wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2026 – 08:15 pm BST

PDB ID	:	8CGS / pdb_00008cgs
Title	:	Crystal structure of arsenite oxidase from Alcaligenes faecalis (Af Aio) bound to antimony oxyanion
Authors	:	Engrola, F.; Correia, M.A.S.; Romao, M.J.; Santos-Silva, T.
Deposited on	:	2023-02-06
Resolution	:	1.84 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:

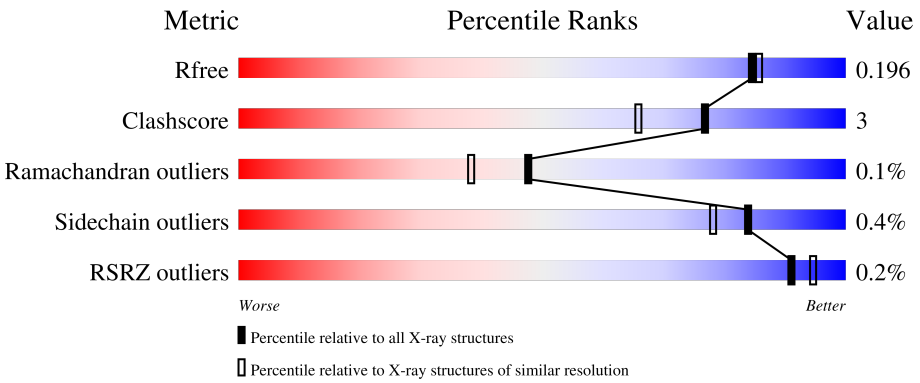
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

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 1.84 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}	180053	1296 (1.84-1.84)
Clashscore	190562	1329 (1.84-1.84)
Ramachandran outliers	187476	1318 (1.84-1.84)
Sidechain outliers	187428	1318 (1.84-1.84)
RSRZ outliers	180081	1296 (1.84-1.84)

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	822	93% 7%
1	C	822	94% 6%
1	E	822	95% 5%
1	G	822	94% 6%
2	B	133	96% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	133	97%
2	F	133	95% 5%
2	H	133	96%

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
13	TRS	A	5017	-	-	X	-
13	TRS	C	5112	-	X	X	-
13	TRS	G	5006	-	X	X	-
15	EDO	E	5010	-	-	X	-
15	EDO	E	5012	-	-	X	-
8	ACT	A	5009	-	-	X	-

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 33900 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 Arsenite oxidase subunit AioA.

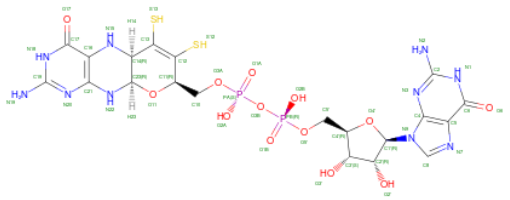
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	822	6509	4096	1153	1219	41	0	7	0
1	C	822	6491	4084	1152	1215	40	0	4	0
1	E	822	6502	4093	1152	1215	42	0	6	0
1	G	822	6484	4080	1147	1217	40	0	3	0

- Molecule 2 is a protein called AioB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	133	998	625	167	197	9	0	2	0
2	D	133	1001	627	167	198	9	0	2	0
2	F	133	995	623	167	196	9	0	1	0
2	H	133	995	623	167	196	9	0	1	0

- Molecule 3 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 (CCD ID: MGD) (formula: $C_{20}H_{26}N_{10}O_{13}P_2S_2$) (labeled as "Ligand of Interest" by depositor).

MGD



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
3	A	1	47	20	10	13	2	2	0	0
3	A	1	47	20	10	13	2	2	0	0
3	C	1	47	20	10	13	2	2	0	0
3	C	1	47	20	10	13	2	2	0	0
3	E	1	47	20	10	13	2	2	0	0
3	E	1	47	20	10	13	2	2	0	0
3	G	1	47	20	10	13	2	2	0	0
3	G	1	47	20	10	13	2	2	0	0

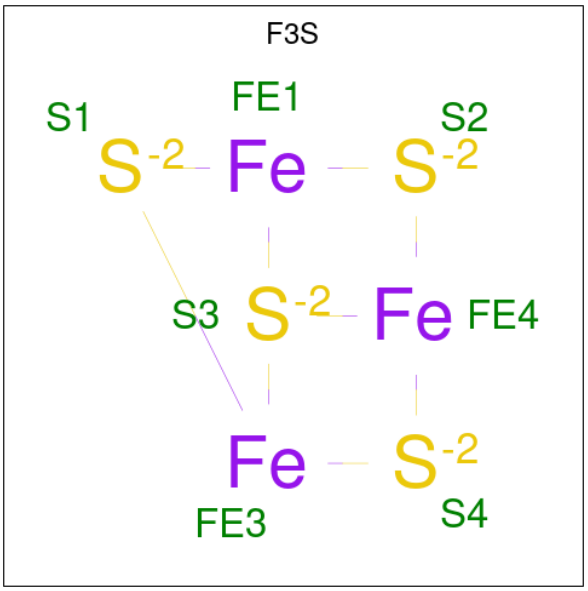
- Molecule 4 is OXYGEN ATOM (CCD ID: O) (formula: O).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	1	1	1	0	0
4	C	1	1	1	0	0
4	E	1	1	1	0	0
4	G	1	1	1	0	0

- Molecule 5 is MOLYBDENUM(IV) ION (CCD ID: 4MO) (formula: Mo) (labeled as "Ligand of Interest" by depositor).

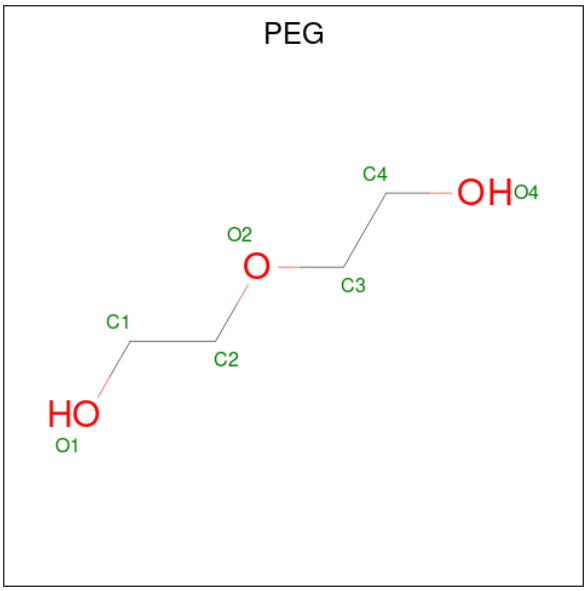
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mo 1 1	0	0
5	C	1	Total Mo 1 1	0	0
5	E	1	Total Mo 1 1	0	0
5	G	1	Total Mo 1 1	0	0

- Molecule 6 is FE3-S4 CLUSTER (CCD ID: F3S) (formula: Fe_3S_4) (labeled as "Ligand of Interest" by depositor).



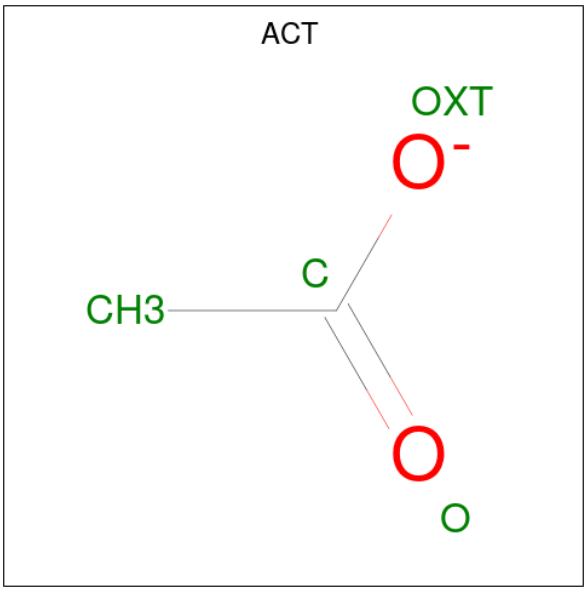
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Fe S 7 3 4	0	0
6	C	1	Total Fe S 7 3 4	0	0
6	E	1	Total Fe S 7 3 4	0	0
6	G	1	Total Fe S 7 3 4	0	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



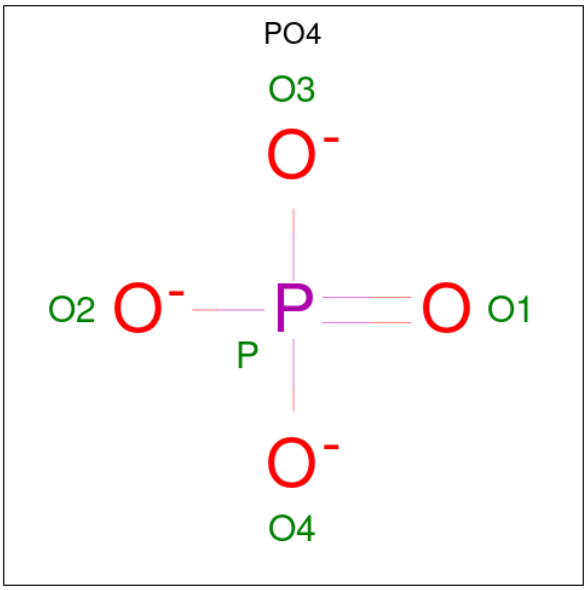
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
7	A	1	7	4	3		0	0
			Total	C	O			
7	A	1	7	4	3		0	0
			Total	C	O			
7	A	1	7	4	3		0	0
			Total	C	O			
7	C	1	7	4	3		0	0
			Total	C	O			
7	E	1	7	4	3		0	0
			Total	C	O			
7	G	1	7	4	3		0	0

- Molecule 8 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



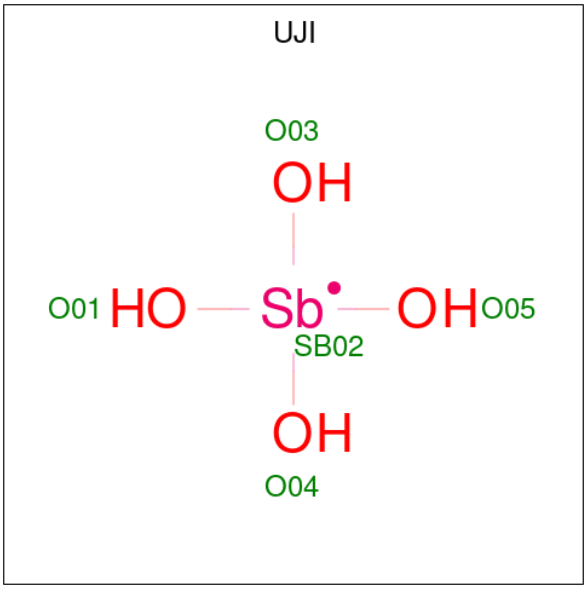
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
8	A	1	4	2	2		0	0
			Total	C	O			
8	A	1	4	2	2		0	0
			Total	C	O			
8	G	1	4	2	2		0	0
			Total	C	O			
8	G	1	4	2	2		0	0

- Molecule 9 is PHOSPHATE ION (CCD ID: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
9	A	1	5	4	1	0	0
9	C	1	5	4	1	0	0
9	C	1	5	4	1	0	0
9	E	1	5	4	1	0	0
9	E	1	5	4	1	0	0
9	G	1	5	4	1	0	0

- Molecule 10 is tetrakis(oxidanyl)antimony (CCD ID: UJI) (formula: H_4O_4Sb) (labeled as "Ligand of Interest" by depositor).

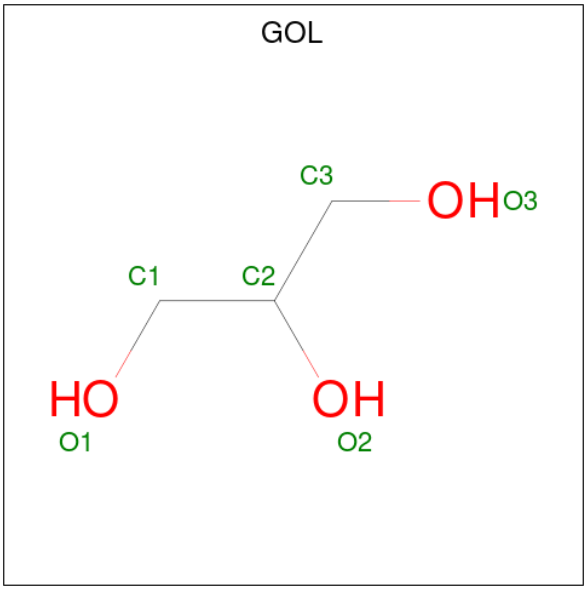


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	Sb		
10	A	1	5	4	1	0	0
10	C	1	5	4	1	0	0
10	E	1	5	4	1	0	0
10	G	1	5	4	1	0	0

- Molecule 11 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

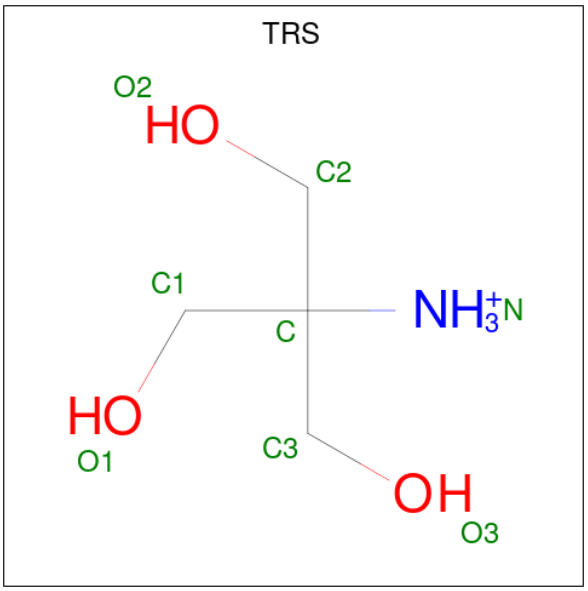
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
11	A	2	2	2	0	0
			Total	Cl		
11	C	1	1	1	0	0
			Total	Cl		
11	E	2	2	2	0	0

- Molecule 12 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



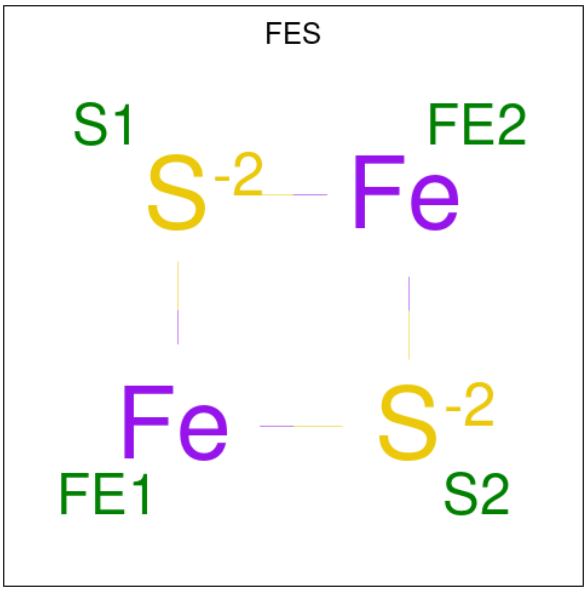
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
12	A	1	6	3	3	0	0
			Total	C	O		
12	A	1	6	3	3	0	0
			Total	C	O		
12	C	1	6	3	3	0	0
			Total	C	O		
12	E	1	6	3	3	0	0

- Molecule 13 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (CCD ID: TRS) (formula: $C_4H_{12}NO_3$).



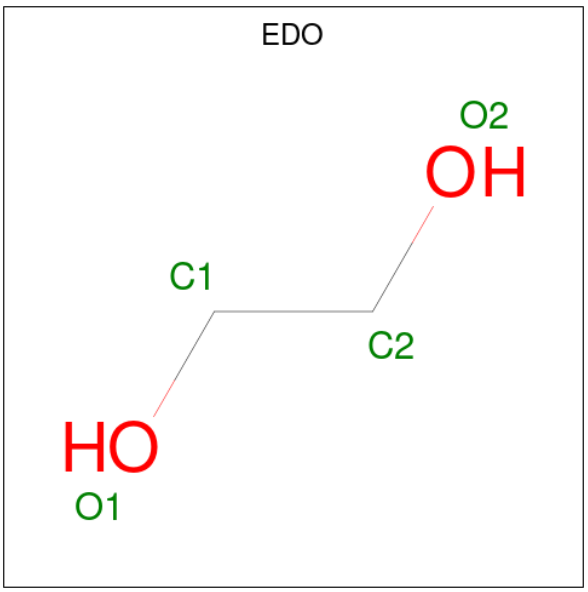
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
13	A	1	8	4	1	3	0	0
			Total	C	N	O		
13	C	1	8	4	1	3	0	0
			Total	C	N	O		
13	G	1	8	4	1	3	0	0

- Molecule 14 is FE2/S2 (INORGANIC) CLUSTER (CCD ID: FES) (formula: Fe_2S_2) (labeled as "Ligand of Interest" by depositor).



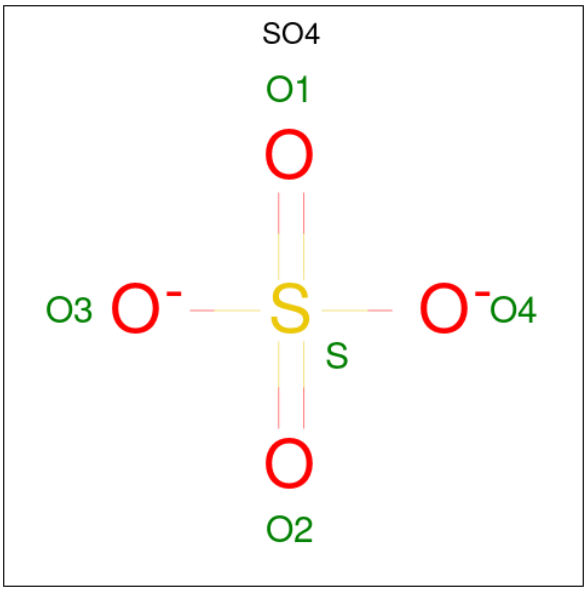
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
14	B	1	4	2	2	0	0
14	D	1	4	2	2	0	0
14	F	1	4	2	2	0	0
14	H	1	4	2	2	0	0

- Molecule 15 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
15	E	1	4	2	2	0	0
15	E	1	4	2	2	0	0

- Molecule 16 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	E	1	Total O S 5 4 1	0	0

- Molecule 17 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	G	1	Total Na 1 1	0	0

- Molecule 18 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	A	738	Total O 738 738	0	0
18	B	116	Total O 116 116	0	0
18	C	726	Total O 726 726	0	0
18	D	121	Total O 121 121	0	0
18	E	692	Total O 692 692	0	0
18	F	115	Total O 115 115	0	0
18	G	687	Total O 687 687	0	0

Continued on next page...

Continued from previous page...

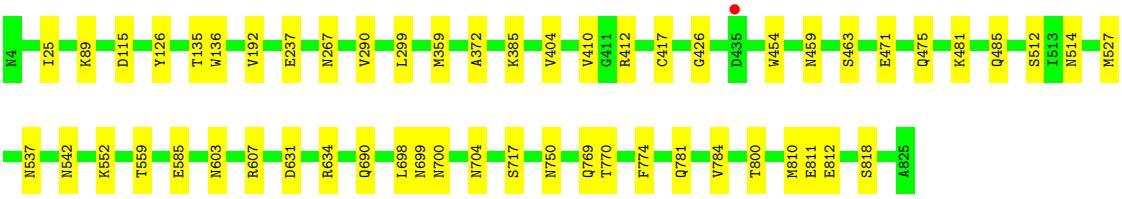
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
18	H	127	127	127	0	0

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.

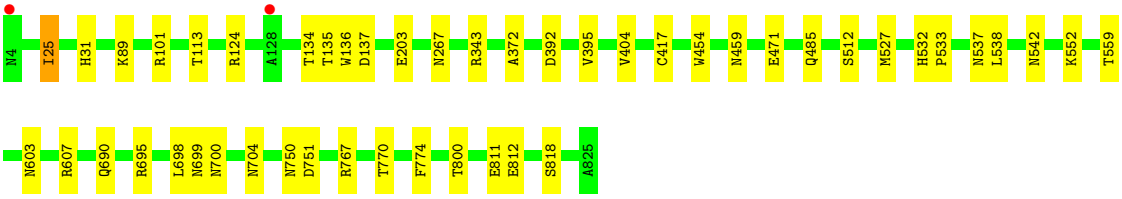
• Molecule 1: Arsenite oxidase subunit AioA

Chain A: 93% 7%



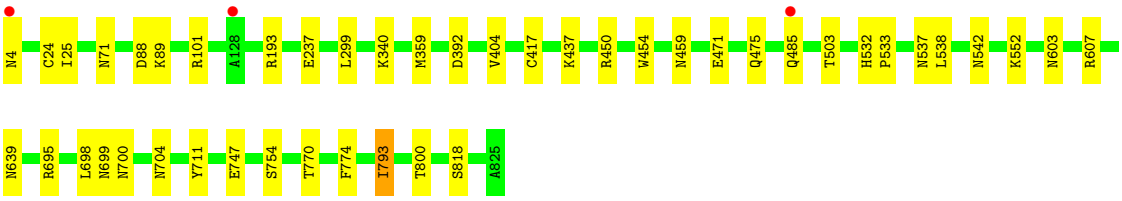
• Molecule 1: Arsenite oxidase subunit AioA

Chain C: 94% 6%



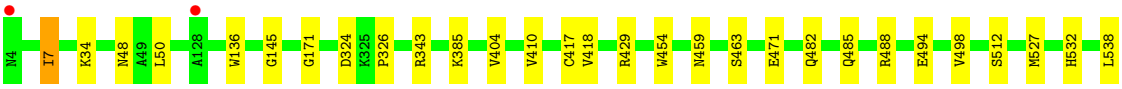
• Molecule 1: Arsenite oxidase subunit AioA

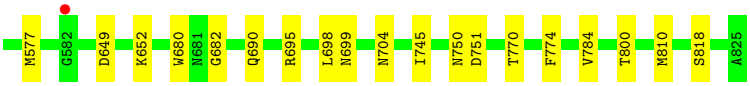
Chain E: 95% 5%



• Molecule 1: Arsenite oxidase subunit AioA

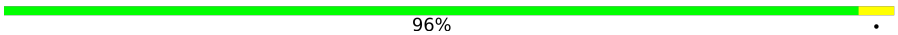
Chain G: 94% 6%

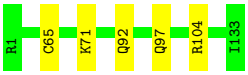




- Molecule 2: AioB

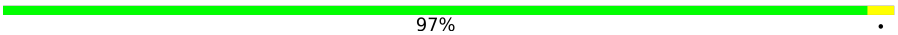
Chain B:

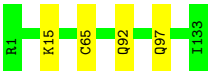




- Molecule 2: AioB

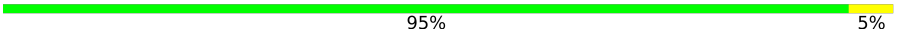
Chain D:

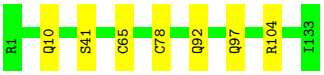




- Molecule 2: AioB

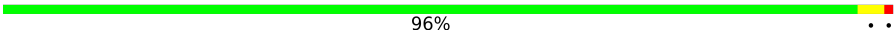
Chain F:





- Molecule 2: AioB

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	90.38Å 109.21Å 117.34Å 97.71° 90.06° 96.28°	Depositor
Resolution (Å)	116.27 – 1.84 116.27 – 1.84	Depositor EDS
% Data completeness (in resolution range)	95.3 (116.27-1.84) 95.3 (116.27-1.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.154 , 0.189 0.165 , 0.196	Depositor DCC
R_{free} test set	18174 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å^2)	18.5	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 42.6	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	0.96	EDS
Total number of atoms	33900	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.76% 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: PO4, CL, SO4, 4MO, O, TRS, PEG, F3S, EDO, GOL, FES, ACT, NA, UJI, MGD

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	1.10	3/6688 (0.0%)	0.97	1/9063 (0.0%)
1	C	1.12	2/6664 (0.0%)	0.99	0/9032
1	E	1.09	2/6681 (0.0%)	0.98	2/9053 (0.0%)
1	G	1.11	2/6654 (0.0%)	0.98	1/9020 (0.0%)
2	B	1.10	0/1025	1.02	1/1396 (0.1%)
2	D	1.11	0/1028	1.02	0/1400
2	F	1.10	0/1019	1.04	2/1388 (0.1%)
2	H	1.11	0/1019	1.06	1/1388 (0.1%)
All	All	1.11	9/30778 (0.0%)	0.99	8/41740 (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	E	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	488	ARG	CA-C	-7.00	1.44	1.52
1	A	514	ASN	C-O	6.77	1.31	1.23
1	E	193	ARG	C-O	6.17	1.32	1.23
1	A	192	VAL	C-O	5.81	1.30	1.24
1	A	290	VAL	C-O	-5.46	1.18	1.24

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	71	LYS	N-CA-C	5.60	118.11	111.33
1	E	340	LYS	N-CA-C	5.47	117.05	111.14
2	H	104	ARG	N-CA-CB	5.34	118.89	110.23
2	F	78	CYS	CA-C-N	-5.31	114.84	121.00
2	F	78	CYS	C-N-CA	-5.31	114.84	121.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	4	ASN	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	6509	0	6308	46	0
1	C	6491	0	6287	42	0
1	E	6502	0	6307	28	2
1	G	6484	0	6270	32	2
2	B	998	0	984	6	0
2	D	1001	0	985	5	0
2	F	995	0	979	7	0
2	H	995	0	979	7	0
3	A	94	0	44	2	0
3	C	94	0	44	2	0
3	E	94	0	44	3	0
3	G	94	0	44	1	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
6	A	7	0	0	0	0
6	C	7	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	E	7	0	0	0	0
6	G	7	0	0	0	0
7	A	21	0	30	5	0
7	C	7	0	10	1	0
7	E	7	0	10	2	0
7	G	7	0	10	0	0
8	A	8	0	6	3	0
8	G	8	0	6	0	0
9	A	5	0	0	0	0
9	C	10	0	0	0	0
9	E	10	0	0	1	0
9	G	5	0	0	0	0
10	A	5	0	0	0	0
10	C	5	0	0	1	0
10	E	5	0	0	0	0
10	G	5	0	0	0	0
11	A	2	0	0	1	0
11	C	1	0	0	0	0
11	E	2	0	0	1	0
12	A	12	0	16	0	0
12	C	6	0	8	0	0
12	E	6	0	8	2	0
13	A	8	0	11	14	0
13	C	8	0	12	7	0
13	G	8	0	12	7	0
14	B	4	0	0	0	0
14	D	4	0	0	0	0
14	F	4	0	0	0	0
14	H	4	0	0	0	0
15	E	8	0	8	9	0
16	E	5	0	0	0	0
17	G	1	0	0	0	0
18	A	738	0	0	5	0
18	B	116	0	0	2	0
18	C	726	0	0	4	0
18	D	121	0	0	2	0
18	E	692	0	0	5	1
18	F	115	0	0	3	0
18	G	687	0	0	5	0
18	H	127	0	0	2	1
All	All	33900	0	29422	171	3

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 3.

The worst 5 of 171 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)
15:E:5010:EDO:C1	15:E:5012:EDO:O1	1.67	1.41
15:E:5010:EDO:C1	15:E:5012:EDO:C1	2.14	1.26
15:E:5010:EDO:C1	15:E:5012:EDO:C2	2.14	1.24
15:E:5010:EDO:C1	15:E:5012:EDO:H22	1.58	1.21
15:E:5010:EDO:C2	15:E:5012:EDO:C1	2.35	1.04

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:471:GLU:OE1	1:G:485[B]:GLN:OE1[1_654]	2.09	0.11
1:E:485:GLN:OE1	1:G:471:GLU:OE1[1_654]	2.13	0.07
18:E:5682:HOH:O	18:H:307:HOH:O[1_654]	2.14	0.06

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	827/822 (101%)	801 (97%)	25 (3%)	1 (0%)	48	38
1	C	824/822 (100%)	794 (96%)	28 (3%)	2 (0%)	43	34
1	E	826/822 (100%)	798 (97%)	26 (3%)	2 (0%)	43	34
1	G	823/822 (100%)	797 (97%)	26 (3%)	0	100	100
2	B	133/133 (100%)	128 (96%)	5 (4%)	0	100	100
2	D	133/133 (100%)	126 (95%)	7 (5%)	0	100	100
2	F	132/133 (99%)	128 (97%)	4 (3%)	0	100	100
2	H	132/133 (99%)	125 (95%)	7 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3830/3820 (100%)	3697 (96%)	128 (3%)	5 (0%)	48 38

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	392	ASP
1	E	392	ASP
1	C	25	ILE
1	E	793	ILE
1	A	426	GLY

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	682/675 (101%)	680 (100%)	2 (0%)	86 81
1	C	679/675 (101%)	676 (100%)	3 (0%)	84 78
1	E	681/675 (101%)	680 (100%)	1 (0%)	88 85
1	G	678/675 (100%)	674 (99%)	4 (1%)	78 72
2	B	113/111 (102%)	113 (100%)	0	100 100
2	D	113/111 (102%)	112 (99%)	1 (1%)	70 60
2	F	112/111 (101%)	111 (99%)	1 (1%)	70 60
2	H	112/111 (101%)	111 (99%)	1 (1%)	70 60
All	All	3170/3144 (101%)	3157 (100%)	13 (0%)	84 78

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	41	SER
1	G	7	ILE
2	H	104	ARG
1	G	532	HIS
1	G	774	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 85 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	688	GLN
1	G	532	HIS
1	E	704	ASN
1	G	82	ASN
1	G	638	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 60 ligands modelled in this entry, 14 are monoatomic - leaving 46 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	MGD	C	5102	5	45,52,52	1.34	7 (15%)	54,81,81	1.82	12 (22%)
9	PO4	E	5007	-	4,4,4	0.83	0	6,6,6	0.66	0
14	FES	D	201	2	0,4,4	-	-	-		
9	PO4	C	5107	-	4,4,4	0.74	0	6,6,6	0.65	0
12	GOL	C	5108	-	5,5,5	0.17	0	5,5,5	0.48	0
14	FES	H	201	2	0,4,4	-	-	-		
7	PEG	A	5006	-	6,6,6	0.84	0	5,5,5	0.91	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	PEG	E	5006	-	6,6,6	0.29	0	5,5,5	0.78	0
9	PO4	E	5011	-	4,4,4	1.35	0	6,6,6	0.92	0
8	ACT	A	5009	-	3,3,3	0.95	0	3,3,3	0.95	0
9	PO4	A	5008	-	4,4,4	0.84	0	6,6,6	0.81	0
9	PO4	G	5008	-	4,4,4	0.80	0	6,6,6	0.73	0
12	GOL	A	5016	-	5,5,5	0.80	0	5,5,5	1.48	1 (20%)
14	FES	F	201	2	0,4,4	-	-	-		
6	F3S	G	5005	1	0,9,9	-	-	-		
8	ACT	G	5011	-	3,3,3	0.97	0	3,3,3	1.33	0
6	F3S	A	5005	1	0,9,9	-	-	-		
6	F3S	E	5005	1	0,9,9	-	-	-		
8	ACT	G	5009	-	3,3,3	0.69	0	3,3,3	1.48	1 (33%)
3	MGD	E	5002	5	45,52,52	1.56	6 (13%)	54,81,81	1.58	11 (20%)
8	ACT	A	5007	-	3,3,3	0.83	0	3,3,3	0.61	0
6	F3S	C	5105	1	0,9,9	-	-	-		
7	PEG	A	5011	-	6,6,6	0.52	0	5,5,5	0.83	0
3	MGD	E	5001	5	45,52,52	1.44	7 (15%)	54,81,81	1.79	14 (25%)
13	TRS	G	5006	-	7,7,7	1.49	1 (14%)	9,9,9	2.13	4 (44%)
3	MGD	G	5002	5	45,52,52	1.67	8 (17%)	54,81,81	1.59	11 (20%)
10	UJI	C	5109	5,4	0,4,4	-	-	-		
15	EDO	E	5010	-	3,3,3	1.41	0	2,2,2	1.62	1 (50%)
10	UJI	E	5009	5,4	0,4,4	-	-	-		
13	TRS	C	5112	-	7,7,7	1.57	2 (28%)	9,9,9	2.01	4 (44%)
10	UJI	A	5010	4	0,4,4	-	-	-		
3	MGD	A	5001	5	45,52,52	1.57	8 (17%)	54,81,81	1.60	11 (20%)
15	EDO	E	5012	-	3,3,3	1.34	0	2,2,2	0.64	0
9	PO4	C	5110	-	4,4,4	1.08	0	6,6,6	0.69	0
12	GOL	E	5008	-	5,5,5	0.26	0	5,5,5	1.15	0
16	SO4	E	5015	-	4,4,4	1.22	1 (25%)	6,6,6	0.54	0
7	PEG	G	5007	-	6,6,6	0.51	0	5,5,5	0.74	0
3	MGD	A	5002	5	45,52,52	1.59	9 (20%)	54,81,81	1.85	14 (25%)
3	MGD	C	5101	5	45,52,52	1.57	9 (20%)	54,81,81	1.61	12 (22%)
7	PEG	C	5106	-	6,6,6	1.09	0	5,5,5	1.48	1 (20%)
12	GOL	A	5015	-	5,5,5	0.26	0	5,5,5	0.19	0
7	PEG	A	5012	-	6,6,6	1.33	1 (16%)	5,5,5	1.92	1 (20%)
10	UJI	G	5010	4	0,4,4	-	-	-		
3	MGD	G	5001	5	45,52,52	1.32	4 (8%)	54,81,81	1.84	17 (31%)
13	TRS	A	5017	-	7,7,7	1.98	3 (42%)	9,9,9	1.50	2 (22%)
14	FES	B	201	2	0,4,4	-	-	-		

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	MGD	C	5102	5	-	4/22/66/66	0/6/6/6
14	FES	D	201	2	-	-	0/1/1/1
12	GOL	C	5108	-	-	0/4/4/4	-
14	FES	H	201	2	-	-	0/1/1/1
7	PEG	A	5006	-	-	2/4/4/4	-
7	PEG	E	5006	-	-	2/4/4/4	-
12	GOL	A	5016	-	-	1/4/4/4	-
14	FES	F	201	2	-	-	0/1/1/1
6	F3S	G	5005	1	-	-	0/3/3/3
6	F3S	A	5005	1	-	-	0/3/3/3
6	F3S	E	5005	1	-	-	0/3/3/3
3	MGD	E	5002	5	-	4/22/66/66	0/6/6/6
6	F3S	C	5105	1	-	-	0/3/3/3
7	PEG	A	5011	-	-	1/4/4/4	-
3	MGD	E	5001	5	-	3/22/66/66	0/6/6/6
13	TRS	G	5006	-	-	6/9/9/9	-
3	MGD	G	5002	5	-	4/22/66/66	0/6/6/6
15	EDO	E	5010	-	-	1/1/1/1	-
13	TRS	C	5112	-	-	5/9/9/9	-
15	EDO	E	5012	-	-	1/1/1/1	-
3	MGD	A	5001	5	-	2/22/66/66	0/6/6/6
12	GOL	E	5008	-	-	4/4/4/4	-
7	PEG	G	5007	-	-	4/4/4/4	-
3	MGD	A	5002	5	-	4/22/66/66	0/6/6/6
3	MGD	C	5101	5	-	4/22/66/66	0/6/6/6
7	PEG	C	5106	-	-	1/4/4/4	-
12	GOL	A	5015	-	-	3/4/4/4	-
7	PEG	A	5012	-	-	1/4/4/4	-
3	MGD	G	5001	5	-	3/22/66/66	0/6/6/6
13	TRS	A	5017	-	-	4/9/9/9	-
14	FES	B	201	2	-	-	0/1/1/1

The worst 5 of 66 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	5002	MGD	C16-C21	5.99	1.48	1.38
3	E	5002	MGD	C16-C21	5.32	1.47	1.38
3	A	5001	MGD	C16-C21	5.13	1.47	1.38
3	E	5001	MGD	C16-C21	4.95	1.47	1.38
3	G	5001	MGD	C16-C21	4.71	1.46	1.38

The worst 5 of 117 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	5002	MGD	C5-C4-N3	-5.00	120.34	128.46
3	E	5002	MGD	C5-C4-N3	-4.62	120.97	128.46
3	E	5001	MGD	O11-C23-C14	4.41	111.91	108.96
3	C	5102	MGD	O11-C23-N22	-4.28	104.16	108.57
3	E	5001	MGD	C5-C4-N3	-4.26	121.56	128.46

There are no chirality outliers.

5 of 64 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	5001	MGD	C5'-O5'-PB-O1B
3	A	5002	MGD	C5'-O5'-PB-O1B
3	A	5002	MGD	C5'-O5'-PB-O3B
3	C	5101	MGD	C5'-O5'-PB-O1B
3	C	5102	MGD	PA-O3B-PB-O5'

There are no ring outliers.

19 monomers are involved in 59 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	5102	MGD	1	0
7	E	5006	PEG	2	0
9	E	5011	PO4	1	0
8	A	5009	ACT	3	0
3	E	5002	MGD	1	0
7	A	5011	PEG	2	0
3	E	5001	MGD	2	0
13	G	5006	TRS	7	0
10	C	5109	UJI	1	0
15	E	5010	EDO	9	0
13	C	5112	TRS	7	0
3	A	5001	MGD	2	0
15	E	5012	EDO	9	0

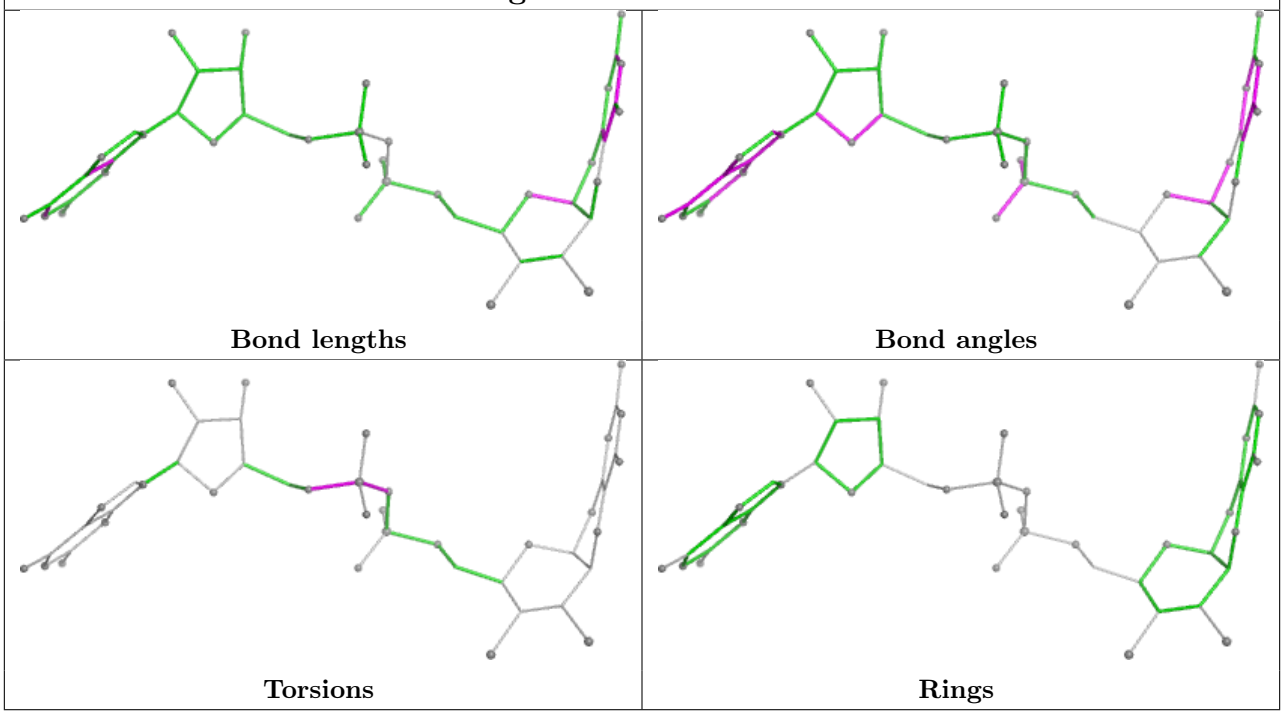
Continued on next page...

Continued from previous page...

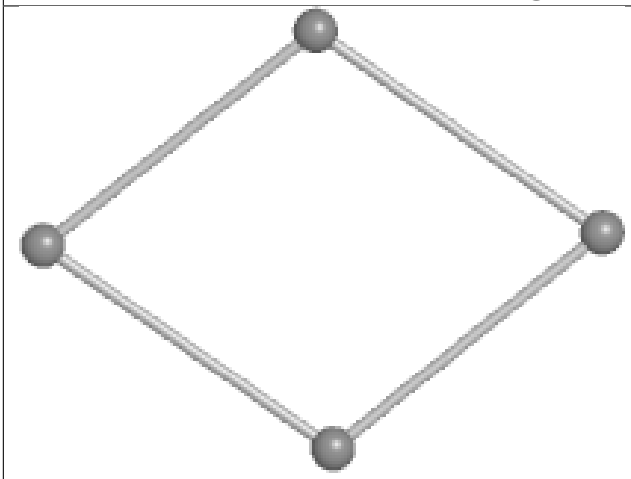
Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	E	5008	GOL	2	0
3	C	5101	MGD	1	0
7	C	5106	PEG	1	0
7	A	5012	PEG	3	0
3	G	5001	MGD	1	0
13	A	5017	TRS	14	0

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.

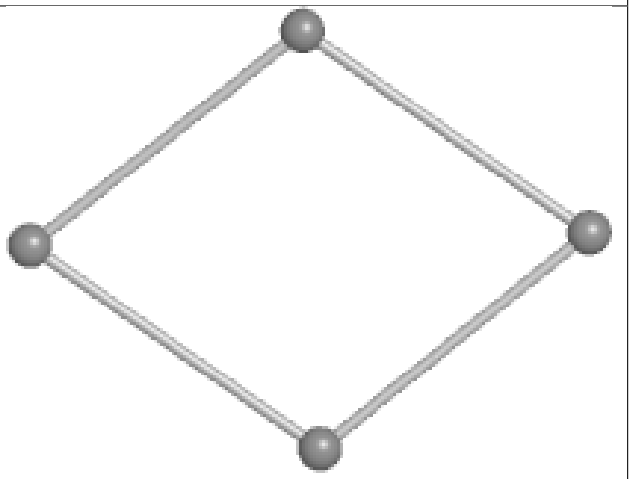
Ligand MGD C 5102



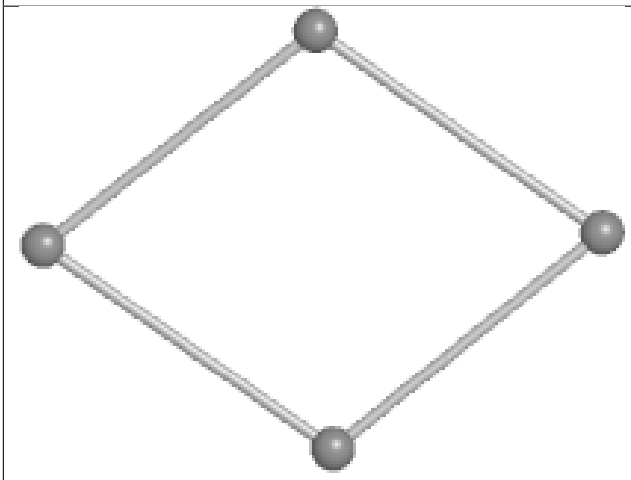
Ligand FES D 201



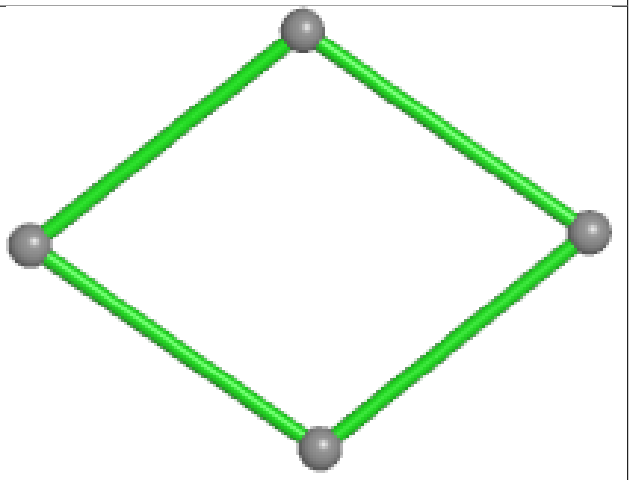
Bond lengths



Bond angles

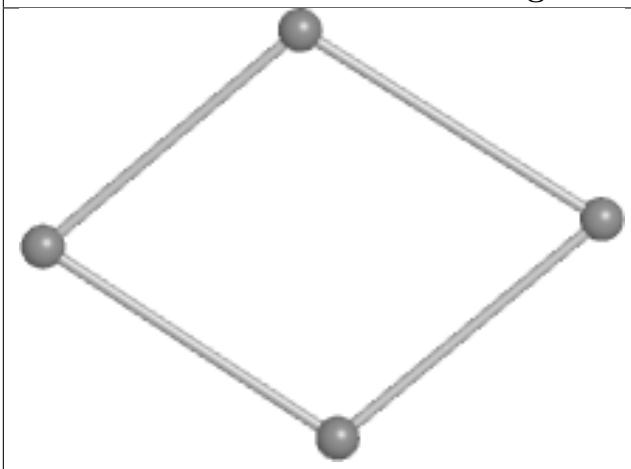


Torsions

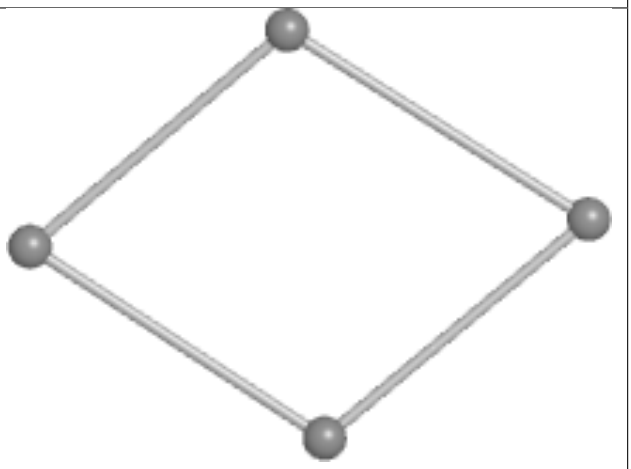


Rings

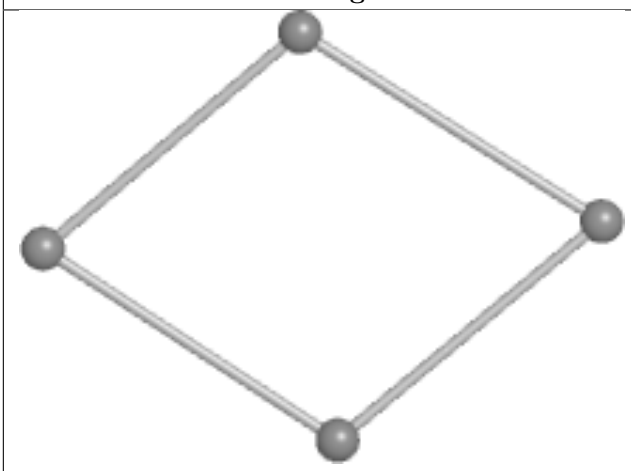
Ligand FES H 201



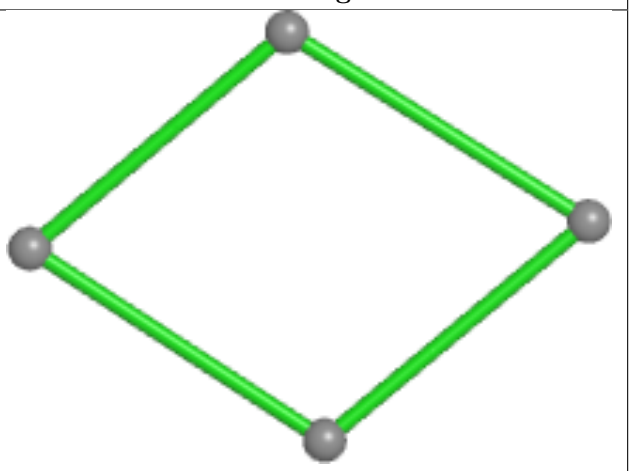
Bond lengths



Bond angles

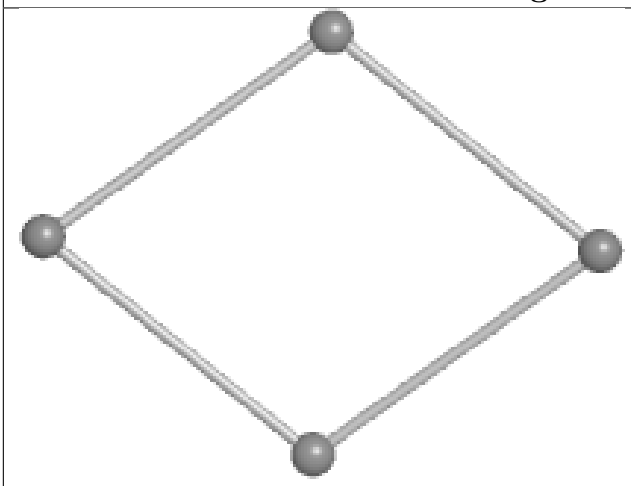


Torsions

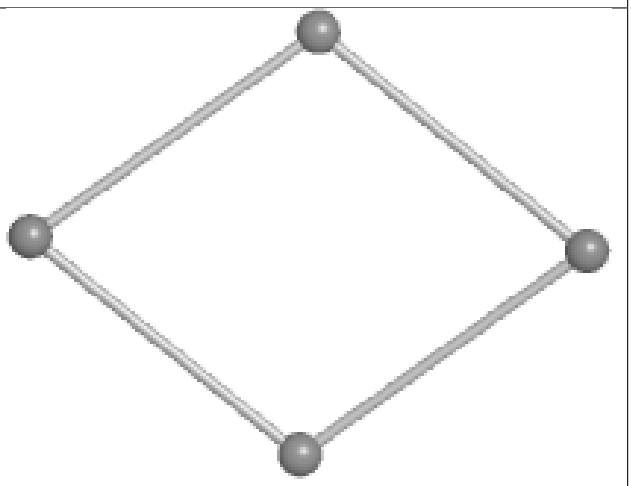


Rings

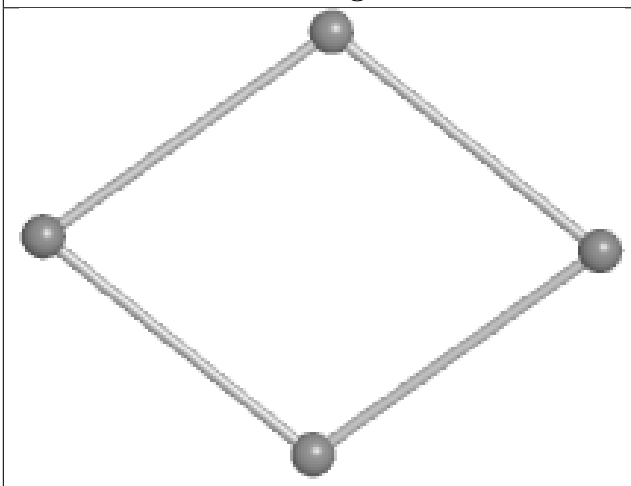
Ligand FES F 201



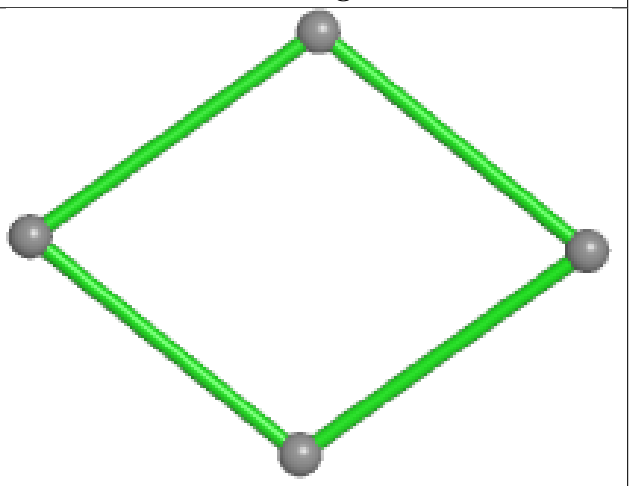
Bond lengths



Bond angles

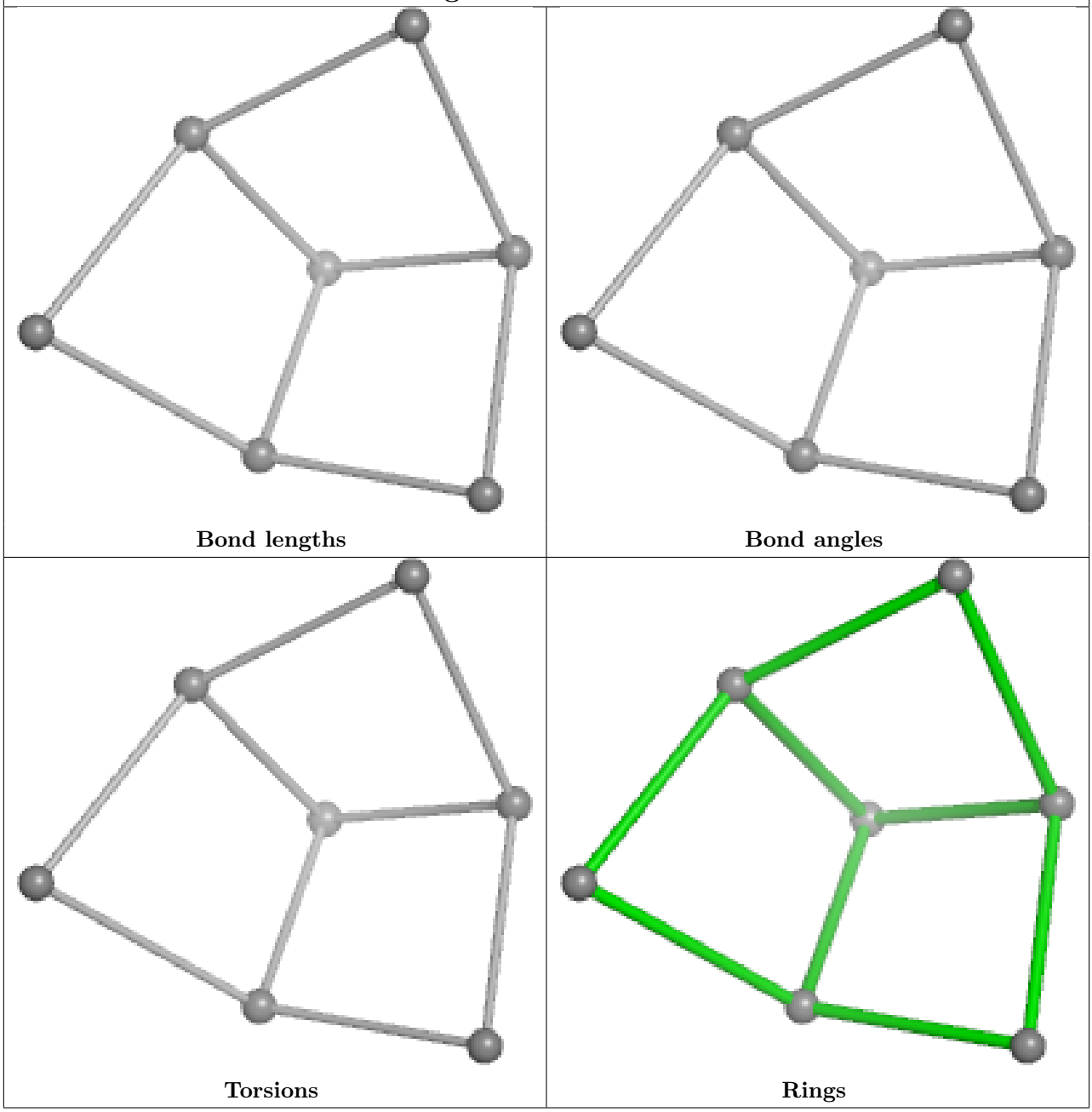


Torsions

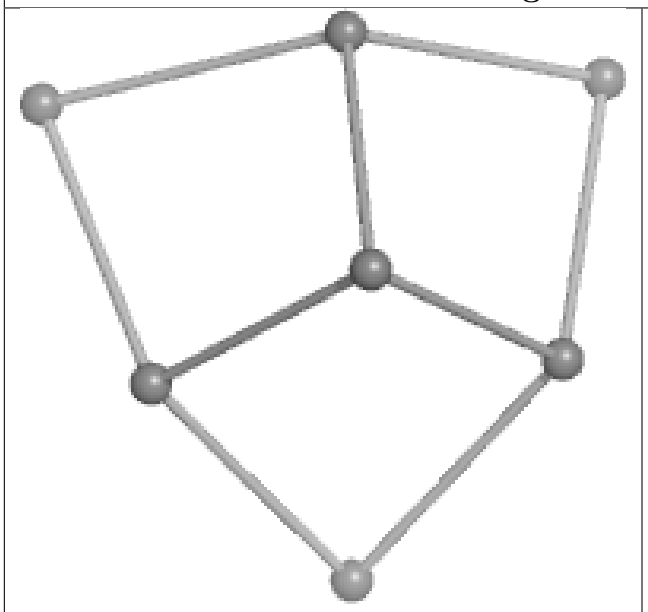


Rings

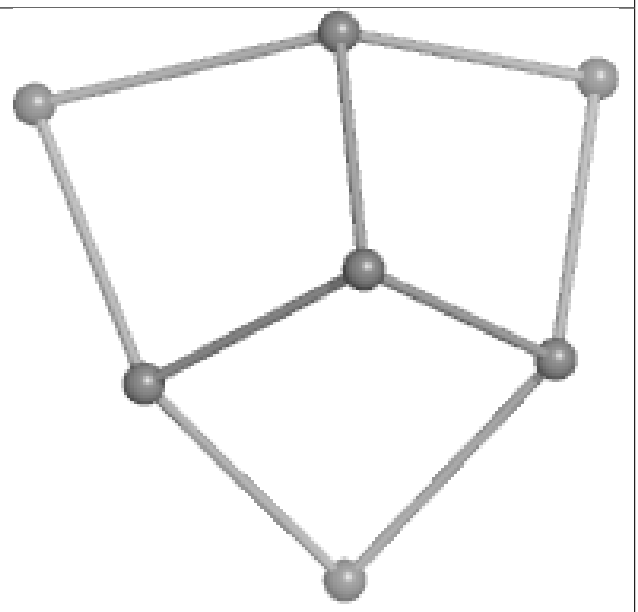
Ligand F3S G 5005



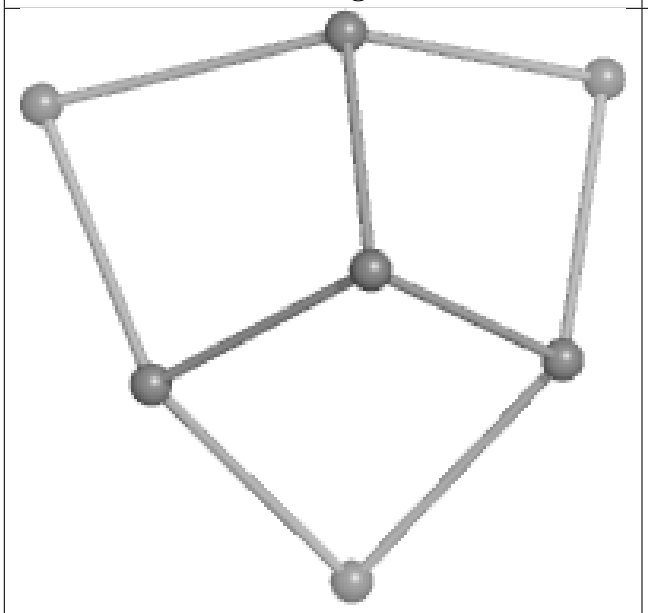
Ligand F3S A 5005



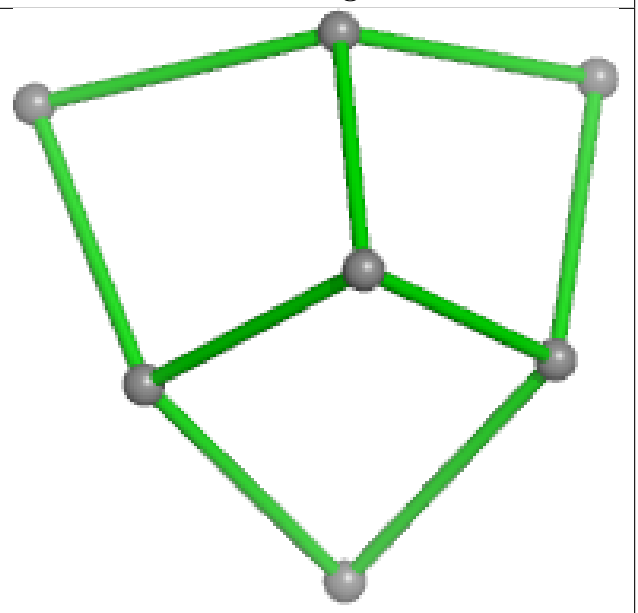
Bond lengths



Bond angles

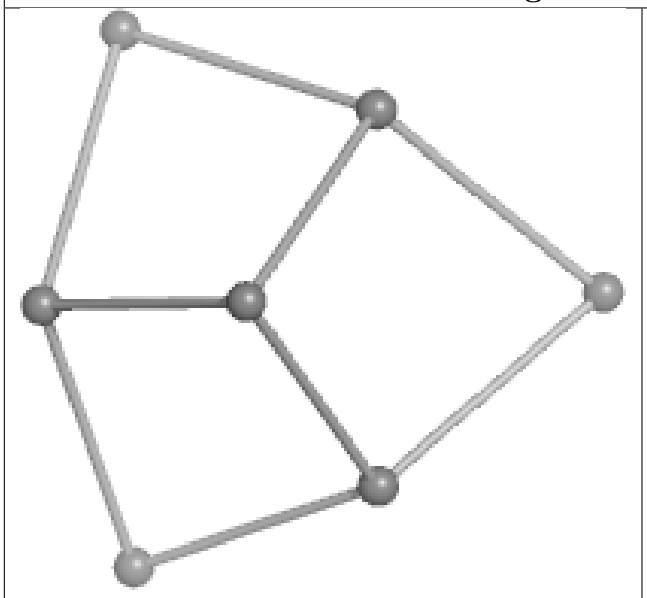


Torsions

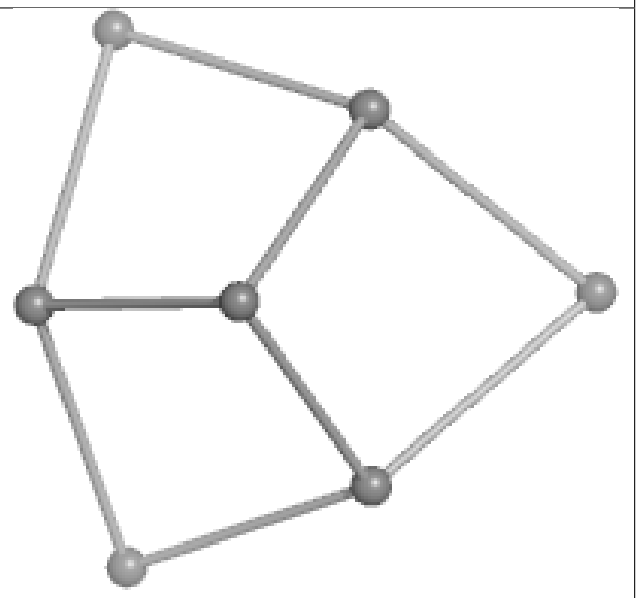


Rings

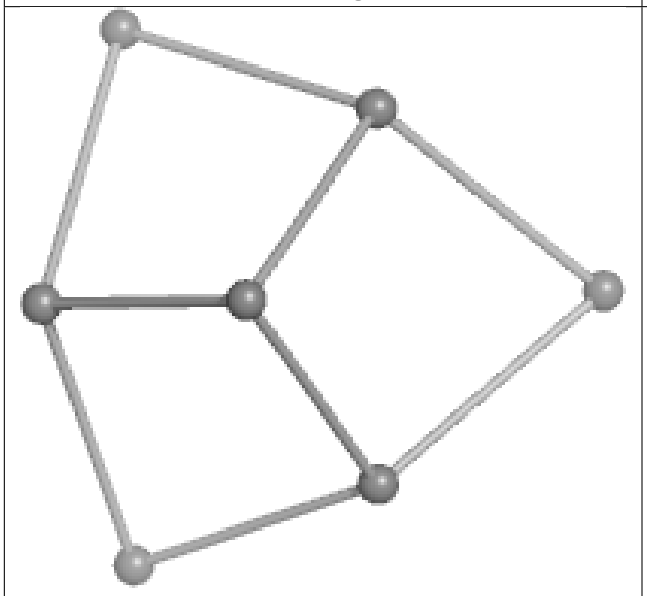
Ligand F3S E 5005



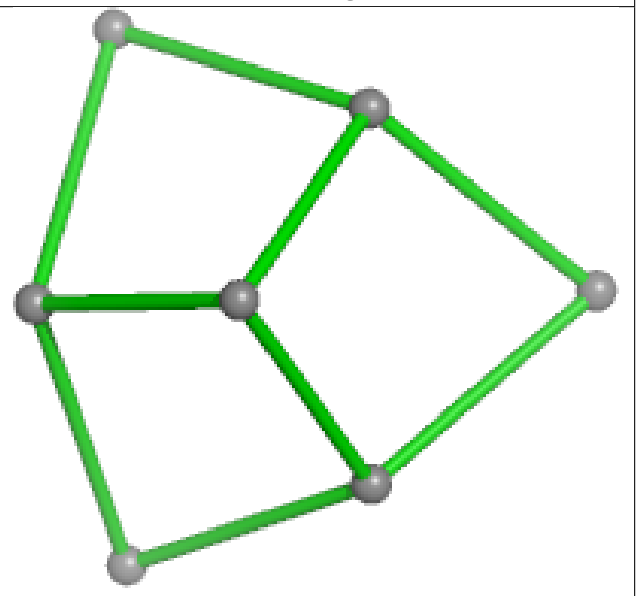
Bond lengths



Bond angles

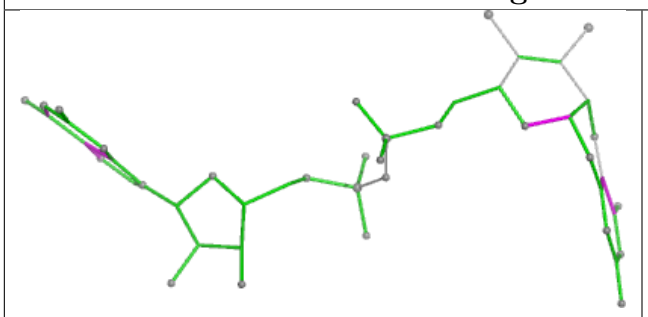


Torsions

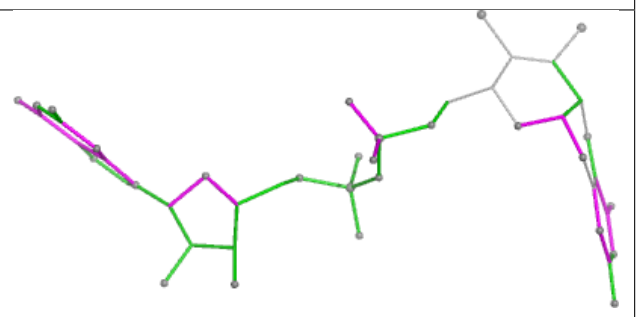


Rings

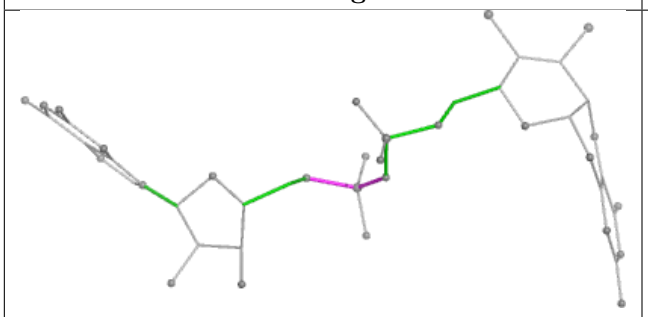
Ligand MGD E 5002



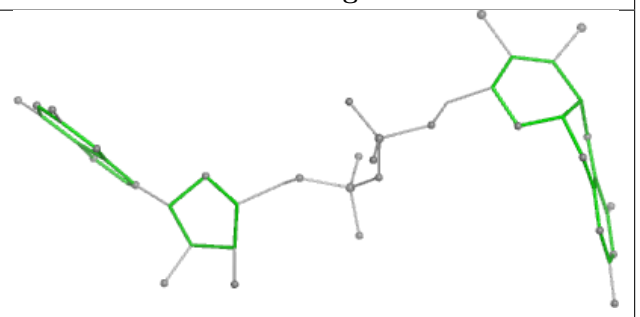
Bond lengths



Bond angles

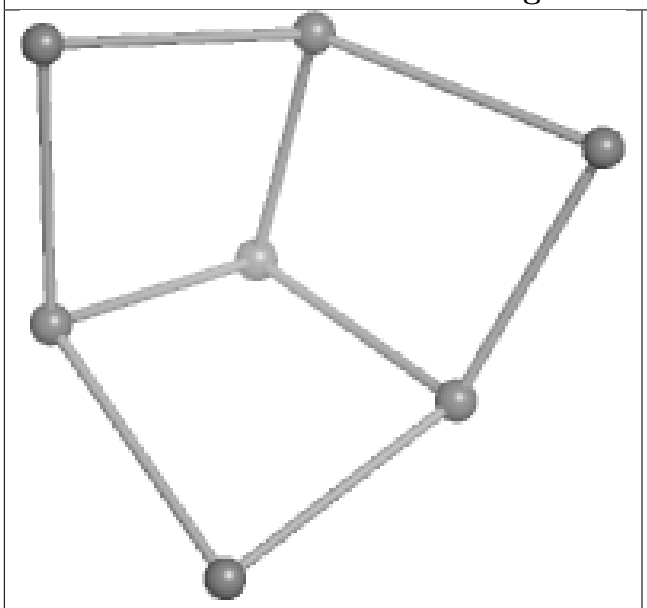


Torsions

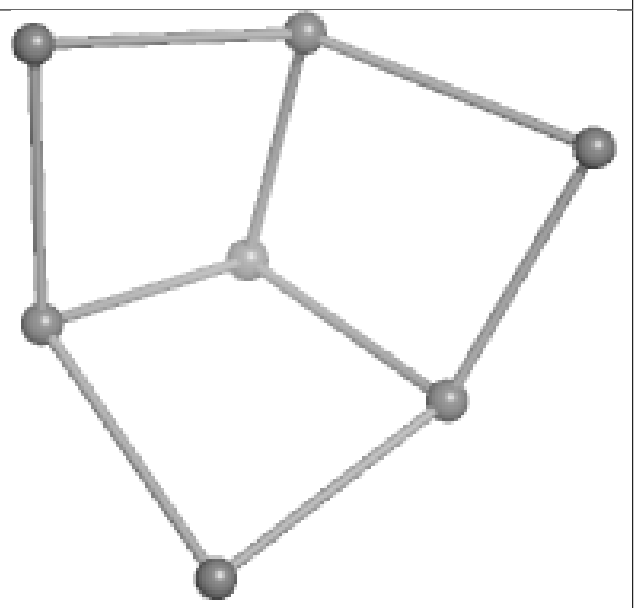


Rings

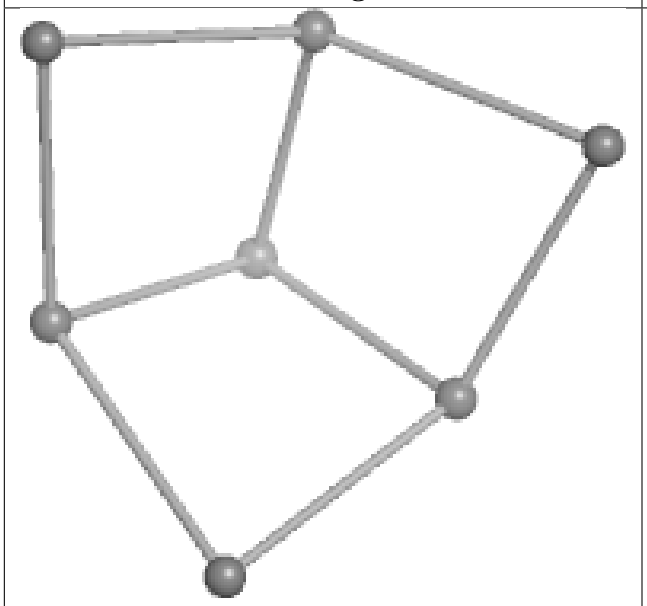
Ligand F3S C 5105



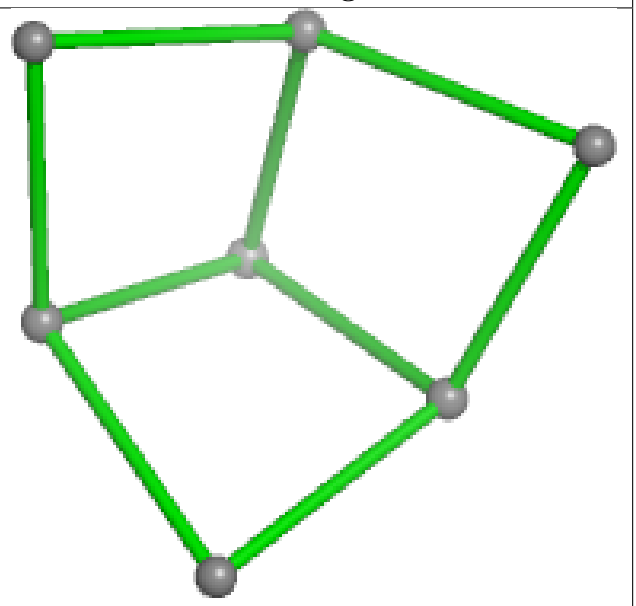
Bond lengths



Bond angles

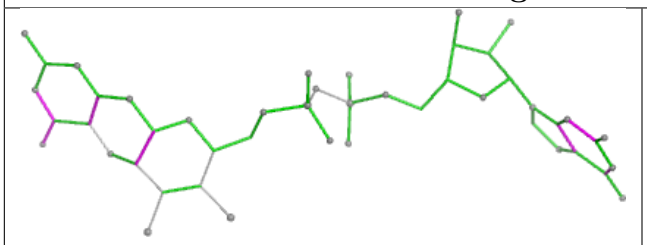


Torsions

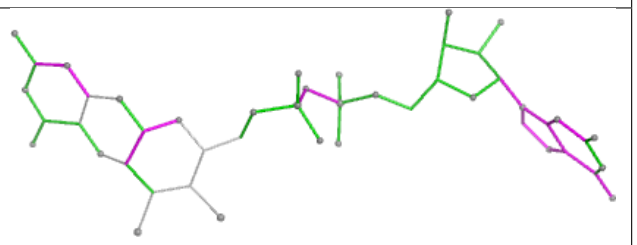


Rings

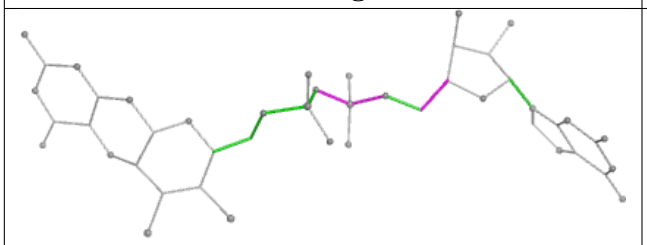
Ligand MGD E 5001



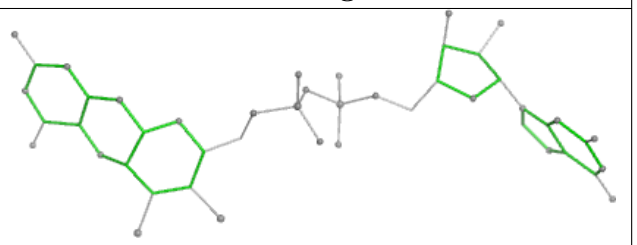
Bond lengths



Bond angles

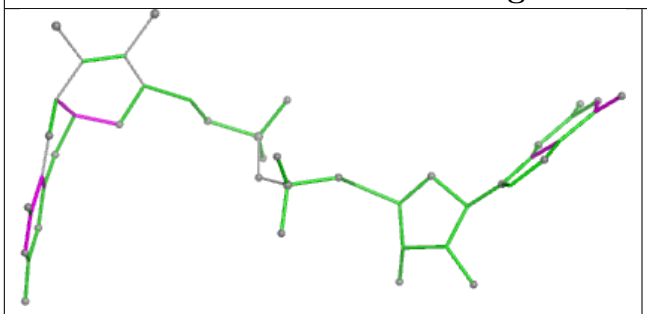


Torsions

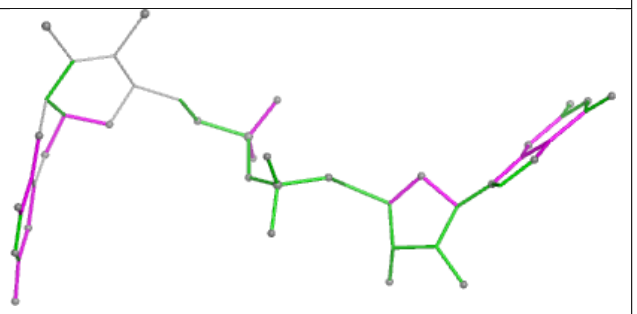


Rings

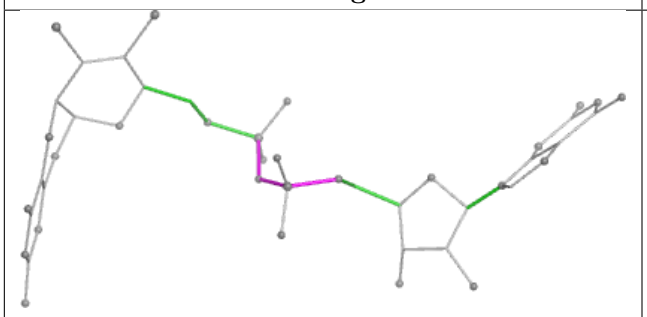
Ligand MGD G 5002



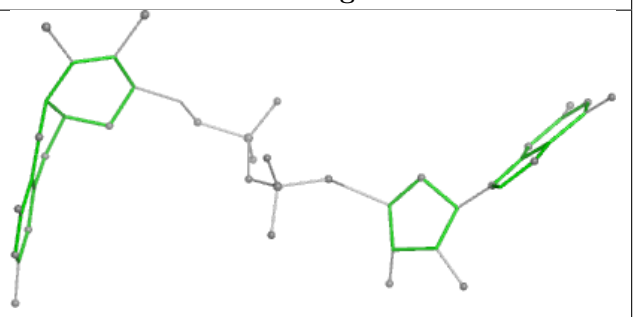
Bond lengths



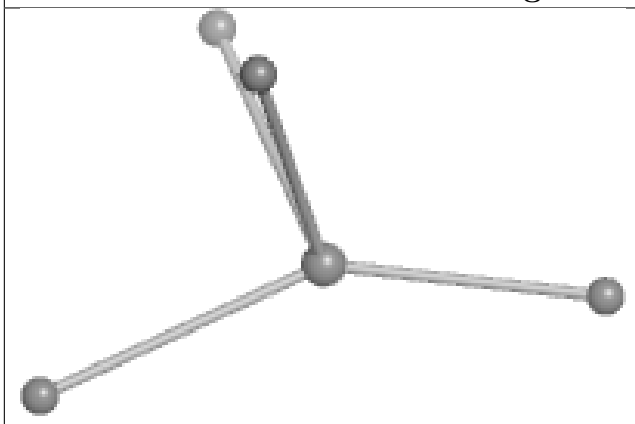
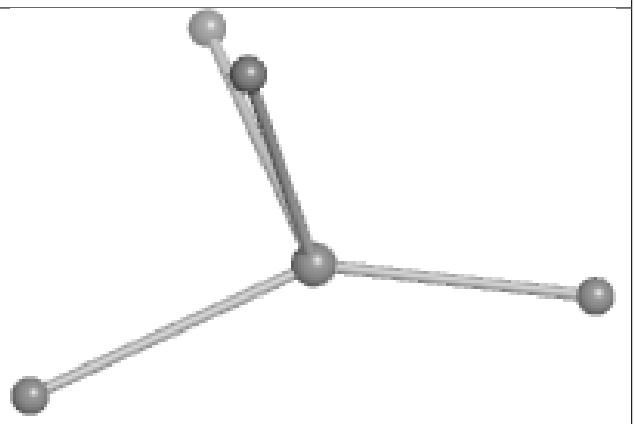
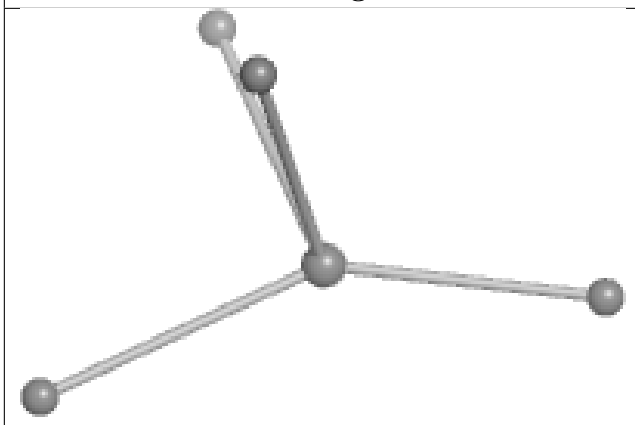
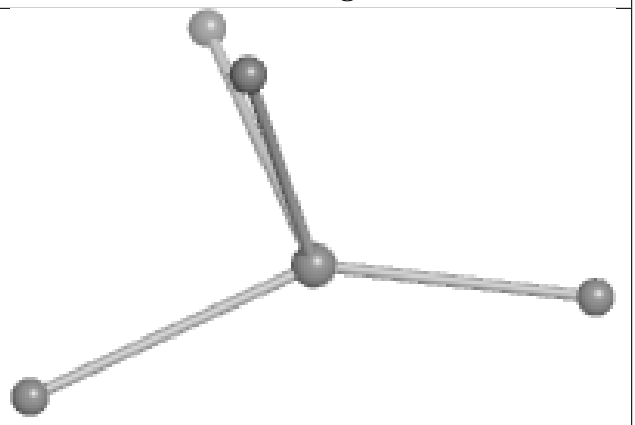
Bond angles



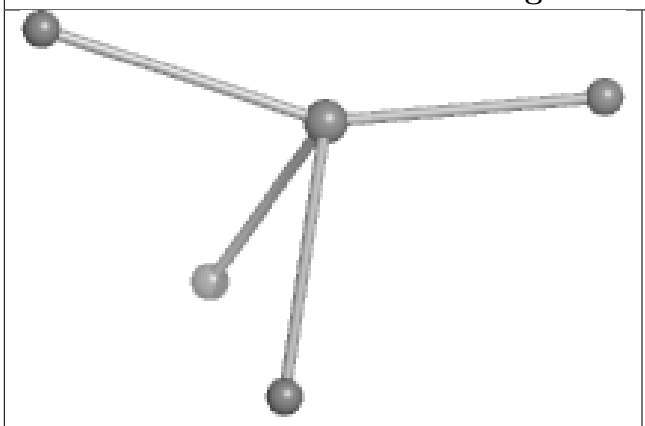
Torsions



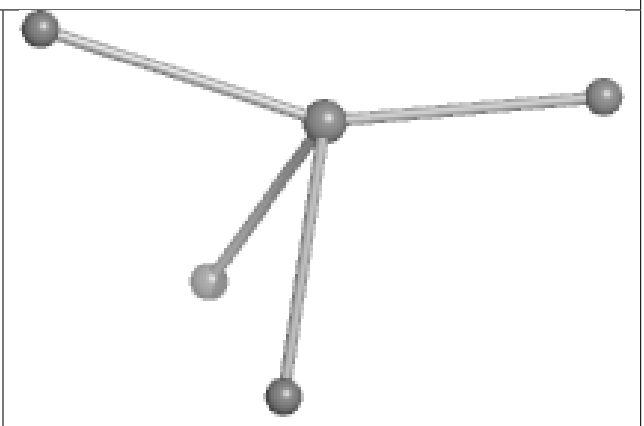
Rings

Ligand UJI C 5109	
 Bond lengths	 Bond angles
 Torsions	 Rings

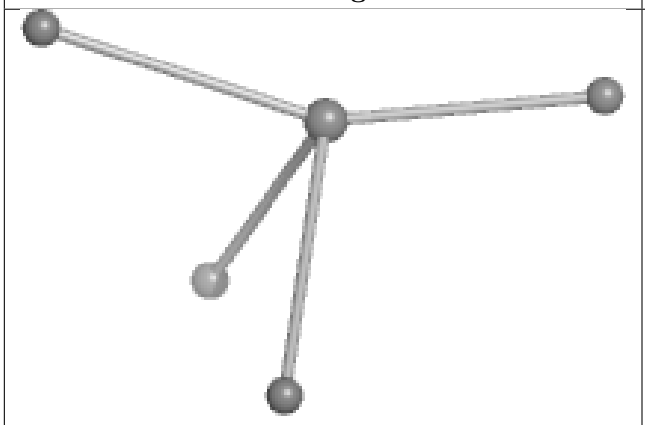
Ligand UJI E 5009



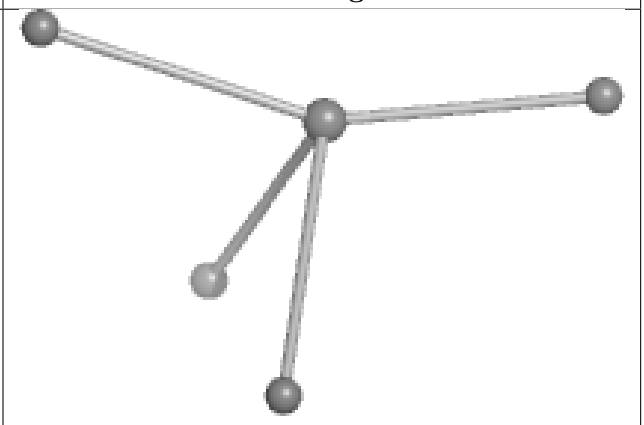
Bond lengths



Bond angles

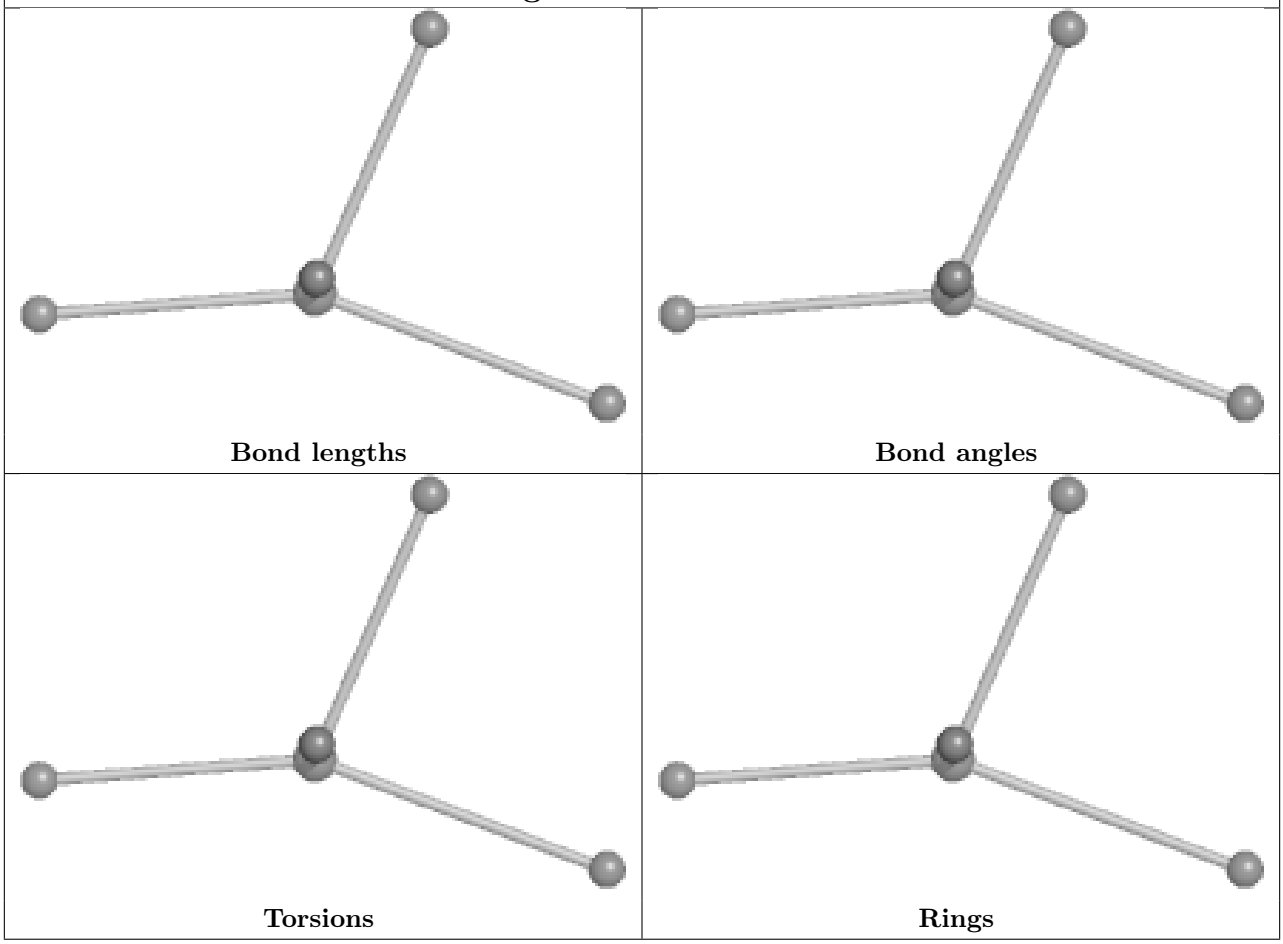


Torsions

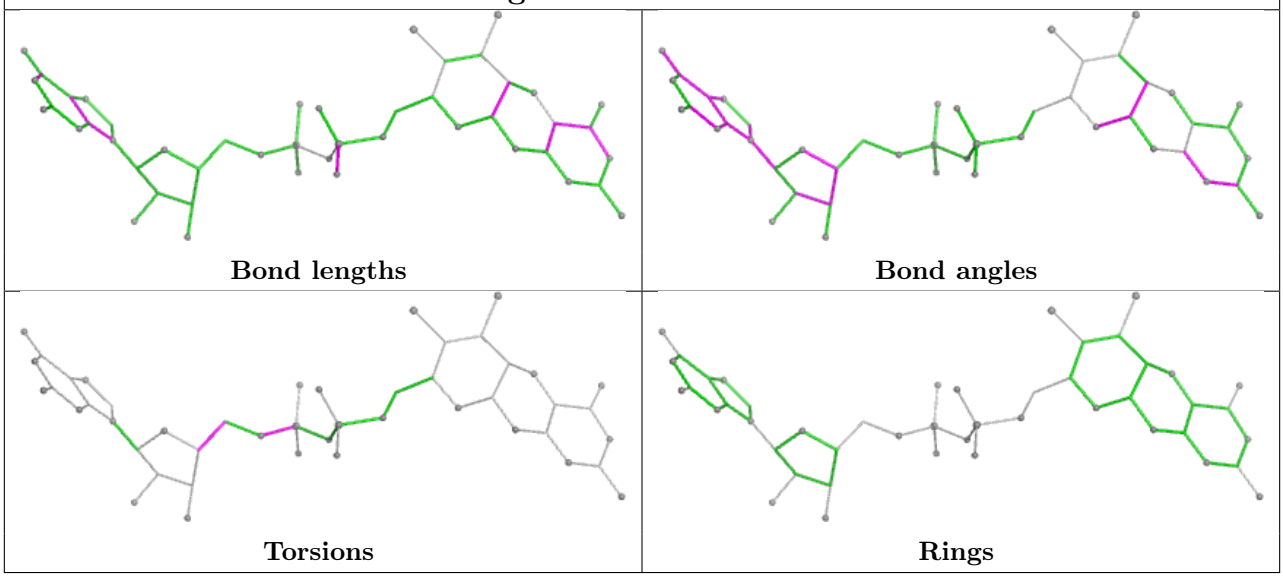


Rings

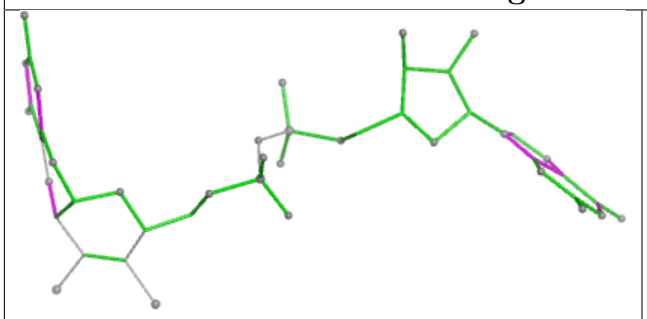
Ligand UJI A 5010



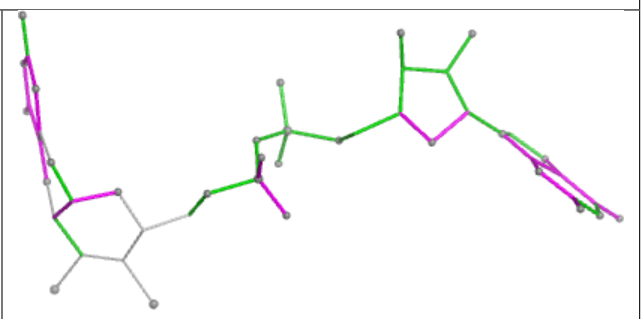
Ligand MGD A 5001



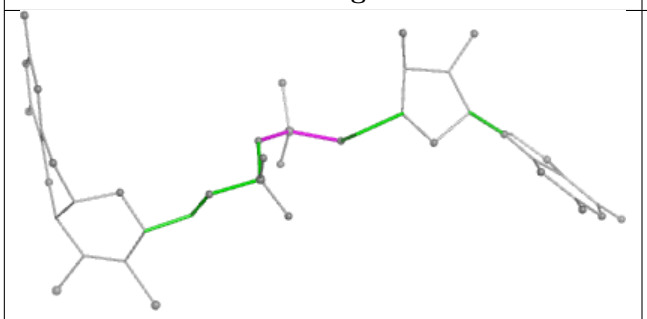
Ligand MGD A 5002



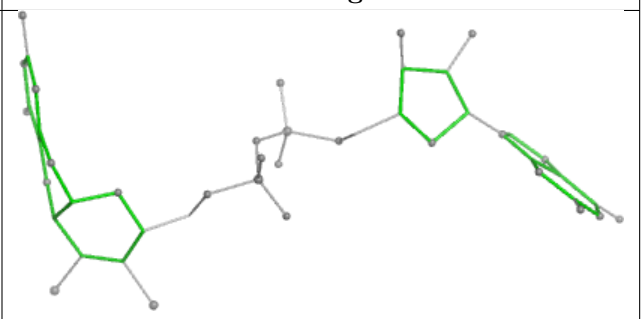
Bond lengths



Bond angles

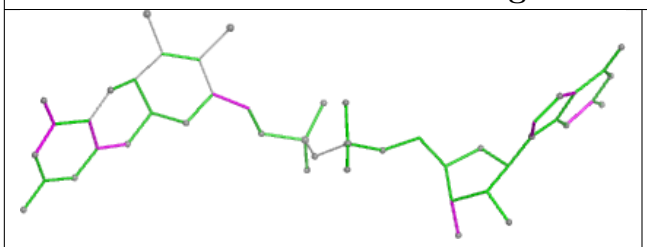


Torsions

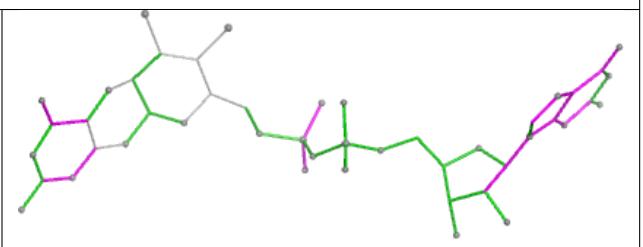


Rings

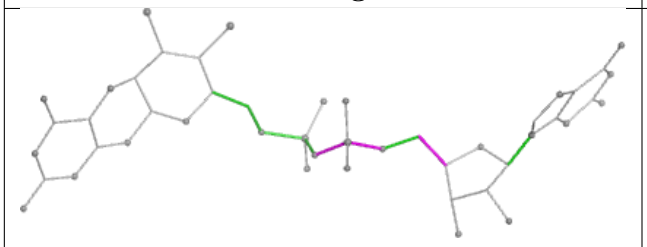
Ligand MGD C 5101



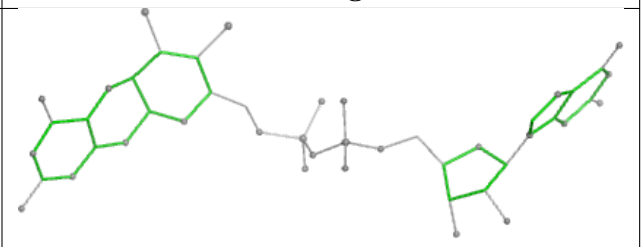
Bond lengths



Bond angles

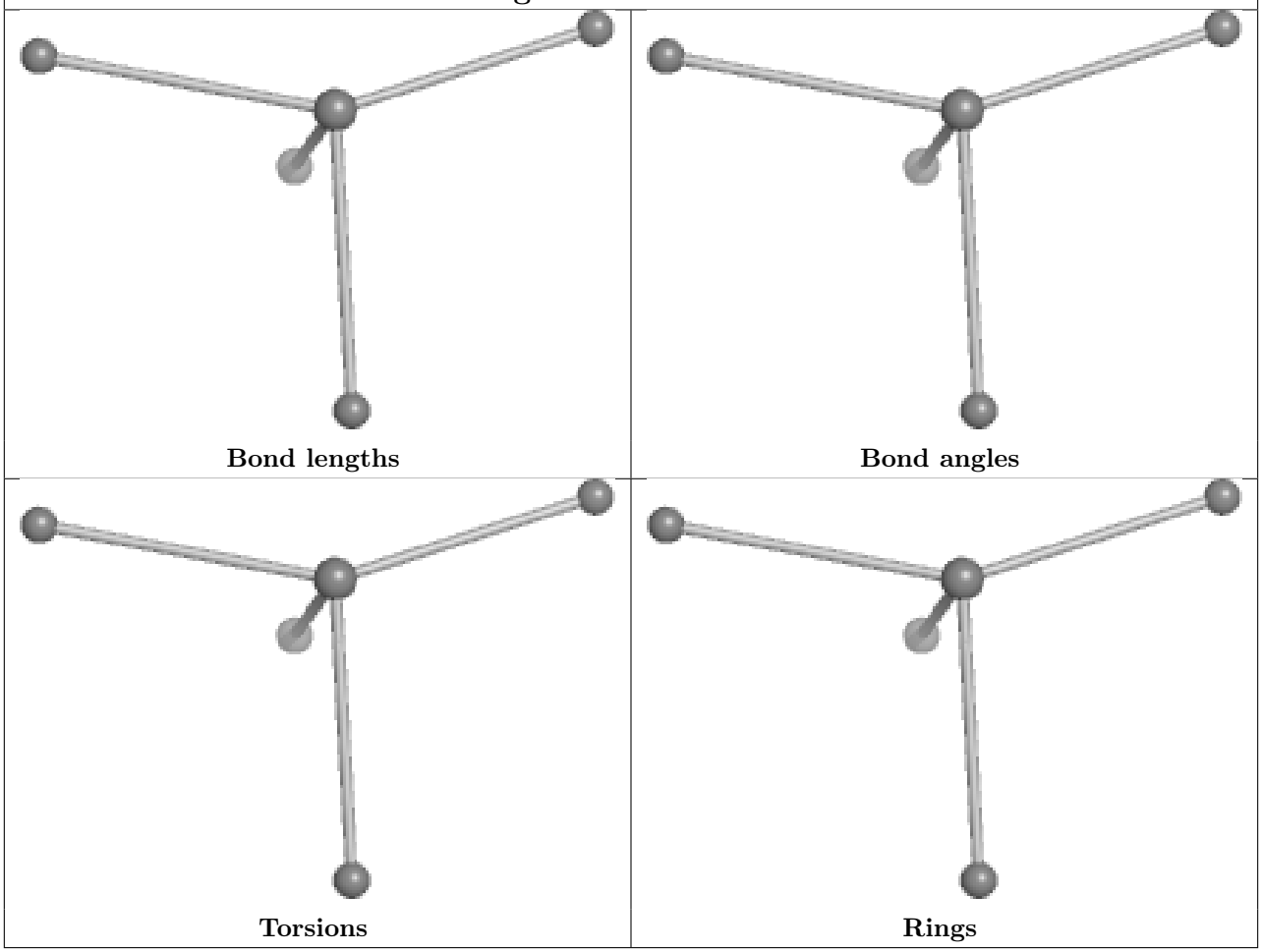


Torsions

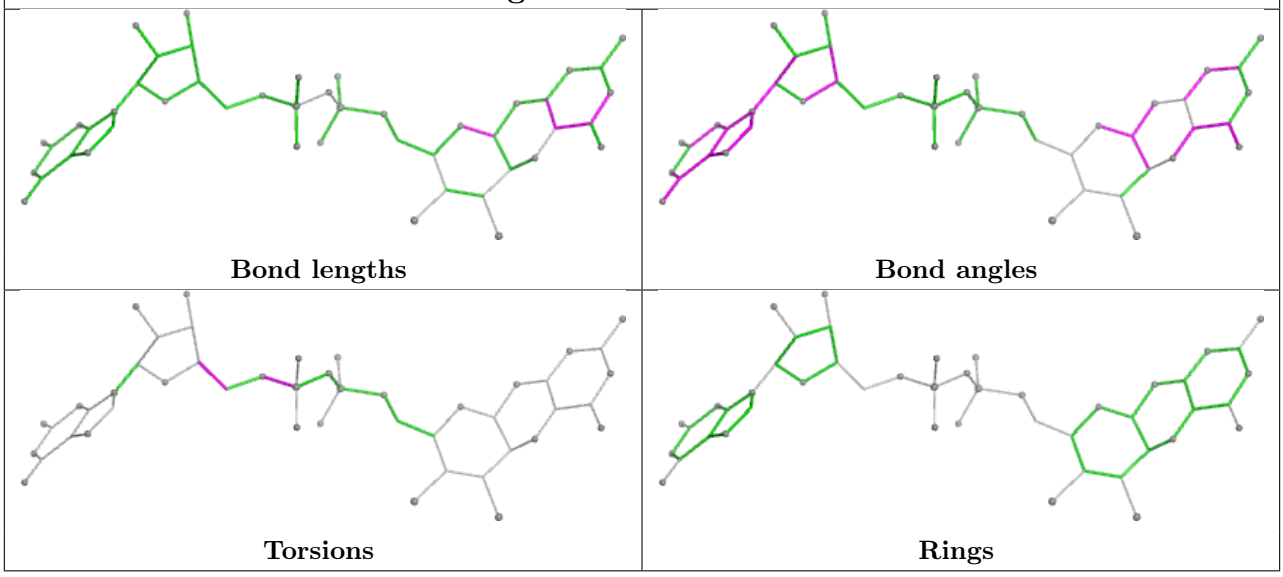


Rings

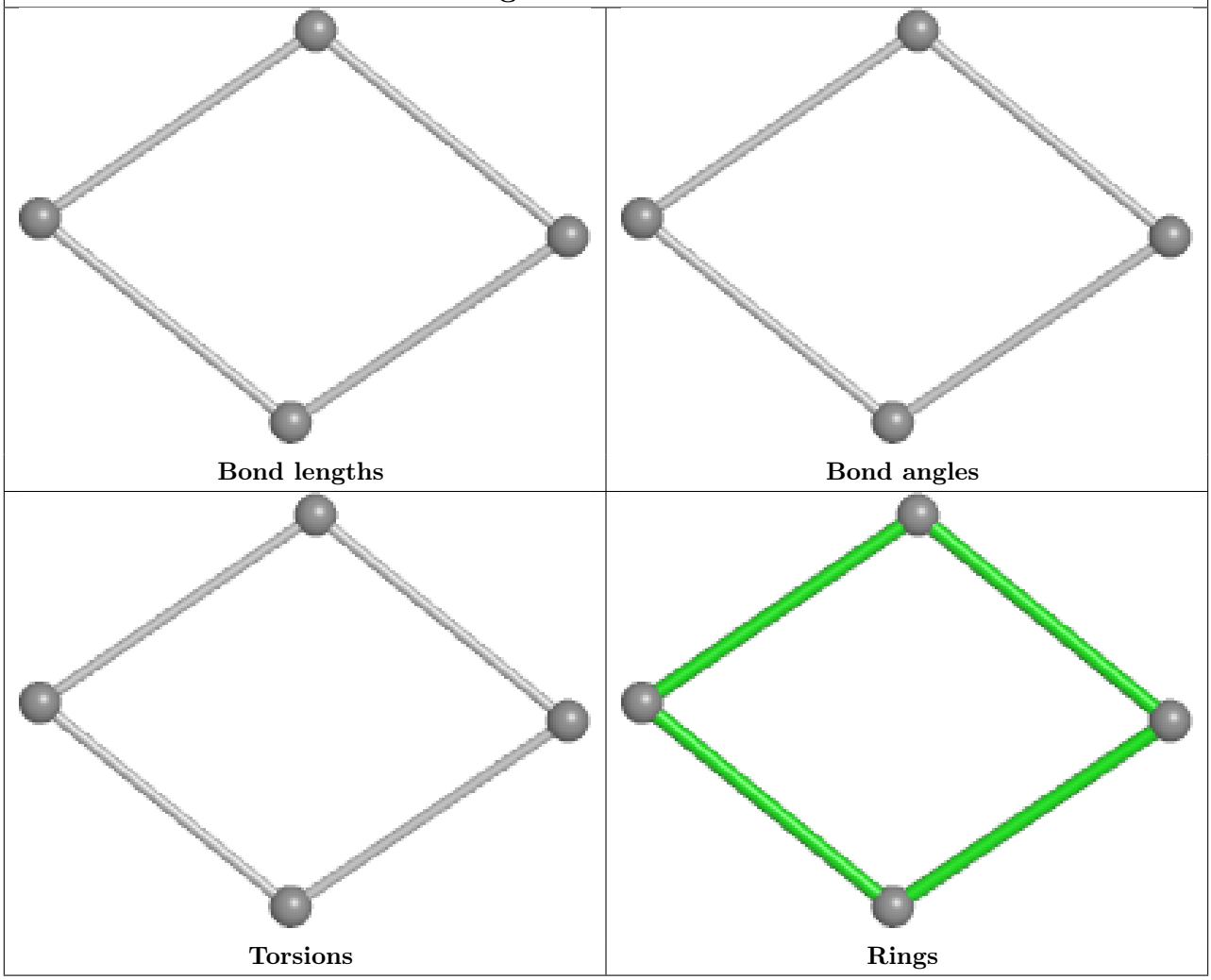
Ligand UJI G 5010



Ligand MGD G 5001



Ligand FES B 201



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	822/822 (100%)	-0.45	1 (0%)	92	95	9, 18, 31, 53	8 (0%)
1	C	822/822 (100%)	-0.46	2 (0%)	91	94	11, 17, 29, 50	5 (0%)
1	E	822/822 (100%)	-0.39	3 (0%)	88	93	11, 18, 31, 54	7 (0%)
1	G	822/822 (100%)	-0.35	3 (0%)	88	93	12, 19, 32, 56	5 (0%)
2	B	133/133 (100%)	-0.34	0	100	100	13, 19, 30, 40	2 (1%)
2	D	133/133 (100%)	-0.34	0	100	100	13, 19, 32, 42	2 (1%)
2	F	133/133 (100%)	-0.32	0	100	100	12, 19, 31, 40	1 (0%)
2	H	133/133 (100%)	-0.29	0	100	100	14, 19, 30, 40	1 (0%)
All	All	3820/3820 (100%)	-0.40	9 (0%)	91	94	9, 18, 31, 56	31 (0%)

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	128	ALA	3.5
1	G	4	ASN	3.1
1	E	128	ALA	2.6
1	C	4	ASN	2.5
1	A	435	ASP	2.5

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 oligosaccharides 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
7	PEG	C	5106	7/7	0.75	0.20	35,37,41,46	0
8	ACT	G	5009	4/4	0.75	0.17	38,42,43,44	0
8	ACT	A	5009	4/4	0.77	0.20	28,39,40,46	0
7	PEG	A	5006	7/7	0.80	0.16	30,36,40,44	0
7	PEG	A	5011	7/7	0.81	0.18	46,48,51,53	0
7	PEG	G	5007	7/7	0.82	0.16	32,41,48,52	0
15	EDO	E	5012	4/4	0.83	0.19	40,43,46,51	0
13	TRS	G	5006	8/8	0.84	0.17	24,25,37,39	0
8	ACT	A	5007	4/4	0.86	0.16	51,56,57,58	0
7	PEG	A	5012	7/7	0.88	0.12	26,27,36,36	0
9	PO4	C	5107	5/5	0.89	0.10	53,54,57,61	0
9	PO4	G	5008	5/5	0.90	0.09	54,55,62,64	0
12	GOL	A	5016	6/6	0.90	0.14	27,31,34,40	0
7	PEG	E	5006	7/7	0.90	0.12	28,35,43,43	0
9	PO4	E	5007	5/5	0.90	0.10	56,59,68,69	0
13	TRS	C	5112	8/8	0.91	0.14	28,30,38,40	0
9	PO4	A	5008	5/5	0.91	0.10	55,58,61,62	0
13	TRS	A	5017	8/8	0.91	0.15	25,28,35,38	0
8	ACT	G	5011	4/4	0.92	0.17	37,37,40,44	0
12	GOL	C	5108	6/6	0.92	0.11	38,39,41,44	0
11	CL	E	5013	1/1	0.93	0.20	53,53,53,53	0
15	EDO	E	5010	4/4	0.93	0.11	31,33,33,35	0
12	GOL	E	5008	6/6	0.93	0.10	23,34,36,45	0
12	GOL	A	5015	6/6	0.94	0.09	38,41,42,46	0
16	SO4	E	5015	5/5	0.94	0.10	24,25,33,34	0
17	NA	G	5012	1/1	0.94	0.12	32,32,32,32	0
9	PO4	E	5011	5/5	0.95	0.08	44,44,46,48	0
11	CL	A	5013	1/1	0.95	0.13	42,42,42,42	0
11	CL	A	5014	1/1	0.95	0.12	61,61,61,61	0
10	UJI	C	5109	5/5	0.96	0.14	29,33,41,41	1
4	O	E	5003	1/1	0.97	0.08	26,26,26,26	0
11	CL	C	5111	1/1	0.97	0.14	42,42,42,42	0
10	UJI	E	5009	5/5	0.97	0.11	25,28,33,37	1
11	CL	E	5014	1/1	0.97	0.18	40,40,40,40	0
10	UJI	A	5010	5/5	0.97	0.12	27,28,35,36	1
3	MGD	G	5002	47/47	0.98	0.04	13,15,16,17	0
4	O	A	5003	1/1	0.98	0.08	23,23,23,23	0

Continued on next page...

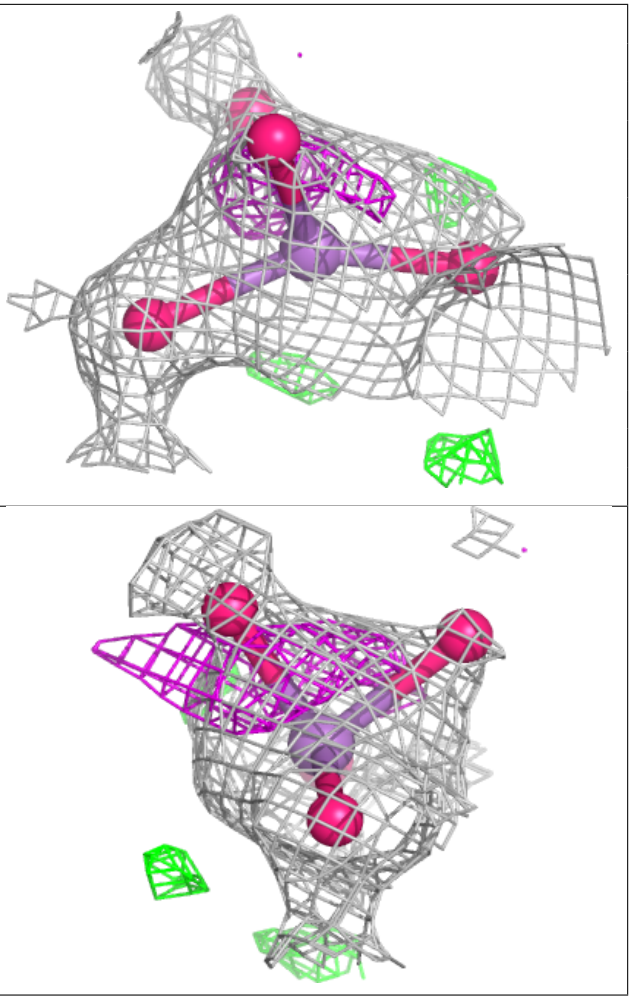
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	O	C	5103	1/1	0.98	0.08	28,28,28,28	0
3	MGD	A	5002	47/47	0.98	0.04	12,14,15,16	0
4	O	G	5003	1/1	0.98	0.06	25,25,25,25	0
10	UJI	G	5010	5/5	0.98	0.12	29,34,37,38	1
3	MGD	C	5101	47/47	0.98	0.05	11,14,16,17	0
3	MGD	E	5001	47/47	0.98	0.05	12,14,17,18	0
3	MGD	E	5002	47/47	0.98	0.04	12,14,16,18	0
9	PO4	C	5110	5/5	0.98	0.06	41,41,47,50	0
3	MGD	G	5001	47/47	0.98	0.04	13,15,18,19	0
6	F3S	A	5005	7/7	0.99	0.02	10,12,13,13	0
3	MGD	A	5001	47/47	0.99	0.04	12,14,16,17	0
3	MGD	C	5102	47/47	0.99	0.04	12,13,15,16	0
14	FES	D	201	4/4	0.99	0.03	17,17,18,18	0
5	4MO	G	5004	1/1	1.00	0.02	19,19,19,19	0
5	4MO	A	5004	1/1	1.00	0.02	19,19,19,19	0
14	FES	B	201	4/4	1.00	0.02	17,18,18,18	0
6	F3S	C	5105	7/7	1.00	0.02	12,13,14,15	0
14	FES	F	201	4/4	1.00	0.03	16,17,18,18	0
14	FES	H	201	4/4	1.00	0.03	18,18,19,19	0
6	F3S	E	5005	7/7	1.00	0.02	12,13,13,15	0
6	F3S	G	5005	7/7	1.00	0.02	13,14,15,17	0
5	4MO	C	5104	1/1	1.00	0.01	18,18,18,18	0
5	4MO	E	5004	1/1	1.00	0.01	19,19,19,19	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.

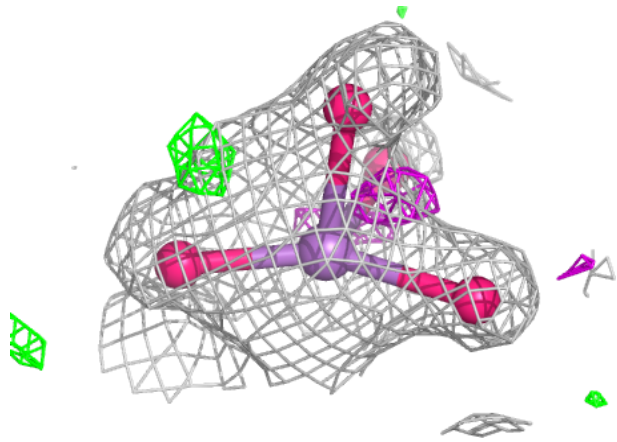
Electron density around UJI C 5109:

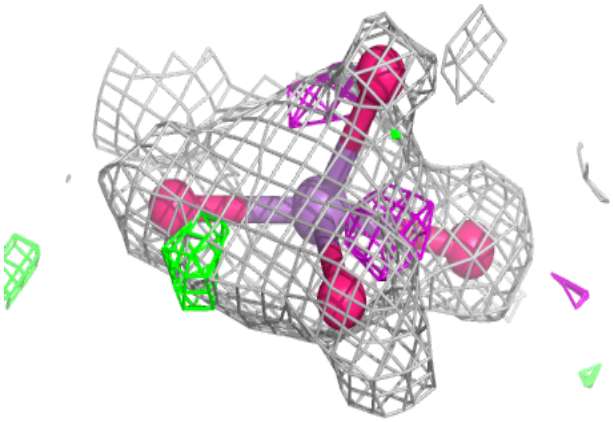
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

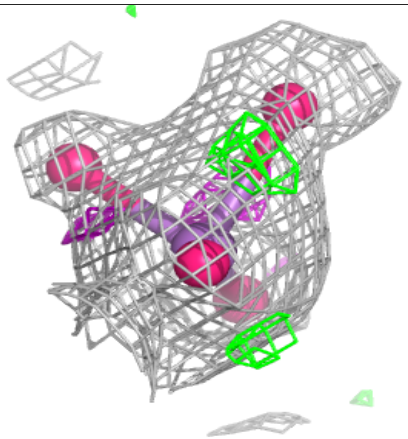


Electron density around UJI E 5009:

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

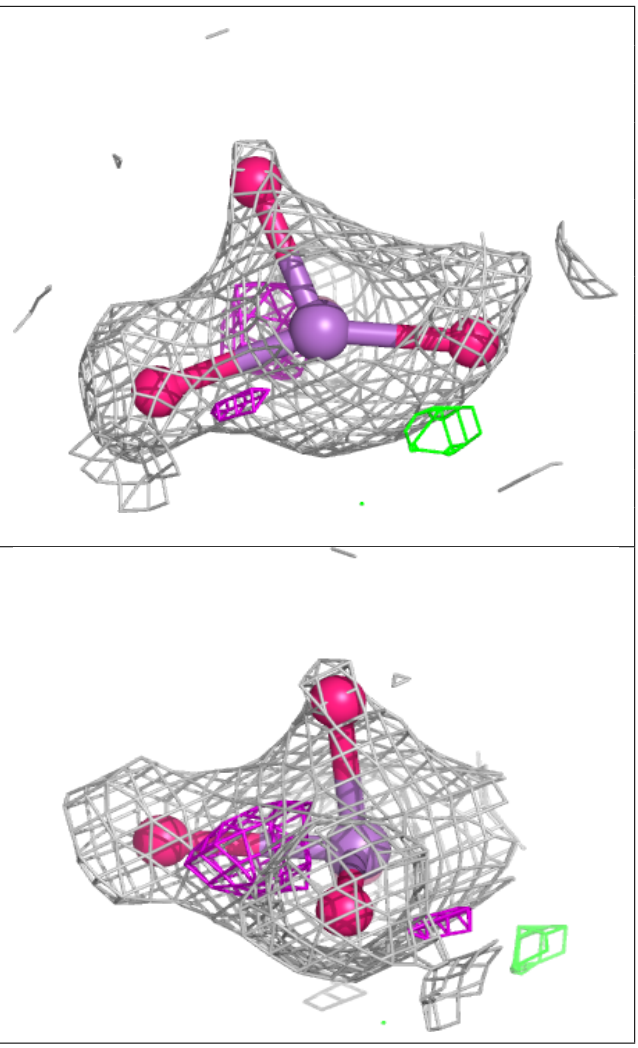






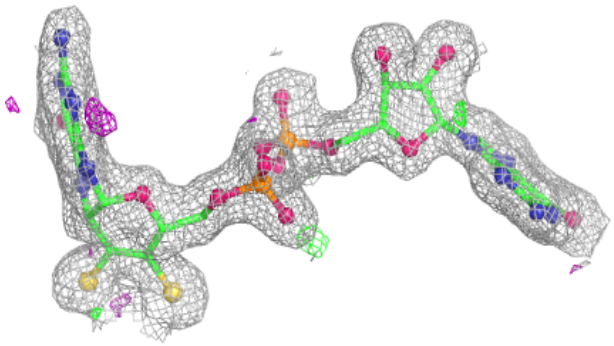
Electron density around UJI A 5010:

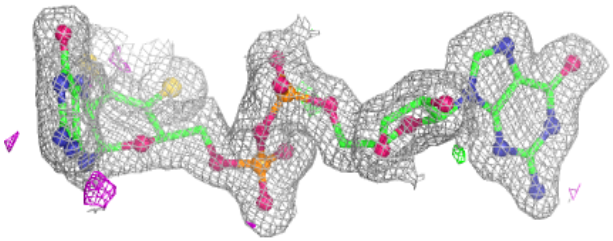
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

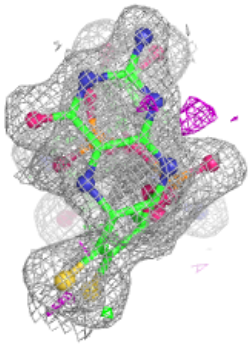


Electron density around MGD G 5002:

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

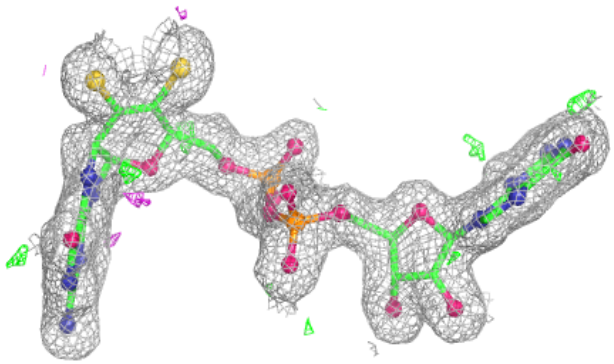


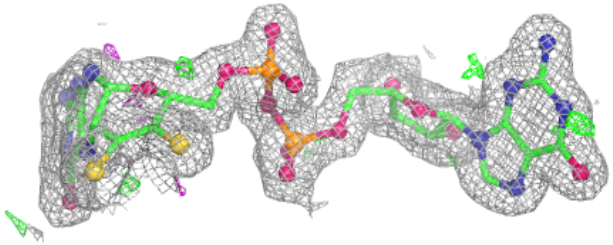


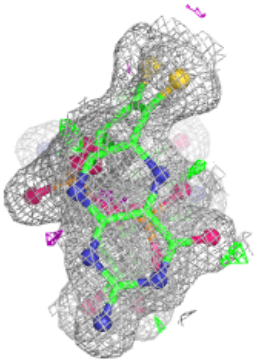


Electron density around MGD A 5002:

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

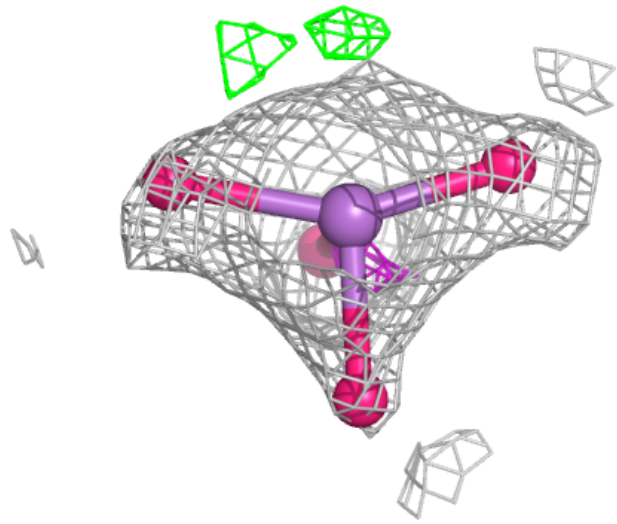


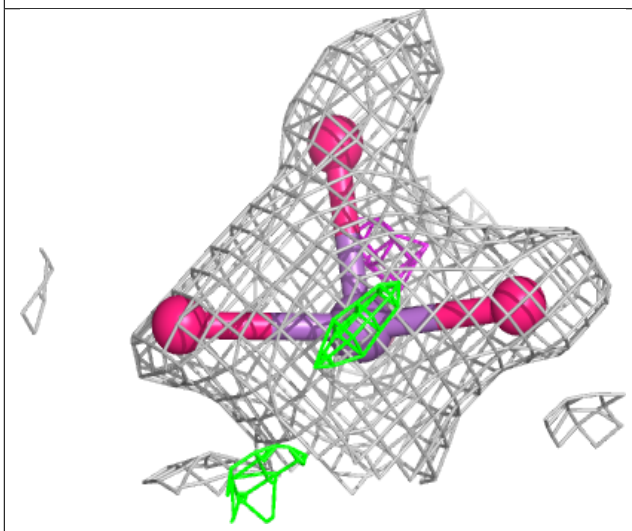


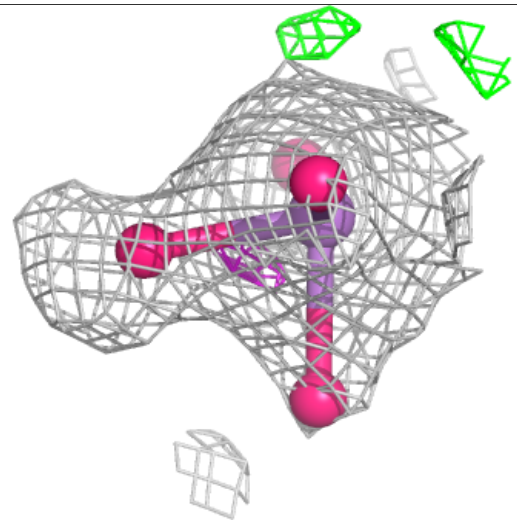


Electron density around UJI G 5010:

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

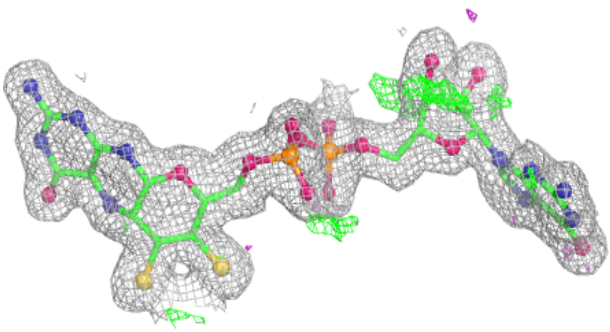


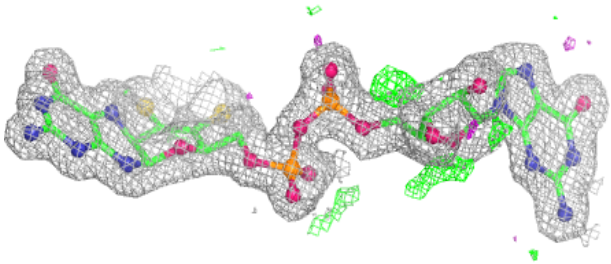


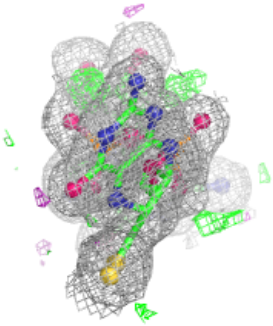


Electron density around MGD C 5101:

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

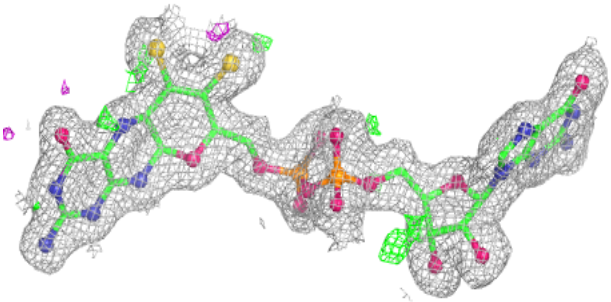


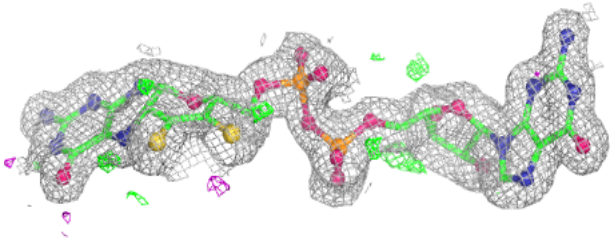


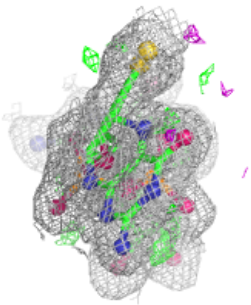


Electron density around MGD E 5001:

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

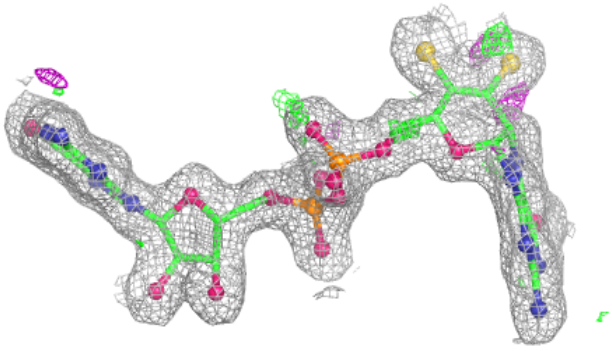


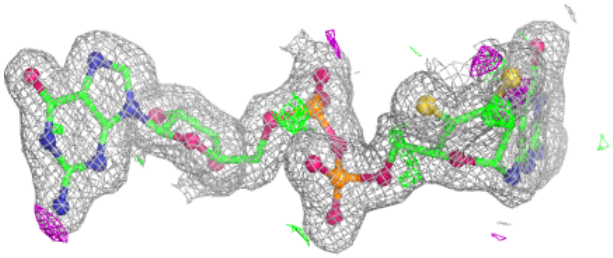


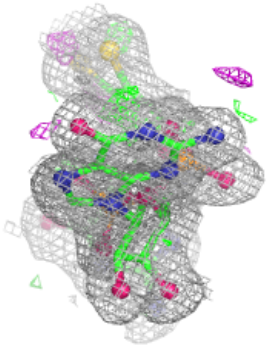


Electron density around MGD E 5002:

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

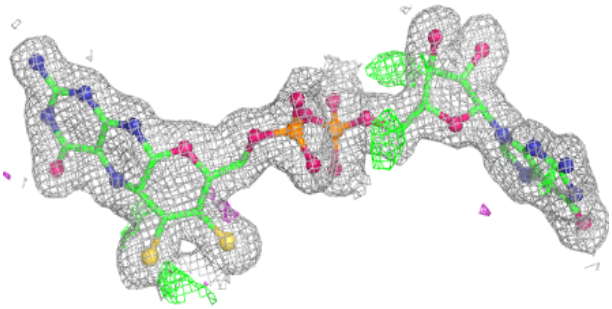


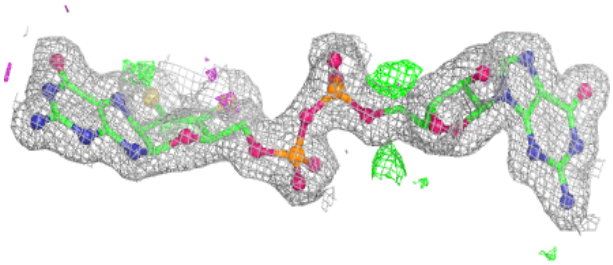


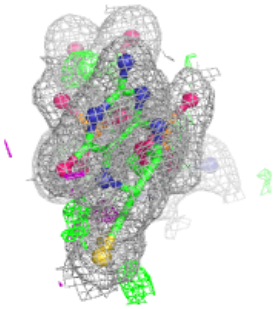


Electron density around MGD G 5001:

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

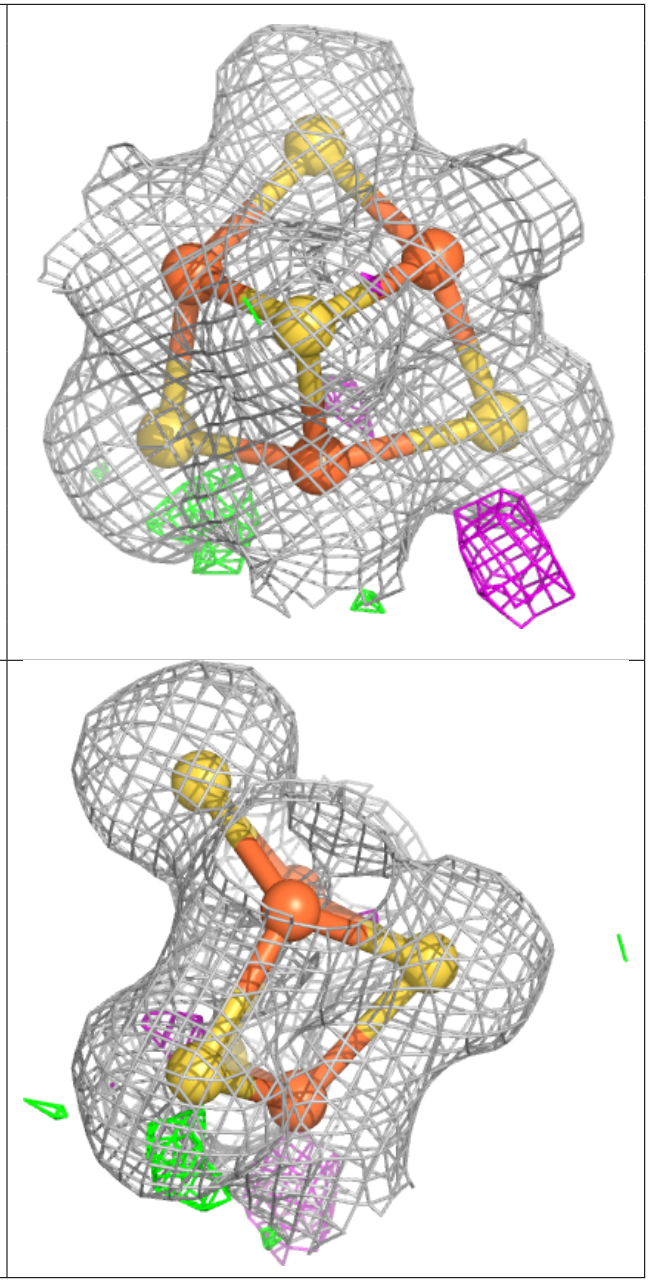






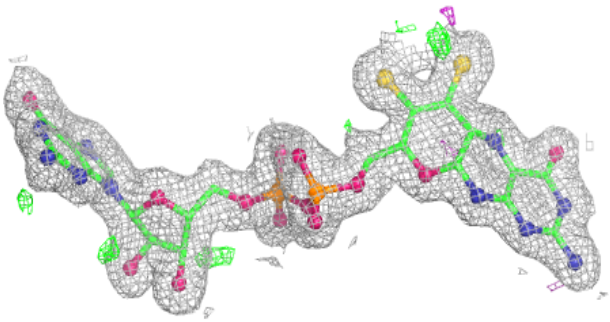
Electron density around F3S A 5005:

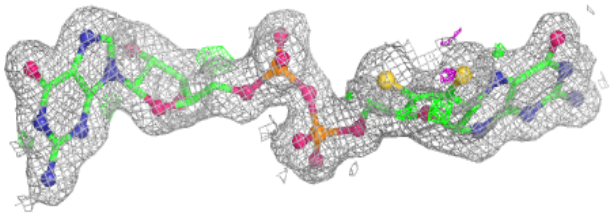
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

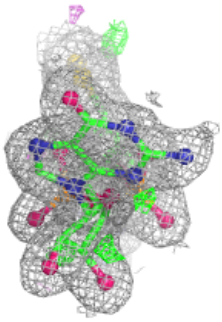


Electron density around MGD A 5001:

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

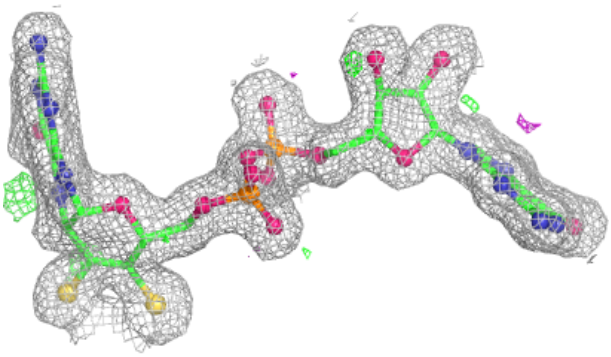


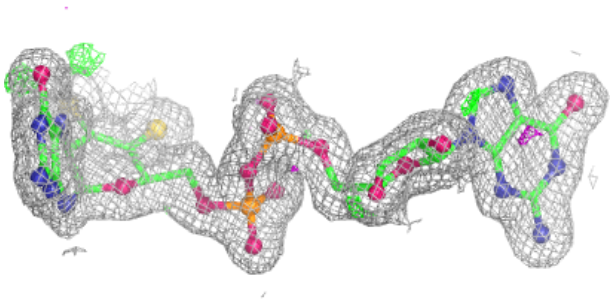


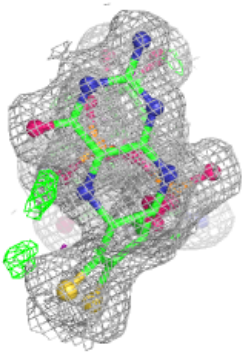


Electron density around MGD C 5102:

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

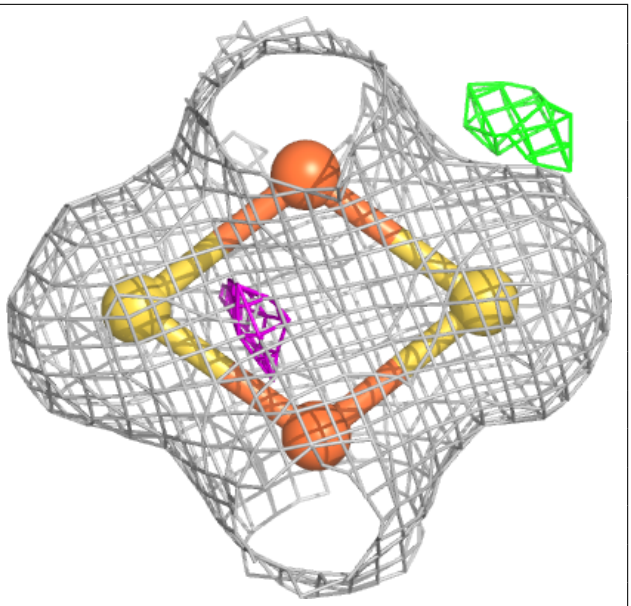


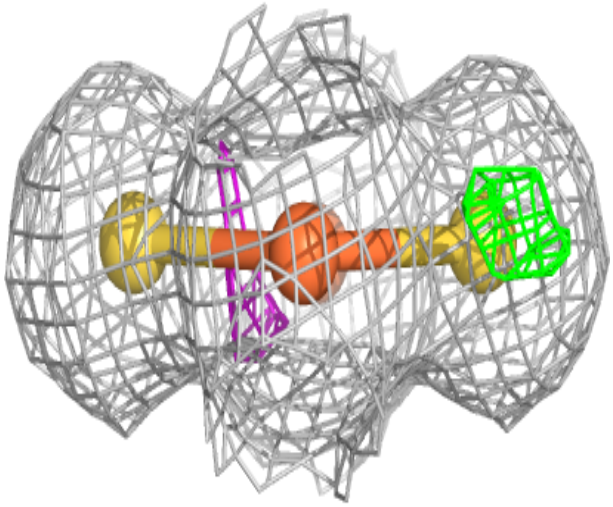


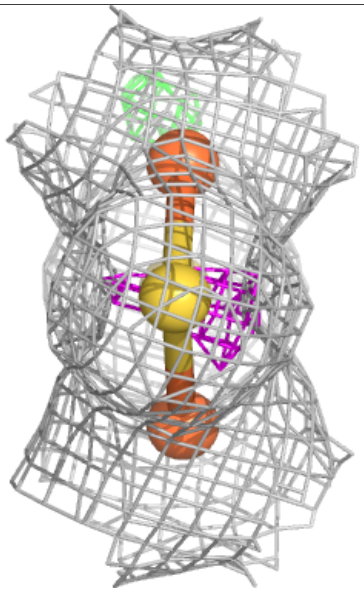


Electron density around FES D 201:

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

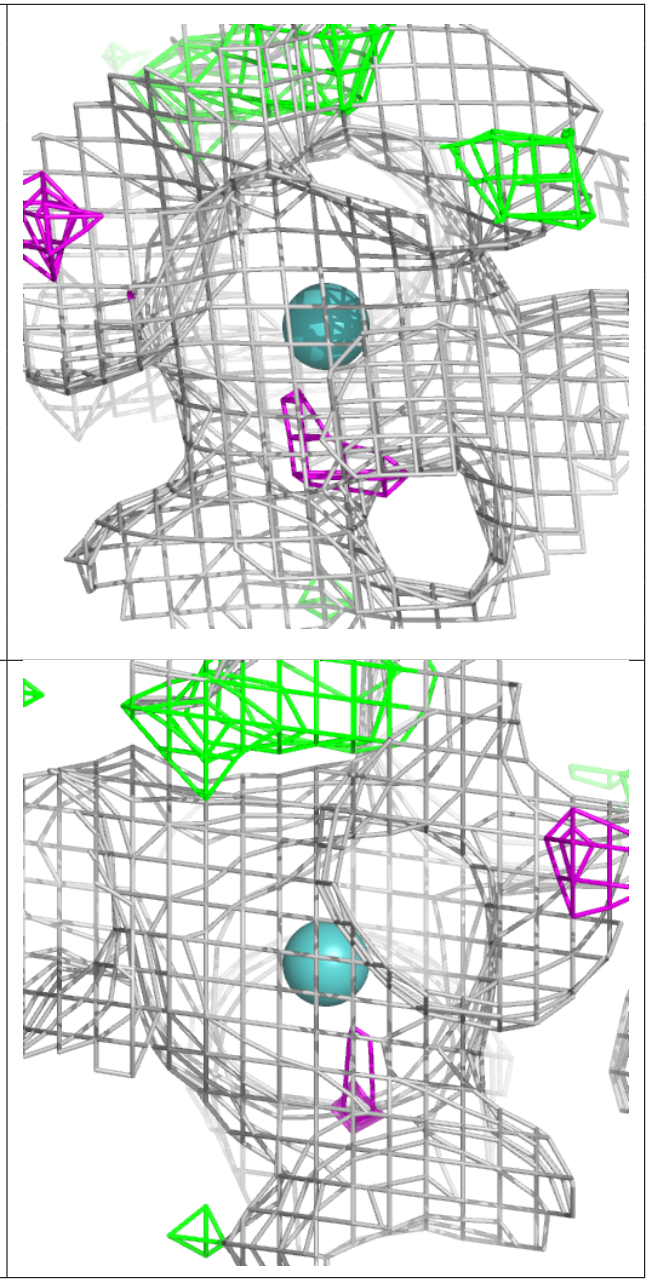






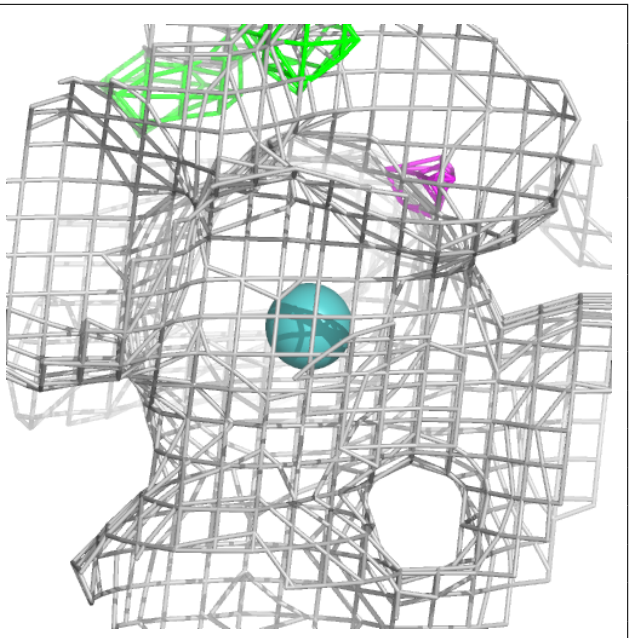
Electron density around 4MO G 5004:

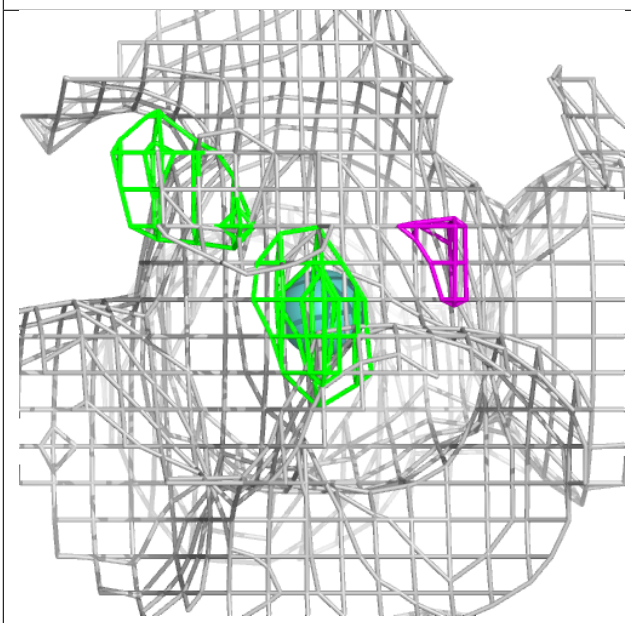
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

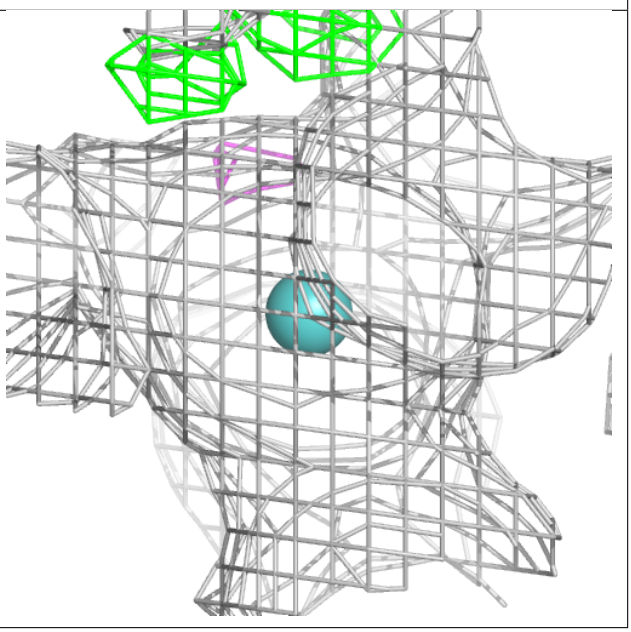


Electron density around 4MO A 5004:

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

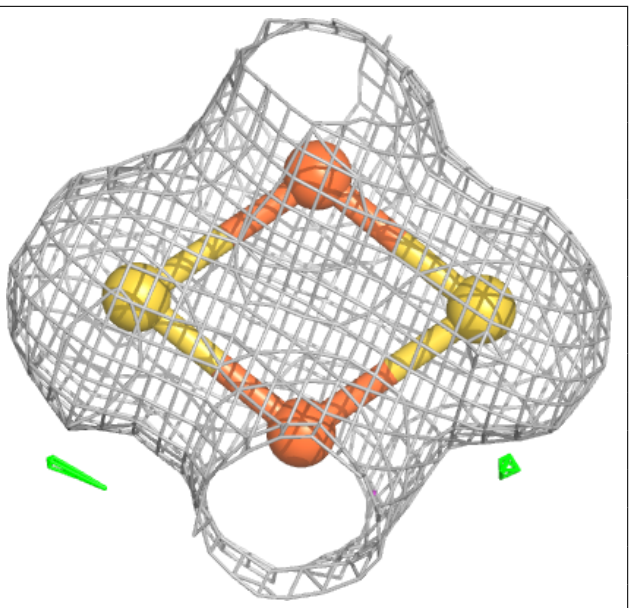


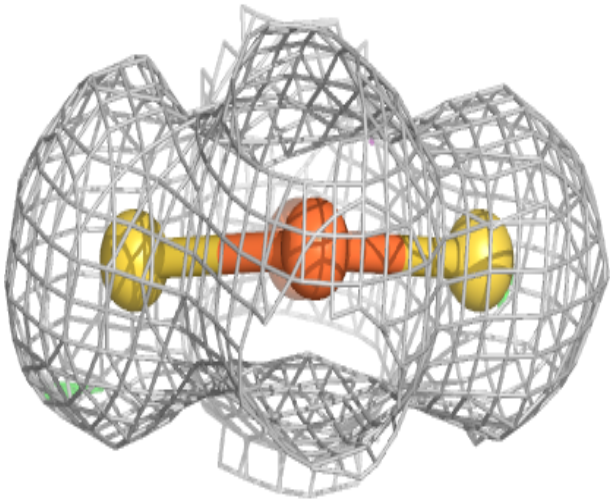


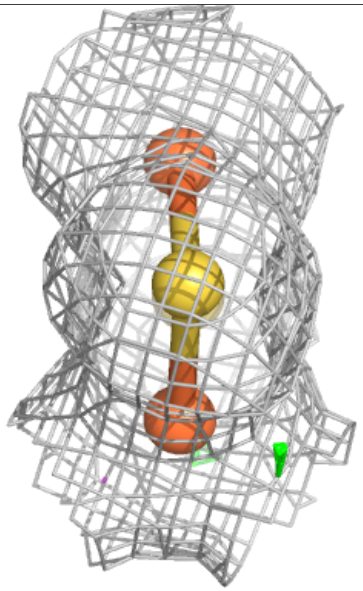


Electron density around FES B 201:

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

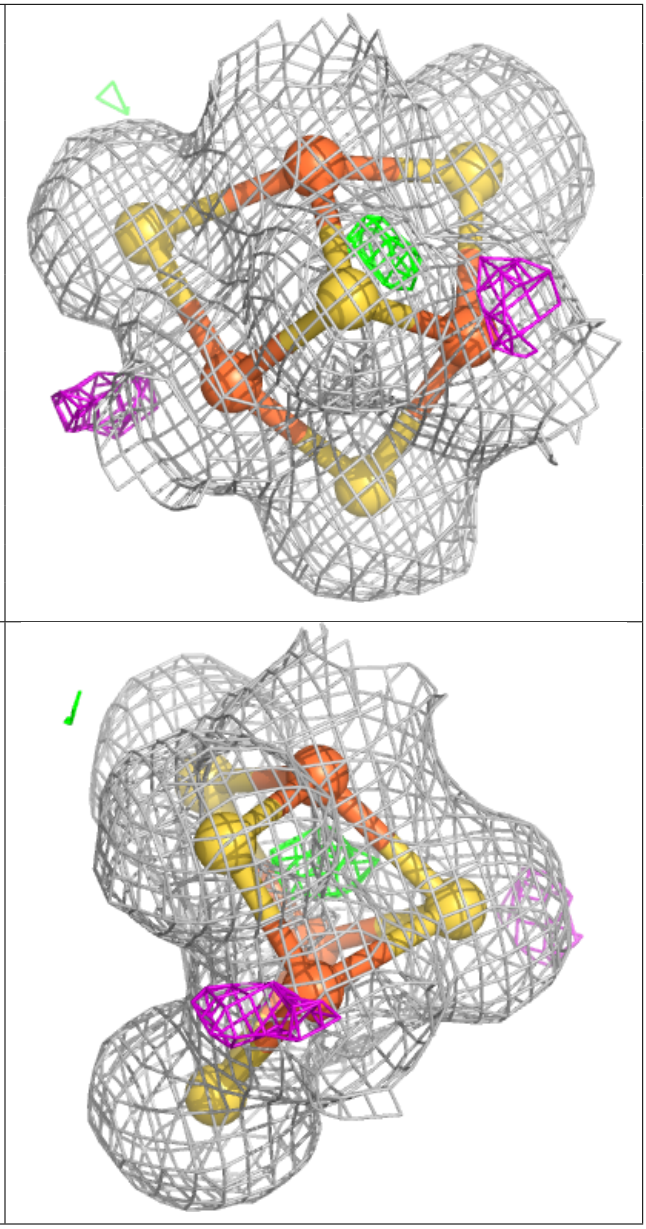






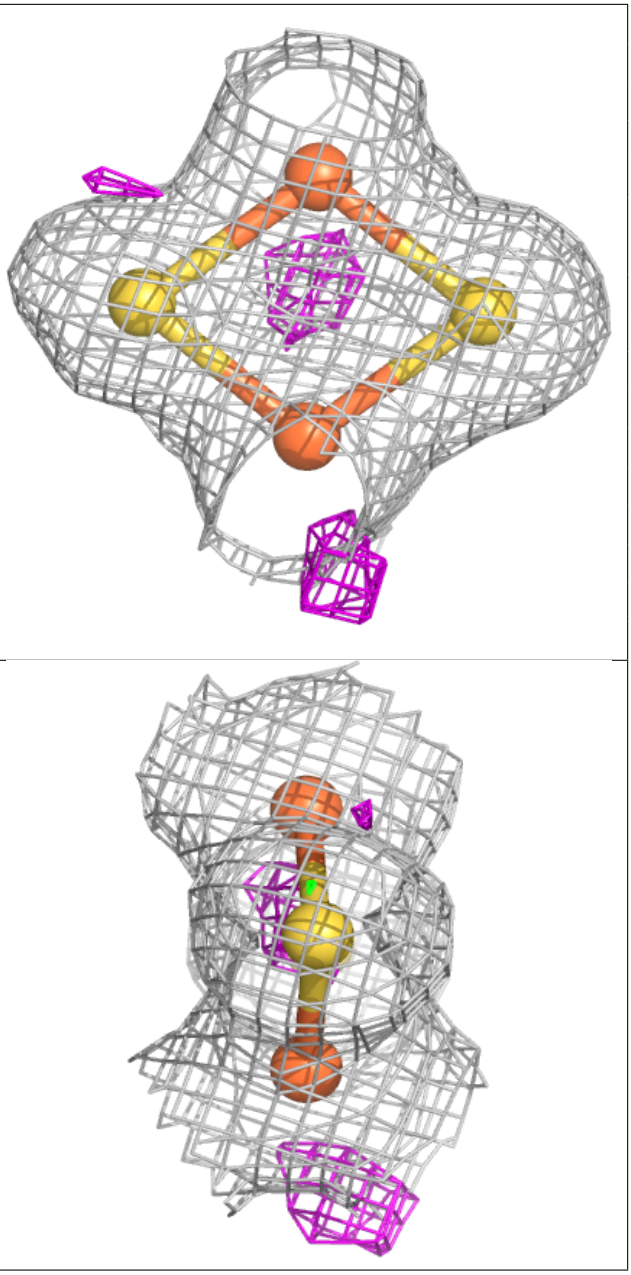
Electron density around F3S C 5105:

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



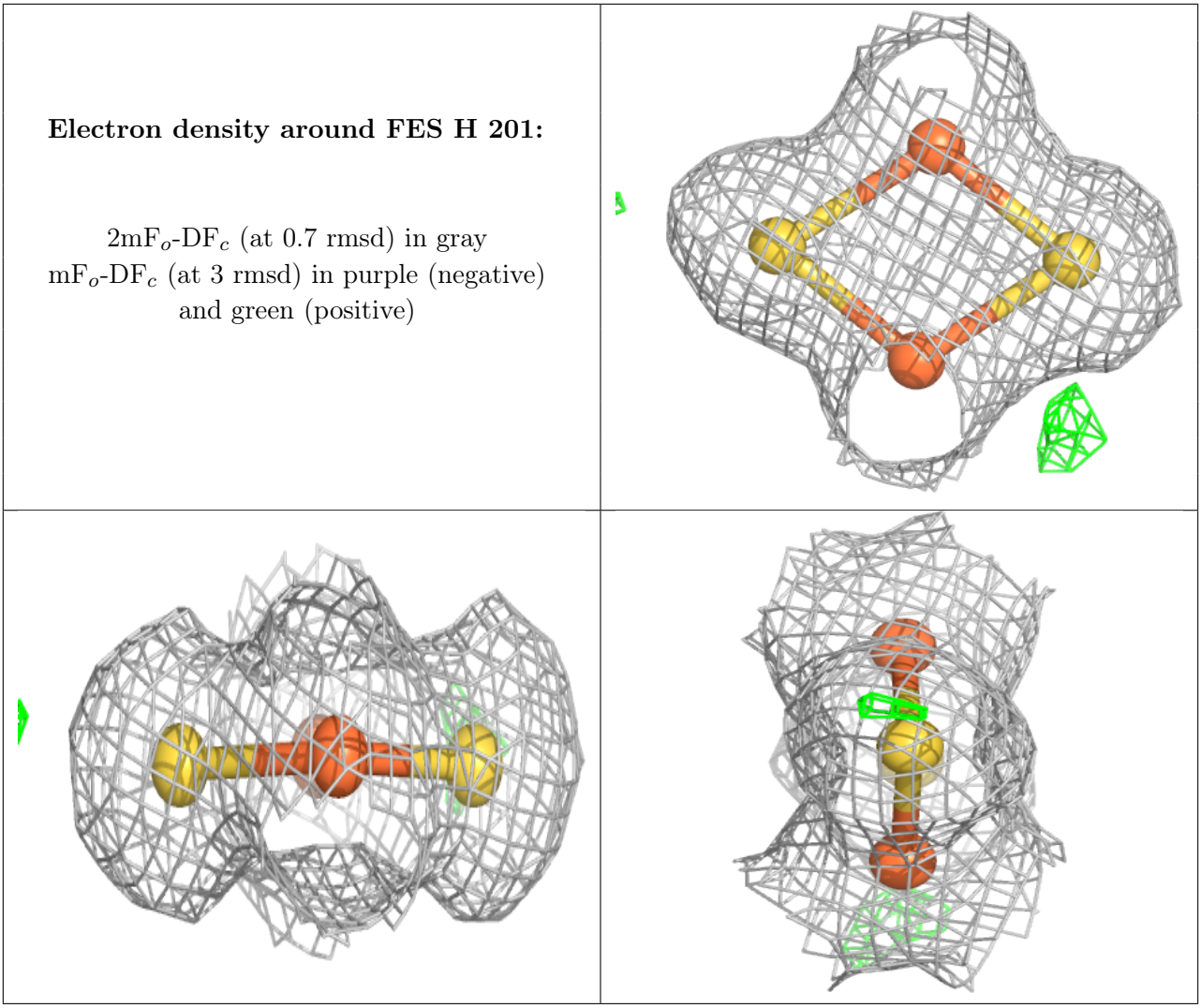
Electron density around FES F 201:

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



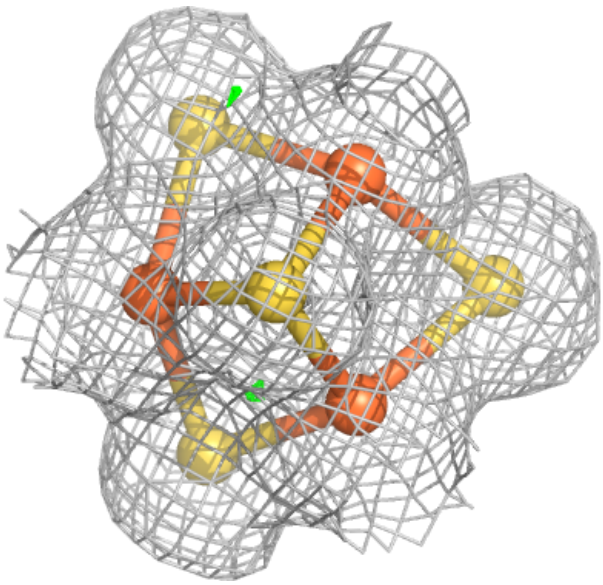
Electron density around FES H 201:

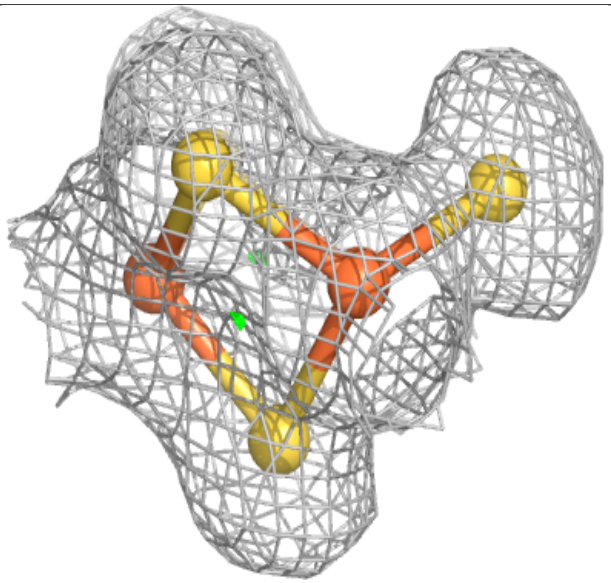
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

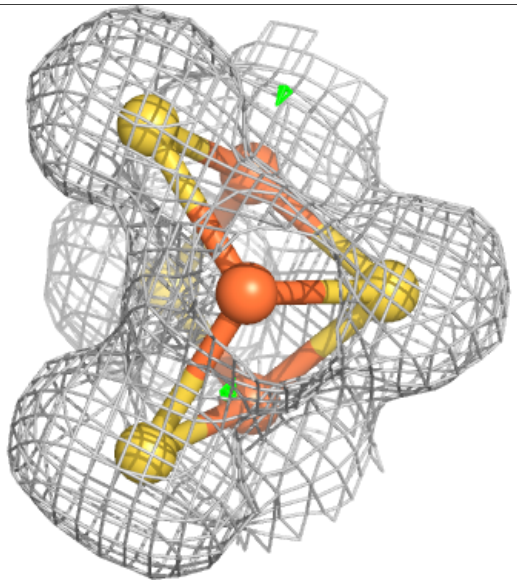


Electron density around F3S E 5005:

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

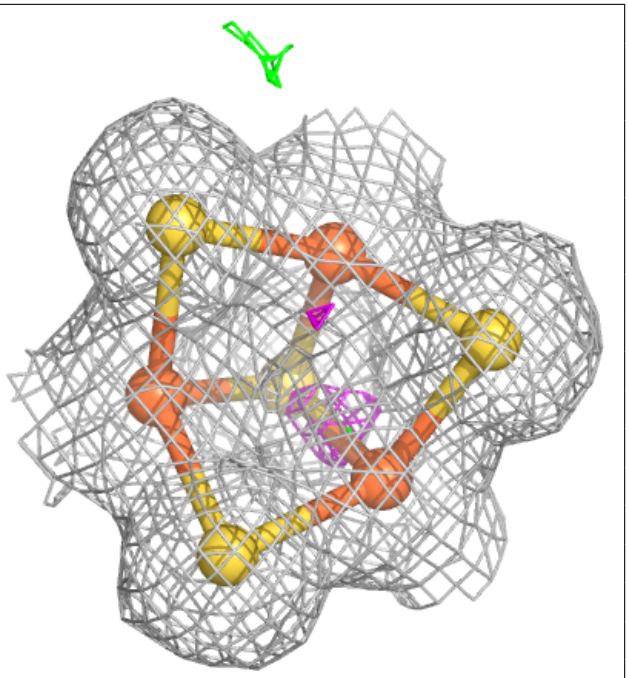


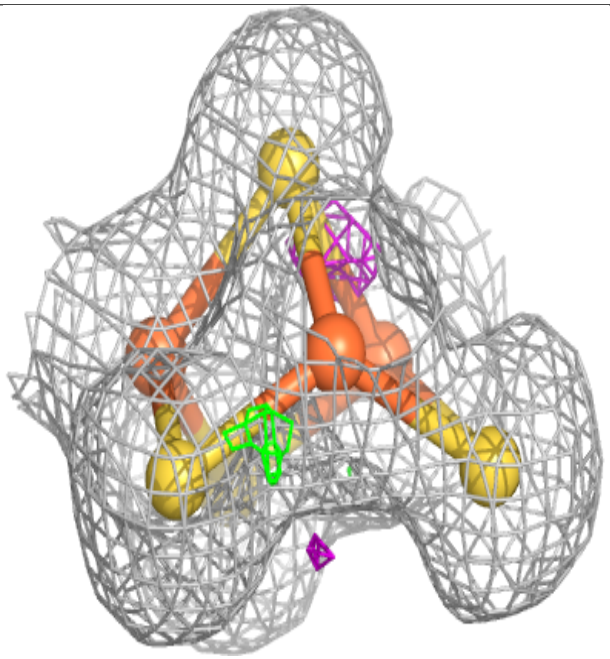


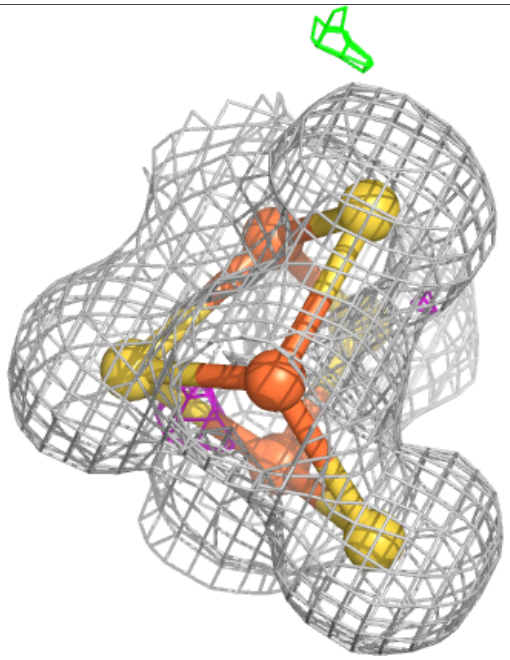


Electron density around F3S G 5005:

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

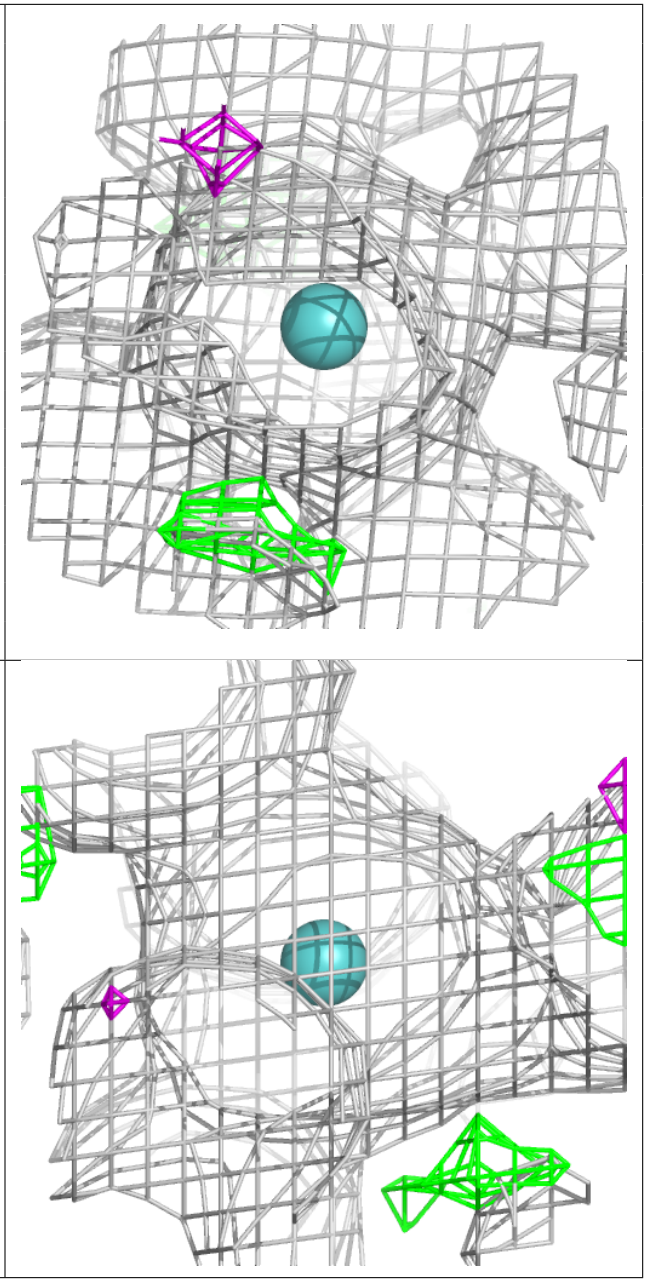






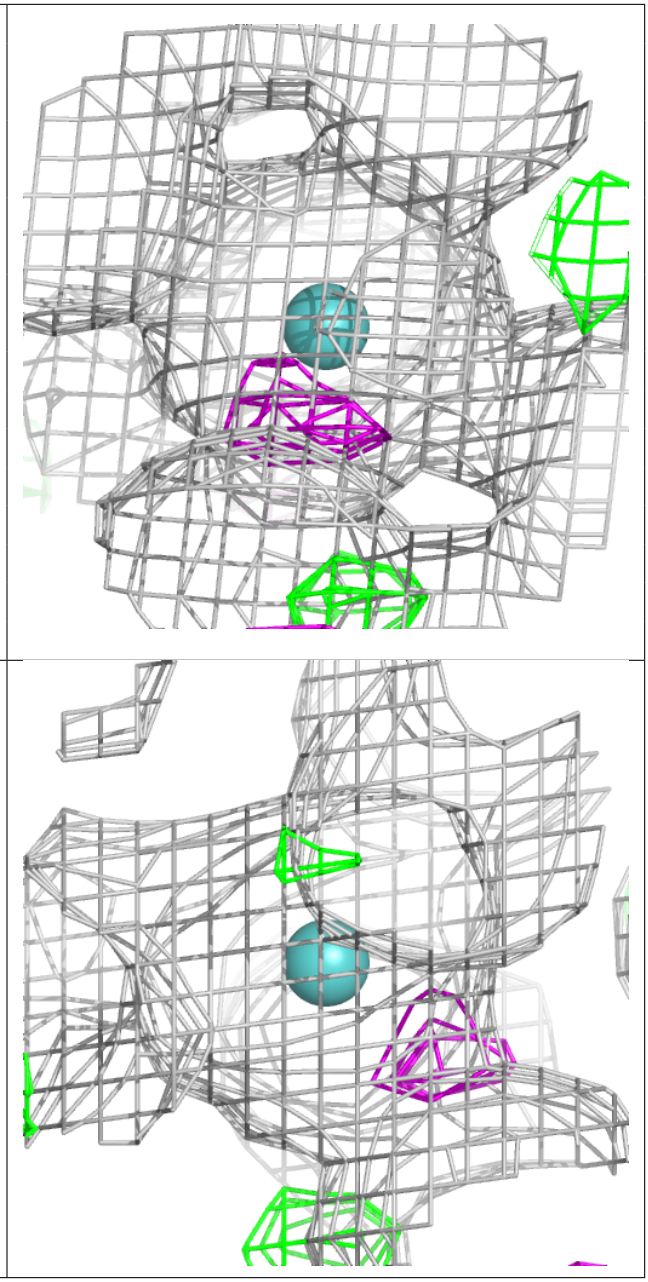
Electron density around 4MO C 5104:

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



Electron density around 4MO E 5004:

$2mF_o$-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.