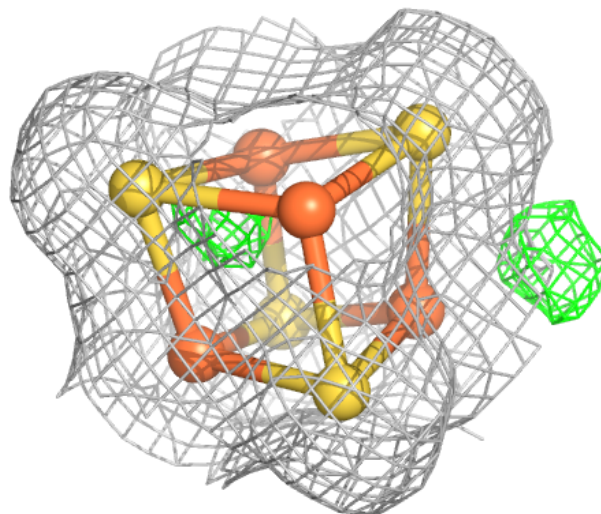
Electron density around W C 1103:

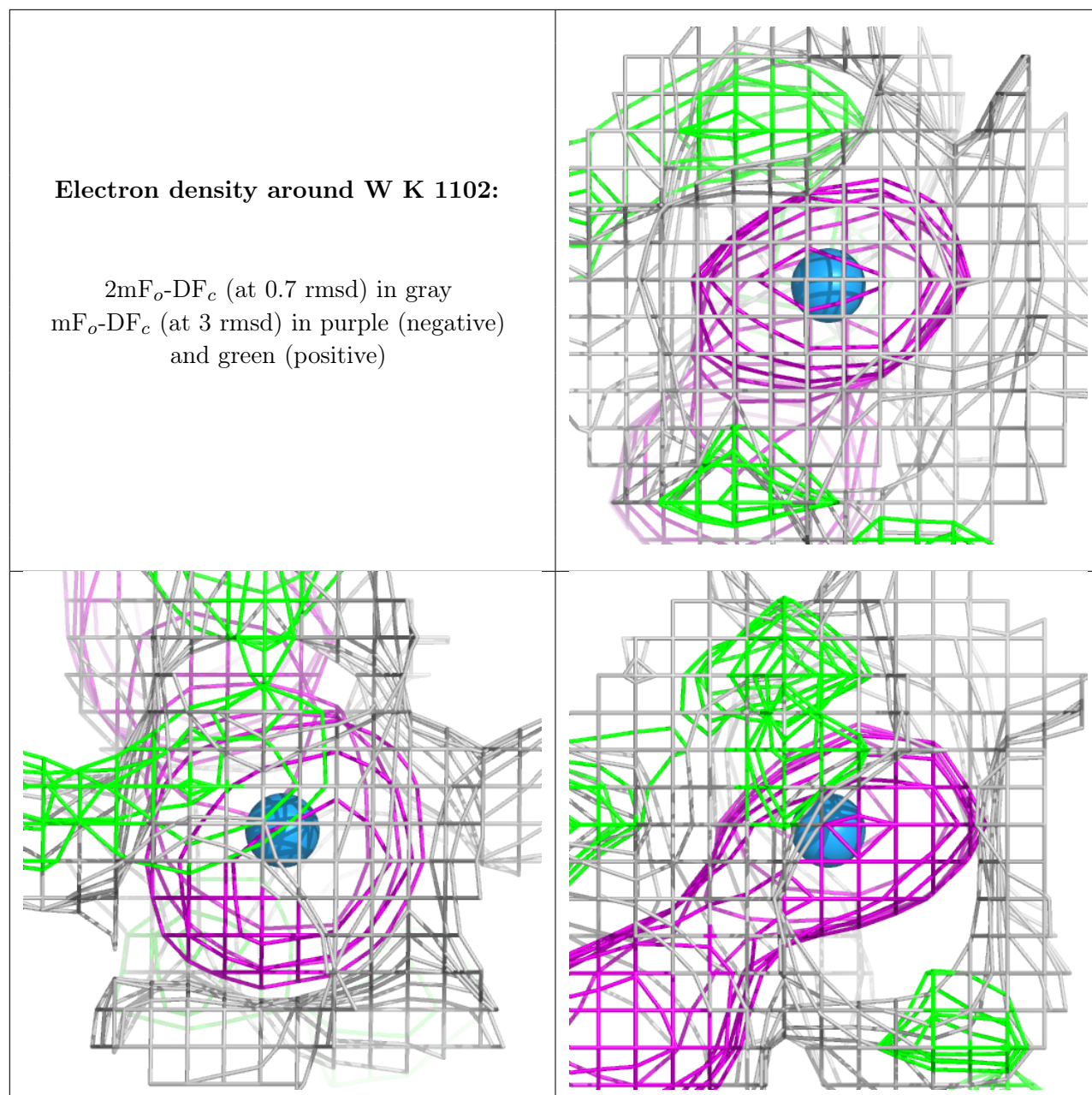
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around SF4 S 305:

$2mF_o-DF_c$ (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.