wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2026 – 03:09 pm BST

PDB ID	:	9G7I / pdb_00009g7i
Title	:	Structure of carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) in complex with acetyl-Coenyzme A from Clostridium autoethanogenum
Authors	:	Lemaire, O.N.; Yin, M.D.; Murphy, B.J.; Wagner, T.
Deposited on	:	2024-07-21
Resolution	:	2.93 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:

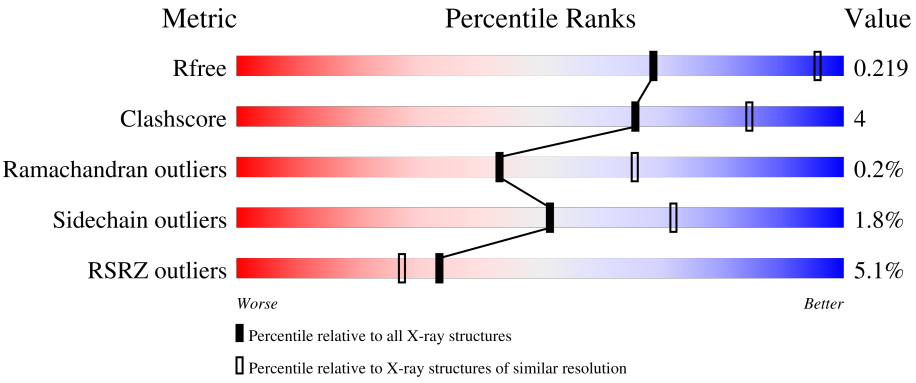
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal 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 2.93 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	1159 (2.96-2.92)
Clashscore	190562	1184 (2.96-2.92)
Ramachandran outliers	187476	1131 (2.96-2.92)
Sidechain outliers	187428	1131 (2.96-2.92)
RSRZ outliers	180081	1159 (2.96-2.92)

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	708	15% 82% 15%
1	D	708	3% 92% 8%
2	B	631	92% 7%
2	C	631	% 92% 8%

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
5	EDO	B	717	-	-	-	X

2 Entry composition (i)

There are 14 unique types of molecules in this entry. The entry contains 20497 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.

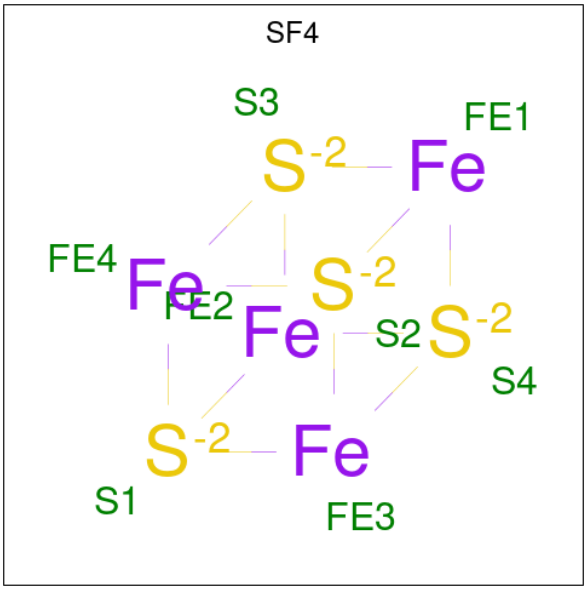
- Molecule 1 is a protein called CO-methylating acetyl-CoA synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	694	Total	C	N	O	S	0	1	0
			5284	3383	860	1006	35			
1	D	708	Total	C	N	O	S	0	0	0
			5392	3449	884	1023	36			

- Molecule 2 is a protein called Carbon monoxide dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	629	Total	C	N	O	S	0	0	0
			4723	2967	814	902	40			
2	C	630	Total	C	N	O	S	0	1	0
			4742	2978	818	906	40			

- Molecule 3 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe_4S_4).

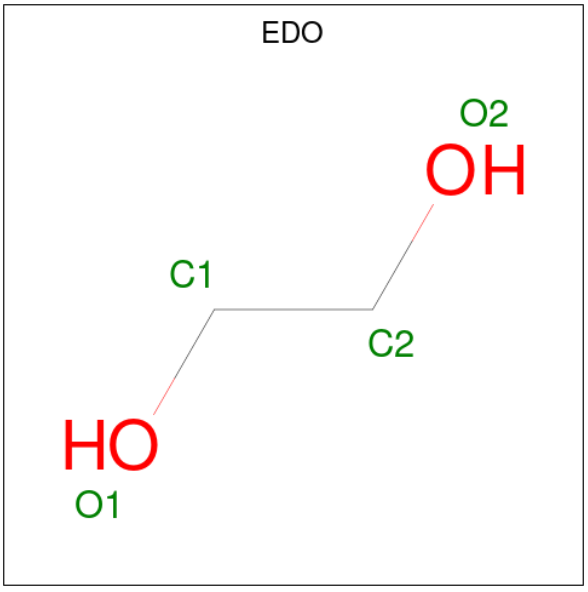


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Fe S 8 4 4	0	0
3	B	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	D	1	Total Fe S 8 4 4	0	0

- Molecule 4 is NICKEL (II) ION (CCD ID: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 5 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



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

Continued on next page...

Continued from previous page...

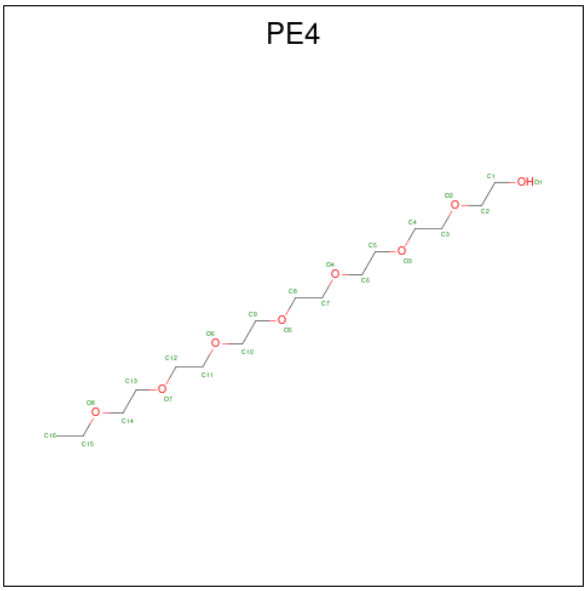
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

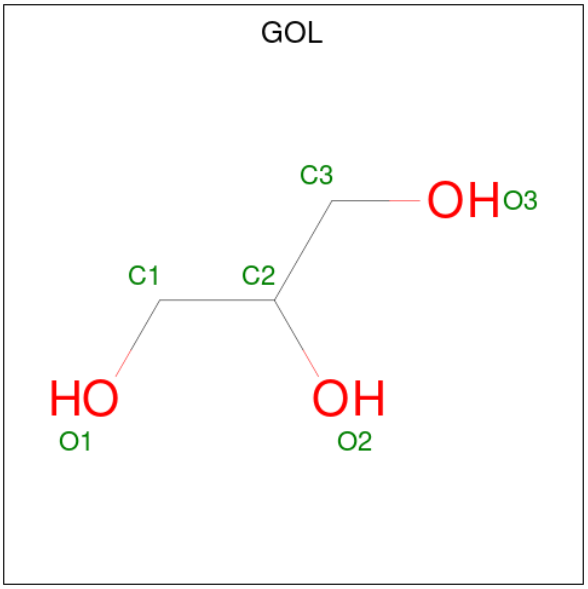
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0

- Molecule 6 is 2-{2-[2-(2-{2-[2-(2-ETHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHOXY)-ETHOXY]-ETHOXY}-ETHANOL (CCD ID: PE4) (formula: $C_{16}H_{34}O_8$).



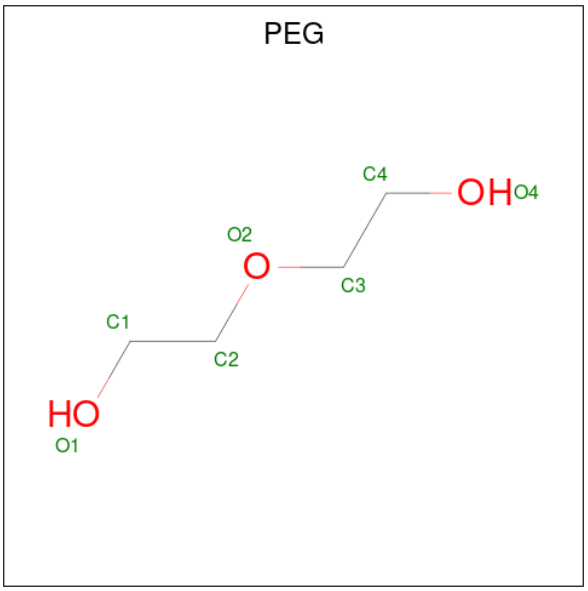
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 10 6 4	0	0

- Molecule 7 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	D	1	Total C O 6 3 3	0	0

- Molecule 8 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 9 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	4	Total Ca 4 4	0	0
9	B	1	Total Ca 1 1	0	0
9	C	3	Total Ca 3 3	0	0
9	D	2	Total Ca 2 2	0	0

- Molecule 10 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

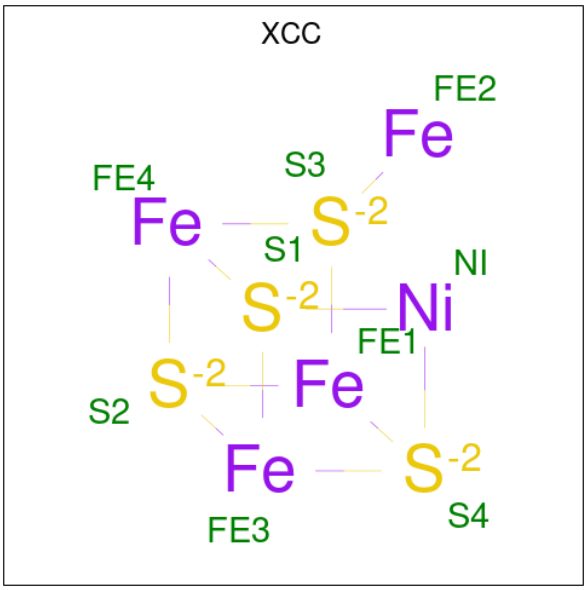
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	2	Total Cl 2 2	0	0
10	B	5	Total Cl 5 5	0	0

Continued on next page...

Continued from previous page...

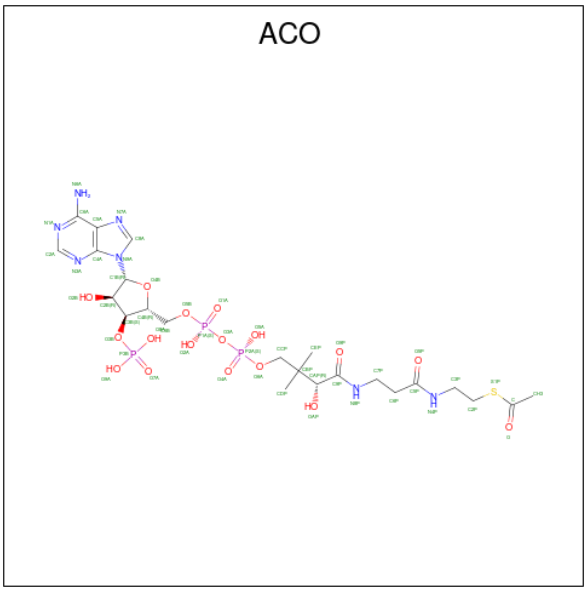
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	C	1	Total Cl 1 1	0	0

- Molecule 11 is FE(4)-NI(1)-S(4) CLUSTER (CCD ID: XCC) (formula: Fe_4NiS_4) (labeled as "Ligand of Interest" by depositor).



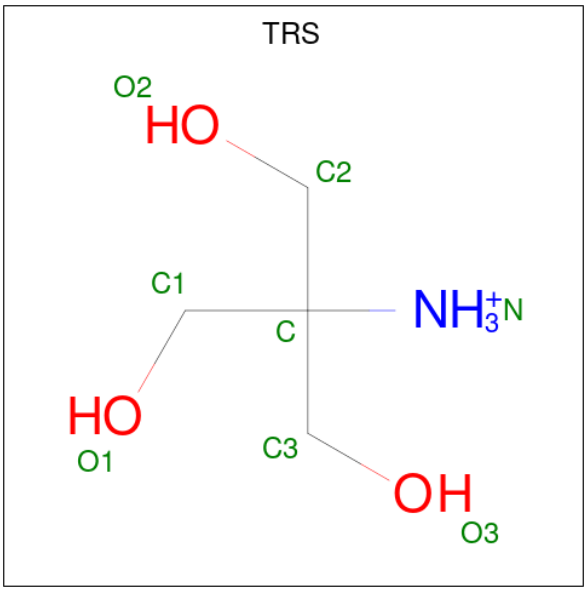
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	B	1	Total Fe Ni S 9 4 1 4	0	0
11	C	1	Total Fe Ni S 9 4 1 4	0	0

- Molecule 12 is ACETYL COENZYME *A (CCD ID: ACO) (formula: $C_{23}H_{38}N_7O_{17}P_3S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
12	D	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

- 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
13	D	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 14 is water.

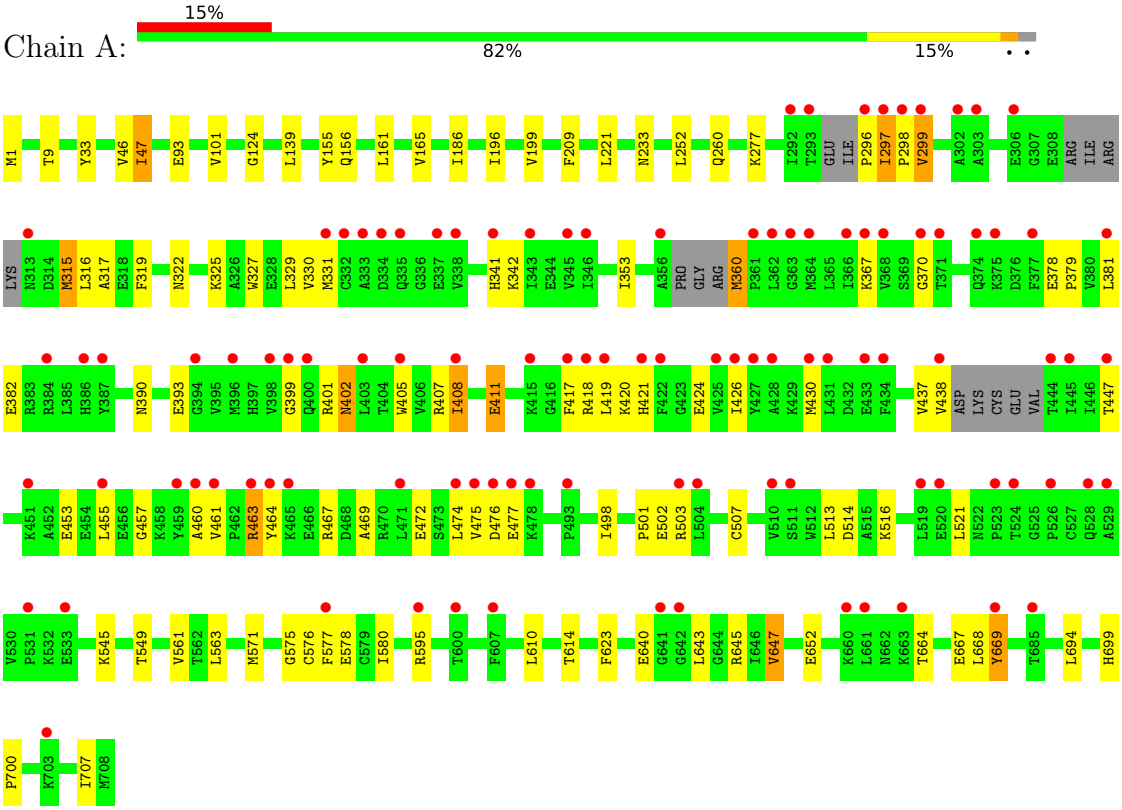
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	A	2	Total O 2 2	0	0
14	B	5	Total O 5 5	0	0
14	C	5	Total O 5 5	0	0
14	D	2	Total O 2 2	0	0

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.

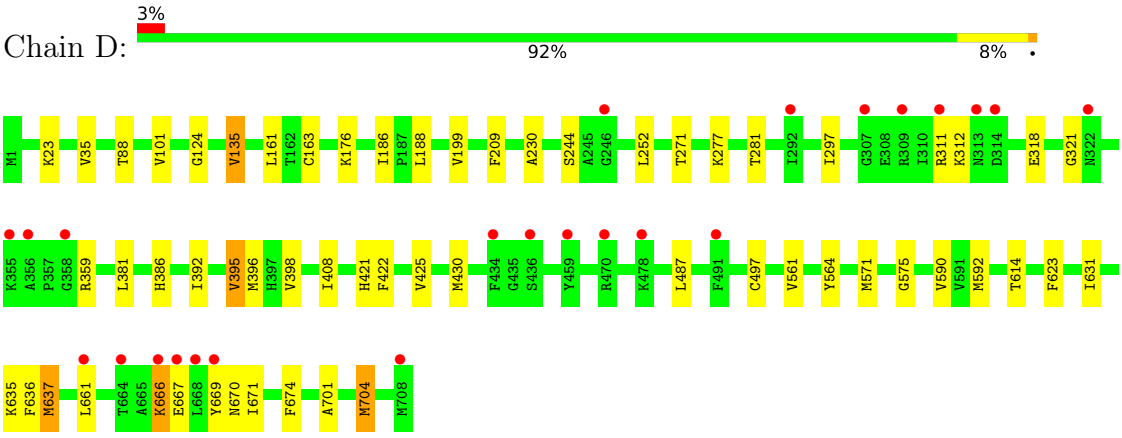
• Molecule 1: CO-methylating acetyl-CoA synthase

Chain A:



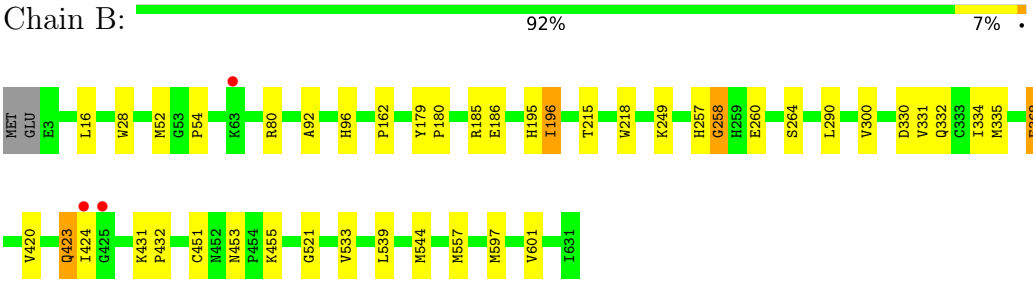
• Molecule 1: CO-methylating acetyl-CoA synthase

Chain D:



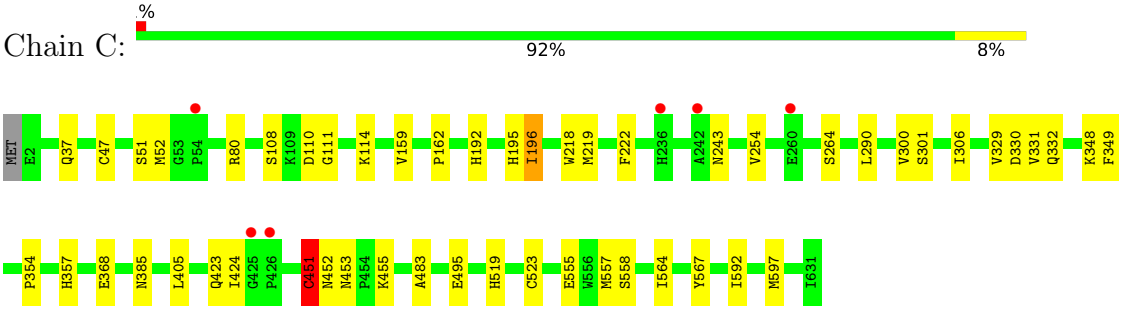
• Molecule 2: Carbon monoxide dehydrogenase

Chain B:



• Molecule 2: Carbon monoxide dehydrogenase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	298.08Å 298.08Å 127.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.68 – 2.93 49.68 – 2.93	Depositor EDS
% Data completeness (in resolution range)	81.4 (49.68-2.93) 81.4 (49.68-2.93)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 2.96Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.190 , 0.220 0.190 , 0.219	Depositor DCC
R_{free} test set	5007 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	84.5	Xtriage
Anisotropy	0.013	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 64.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	20497	wwPDB-VP
Average B, all atoms (Å^2)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, ACO, PE4, TRS, GOL, XCC, SF4, EDO, PEG, CA, NI

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	0.70	0/5375	0.97	5/7271 (0.1%)
1	D	0.64	0/5488	0.90	4/7428 (0.1%)
2	B	0.64	0/4797	0.96	5/6488 (0.1%)
2	C	0.65	0/4817	0.95	2/6515 (0.0%)
All	All	0.66	0/20477	0.95	16/27702 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	402	ASN	N-CA-C	-7.75	101.68	112.45
1	D	176	LYS	N-CA-C	7.14	121.25	111.39
2	B	521	GLY	N-CA-C	7.12	118.75	111.56
1	D	590	VAL	N-CA-C	6.35	118.54	108.89
2	B	52	MET	N-CA-C	-6.22	104.60	111.82

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	5284	0	5365	70	0
1	D	5392	0	5489	33	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	4723	0	4803	31	0
2	C	4742	0	4818	32	0
3	A	8	0	0	1	0
3	B	8	0	0	1	0
3	C	16	0	0	1	0
3	D	8	0	0	1	0
4	A	2	0	0	0	0
4	D	2	0	0	0	0
5	A	32	0	48	0	0
5	B	60	0	90	0	0
5	C	28	0	42	0	0
5	D	16	0	24	0	0
6	A	10	0	12	0	0
7	A	6	0	8	0	0
7	B	24	0	32	0	0
7	D	6	0	8	0	0
8	A	7	0	10	0	0
8	B	7	0	10	0	0
8	D	7	0	10	0	0
9	A	4	0	0	0	0
9	B	1	0	0	0	0
9	C	3	0	0	0	0
9	D	2	0	0	0	0
10	A	2	0	0	0	0
10	B	5	0	0	0	0
10	C	1	0	0	0	0
11	B	9	0	0	0	0
11	C	9	0	0	0	0
12	D	51	0	34	2	0
13	D	8	0	12	0	0
14	A	2	0	0	0	0
14	B	5	0	0	0	0
14	C	5	0	0	0	0
14	D	2	0	0	0	0
All	All	20497	0	20815	157	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 157 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:C:424:ILE:HD11	1:D:281:THR:HB	1.64	0.79
1:A:299:VAL:HG23	1:A:393:GLU:HB3	1.73	0.71
1:A:316:LEU:HB3	1:A:408:ILE:HD11	1.73	0.71
1:D:277:LYS:O	1:D:281:THR:HG23	1.92	0.70
1:D:701:ALA:HA	1:D:704:MET:HE3	1.75	0.69

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	685/708 (97%)	649 (95%)	35 (5%)	1 (0%)	48	71
1	D	706/708 (100%)	670 (95%)	35 (5%)	1 (0%)	48	71
2	B	627/631 (99%)	591 (94%)	35 (6%)	1 (0%)	43	65
2	C	629/631 (100%)	601 (96%)	26 (4%)	2 (0%)	36	59
All	All	2647/2678 (99%)	2511 (95%)	131 (5%)	5 (0%)	43	65

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	332	GLN
1	A	575	GLY
2	B	332	GLN
2	C	243	ASN
1	D	575	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	571/583 (98%)	553 (97%)	18 (3%)	34	59
1	D	583/583 (100%)	573 (98%)	10 (2%)	53	73
2	B	518/520 (100%)	513 (99%)	5 (1%)	68	81
2	C	520/520 (100%)	514 (99%)	6 (1%)	63	78
All	All	2192/2206 (99%)	2153 (98%)	39 (2%)	51	72

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	523	CYS
1	D	637	MET
1	D	23	LYS
1	D	395	VAL
1	D	671	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	423	GLN
2	C	428	GLN
1	D	594	ASN
1	D	5	GLN
1	D	372	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 75 ligands modelled in this entry, 22 are monoatomic - leaving 53 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	$\#\|Z\| > 2$
5	EDO	B	715	-	3,3,3	0.33	0	2,2,2	0.20	0
5	EDO	B	718	-	3,3,3	0.07	0	2,2,2	0.20	0
3	SF4	A	801	1	0,12,12	-	-	-		
5	EDO	B	726	-	3,3,3	0.06	0	2,2,2	0.19	0
5	EDO	C	703	-	3,3,3	0.33	0	2,2,2	0.18	0
5	EDO	C	711	-	3,3,3	0.25	0	2,2,2	0.32	0
5	EDO	B	725	-	3,3,3	0.05	0	2,2,2	0.18	0
3	SF4	B	702	2	0,12,12	-	-	-		
5	EDO	B	716	-	3,3,3	0.26	0	2,2,2	0.28	0
5	EDO	B	717	-	3,3,3	0.16	0	2,2,2	0.48	0
5	EDO	B	722	-	3,3,3	0.04	0	2,2,2	0.12	0
5	EDO	C	710	-	3,3,3	0.30	0	2,2,2	0.18	0
5	EDO	D	812	-	3,3,3	0.29	0	2,2,2	0.20	0
5	EDO	D	813	-	3,3,3	0.06	0	2,2,2	0.18	0
7	GOL	B	701	-	5,5,5	0.07	0	5,5,5	0.20	0
7	GOL	B	707	-	5,5,5	0.09	0	5,5,5	0.33	0
5	EDO	A	804	-	3,3,3	0.18	0	2,2,2	0.16	0
5	EDO	A	820	-	3,3,3	0.06	0	2,2,2	0.19	0
3	SF4	C	701	2	0,12,12	-	-	-		
5	EDO	A	819	-	3,3,3	0.06	0	2,2,2	0.19	0
7	GOL	B	706	-	5,5,5	0.09	0	5,5,5	0.31	0
5	EDO	D	811	-	3,3,3	0.23	0	2,2,2	0.44	0
11	XCC	C	704	2	0,12,12	-	-	-		
11	XCC	B	705	2	0,12,12	-	-	-		
5	EDO	B	723	-	3,3,3	0.07	0	2,2,2	0.21	0
3	SF4	D	801	1	0,12,12	-	-	-		
5	EDO	B	721	-	3,3,3	0.06	0	2,2,2	0.19	0
5	EDO	B	728	-	3,3,3	0.06	0	2,2,2	0.20	0
7	GOL	B	704	-	5,5,5	0.04	0	5,5,5	0.20	0
8	PEG	D	804	-	6,6,6	0.20	0	5,5,5	0.20	0
3	SF4	C	702	2	0,12,12	-	-	-		
5	EDO	C	712	-	3,3,3	0.36	0	2,2,2	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	EDO	D	810	-	3,3,3	0.31	0	2,2,2	0.21	0
7	GOL	A	807	-	5,5,5	0.04	0	5,5,5	0.18	0
5	EDO	B	719	-	3,3,3	0.07	0	2,2,2	0.20	0
5	EDO	C	713	-	3,3,3	0.06	0	2,2,2	0.19	0
13	TRS	D	807	-	7,7,7	0.13	0	9,9,9	0.19	0
5	EDO	B	720	-	3,3,3	0.07	0	2,2,2	0.18	0
5	EDO	A	817	-	3,3,3	0.29	0	2,2,2	0.22	0
5	EDO	B	724	-	3,3,3	0.05	0	2,2,2	0.21	0
5	EDO	A	818	-	3,3,3	0.41	0	2,2,2	0.08	0
5	EDO	C	714	-	3,3,3	0.05	0	2,2,2	0.18	0
5	EDO	B	727	-	3,3,3	0.06	0	2,2,2	0.18	0
7	GOL	D	806	-	5,5,5	0.09	0	5,5,5	0.32	0
8	PEG	B	703	-	6,6,6	0.24	0	5,5,5	0.27	0
5	EDO	A	806	-	3,3,3	0.38	0	2,2,2	0.10	0
12	ACO	D	805	-	48,53,53	1.99	11 (22%)	69,79,79	2.32	12 (17%)
5	EDO	A	816	-	3,3,3	0.26	0	2,2,2	0.31	0
6	PE4	A	805	-	9,9,23	0.49	0	8,8,22	0.63	0
5	EDO	C	709	-	3,3,3	0.33	0	2,2,2	0.20	0
5	EDO	A	815	-	3,3,3	0.30	0	2,2,2	0.32	0
8	PEG	A	808	-	6,6,6	0.17	0	5,5,5	0.18	0
5	EDO	B	714	-	3,3,3	0.32	0	2,2,2	0.22	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
5	EDO	B	715	-	-	1/1/1/1	-
5	EDO	B	718	-	-	1/1/1/1	-
3	SF4	A	801	1	-	-	0/6/5/5
5	EDO	B	726	-	-	0/1/1/1	-
5	EDO	C	703	-	-	1/1/1/1	-
5	EDO	C	711	-	-	0/1/1/1	-
5	EDO	B	725	-	-	0/1/1/1	-
5	EDO	C	710	-	-	0/1/1/1	-
5	EDO	B	716	-	-	0/1/1/1	-
5	EDO	B	717	-	-	0/1/1/1	-
5	EDO	B	722	-	-	1/1/1/1	-
5	EDO	D	812	-	-	1/1/1/1	-
5	EDO	D	813	-	-	1/1/1/1	-
7	GOL	B	701	-	-	0/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	B	707	-	-	1/4/4/4	-
3	SF4	B	702	2	-	-	0/6/5/5
5	EDO	A	804	-	-	1/1/1/1	-
5	EDO	A	820	-	-	0/1/1/1	-
5	EDO	A	819	-	-	1/1/1/1	-
3	SF4	C	701	2	-	-	0/6/5/5
7	GOL	B	706	-	-	0/4/4/4	-
5	EDO	D	811	-	-	0/1/1/1	-
11	XCC	C	704	2	-	-	0/4/4/4
11	XCC	B	705	2	-	-	0/4/4/4
5	EDO	B	723	-	-	0/1/1/1	-
5	EDO	B	728	-	-	1/1/1/1	-
5	EDO	B	721	-	-	0/1/1/1	-
7	GOL	B	704	-	-	0/4/4/4	-
8	PEG	D	804	-	-	1/4/4/4	-
3	SF4	D	801	1	-	-	0/6/5/5
5	EDO	D	810	-	-	1/1/1/1	-
5	EDO	C	712	-	-	1/1/1/1	-
3	SF4	C	702	2	-	-	0/6/5/5
7	GOL	A	807	-	-	0/4/4/4	-
5	EDO	B	719	-	-	1/1/1/1	-
5	EDO	C	713	-	-	0/1/1/1	-
13	TRS	D	807	-	-	8/9/9/9	-
5	EDO	B	720	-	-	0/1/1/1	-
5	EDO	A	817	-	-	0/1/1/1	-
5	EDO	B	724	-	-	1/1/1/1	-
5	EDO	A	818	-	-	1/1/1/1	-
5	EDO	C	714	-	-	1/1/1/1	-
5	EDO	B	727	-	-	0/1/1/1	-
7	GOL	D	806	-	-	2/4/4/4	-
8	PEG	B	703	-	-	2/4/4/4	-
5	EDO	A	806	-	-	1/1/1/1	-
12	ACO	D	805	-	-	24/51/67/67	0/3/3/3
5	EDO	A	816	-	-	0/1/1/1	-
6	PE4	A	805	-	-	3/7/7/21	-
5	EDO	C	709	-	-	1/1/1/1	-
5	EDO	A	815	-	-	0/1/1/1	-
8	PEG	A	808	-	-	1/4/4/4	-
5	EDO	B	714	-	-	1/1/1/1	-

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	D	805	ACO	C9P-N8P	5.80	1.46	1.33
12	D	805	ACO	C5P-N4P	5.64	1.46	1.33
12	D	805	ACO	C1B-N9A	-4.61	1.33	1.46
12	D	805	ACO	C6A-N6A	4.54	1.45	1.34
12	D	805	ACO	C3B-C4B	3.95	1.63	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	D	805	ACO	C4A-N9A-C8A	9.25	115.75	105.73
12	D	805	ACO	N9A-C8A-N7A	-8.43	102.39	113.91
12	D	805	ACO	N3A-C4A-N9A	5.68	136.44	127.08
12	D	805	ACO	N3A-C2A-N1A	-5.59	119.86	128.60
12	D	805	ACO	C1B-N9A-C8A	-5.31	115.15	127.14

There are no chirality outliers.

5 of 60 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	806	EDO	O1-C1-C2-O2
12	D	805	ACO	O4B-C4B-C5B-O5B
12	D	805	ACO	C5B-O5B-P1A-O1A
12	D	805	ACO	C5B-O5B-P1A-O2A
12	D	805	ACO	OAP-CAP-CBP-CCP

There are no ring outliers.

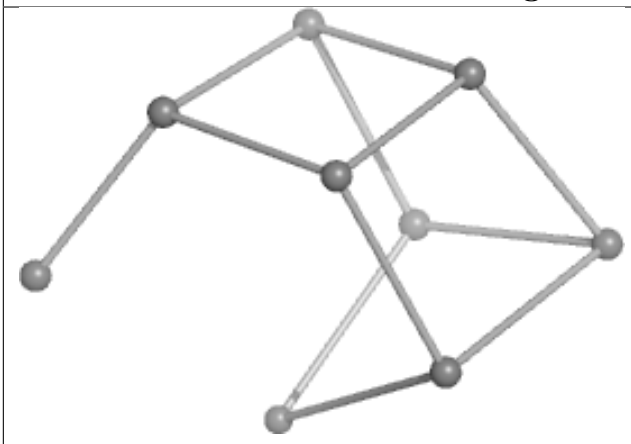
5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	801	SF4	1	0
3	B	702	SF4	1	0
3	D	801	SF4	1	0
3	C	702	SF4	1	0
12	D	805	ACO	2	0

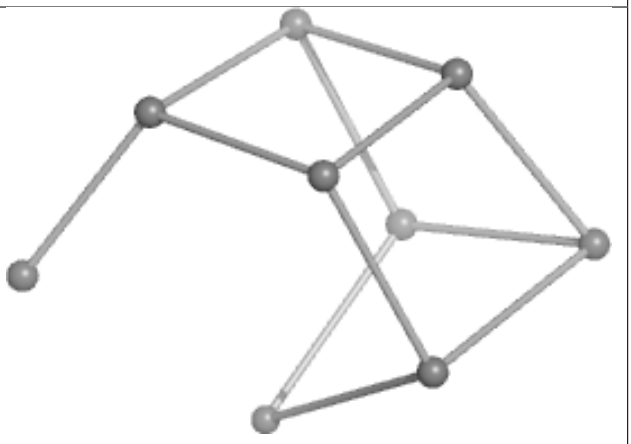
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

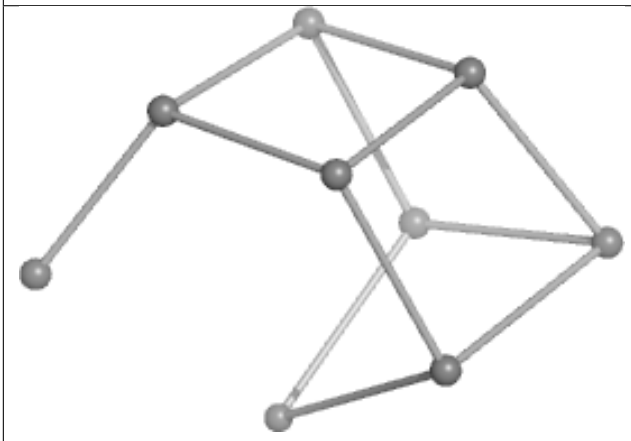
Ligand XCC C 704



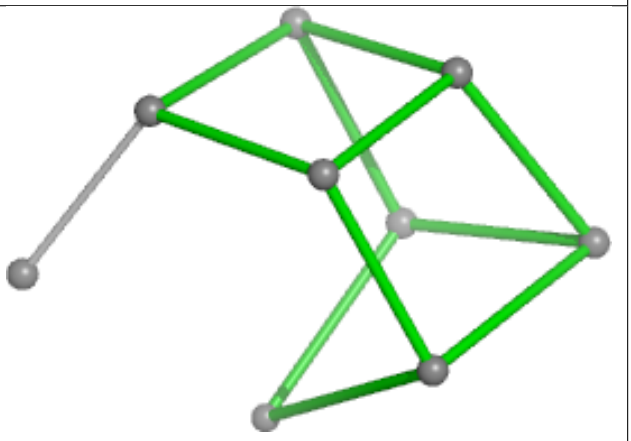
Bond lengths



Bond angles

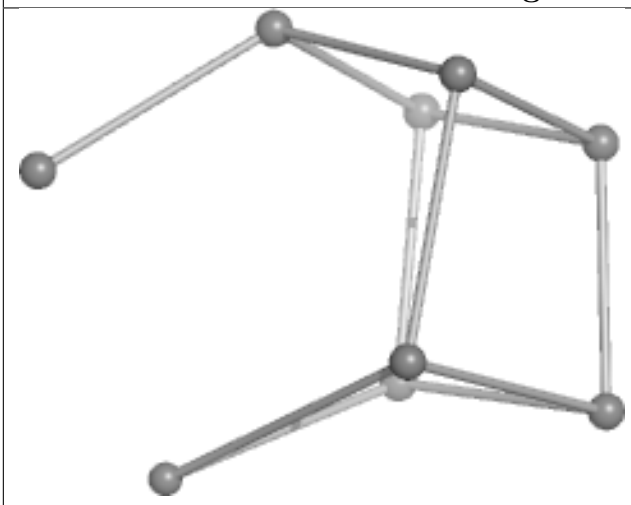


Torsions

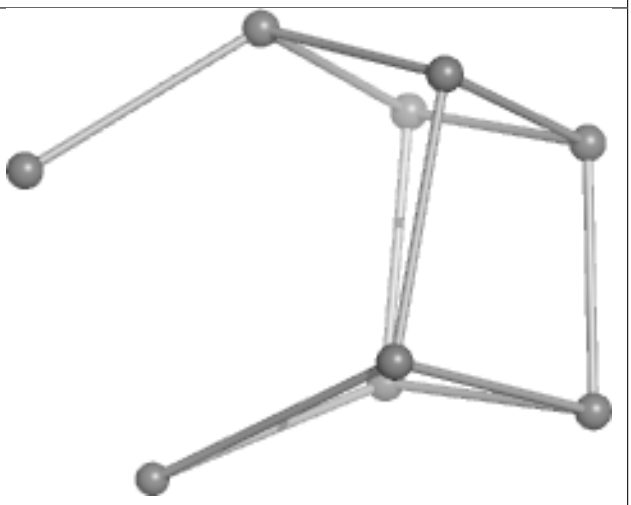


Rings

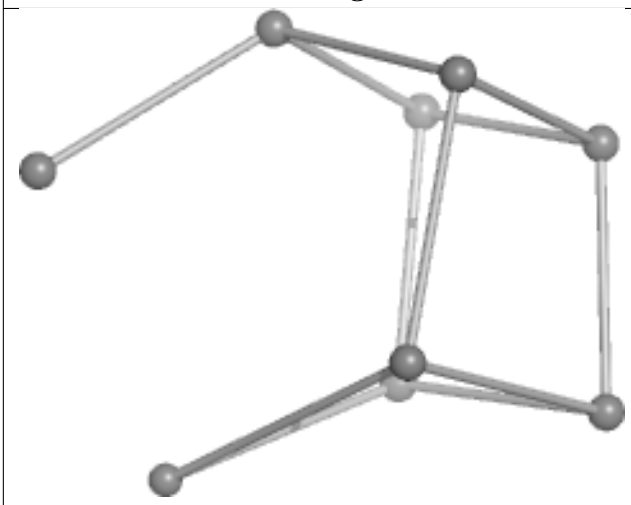
Ligand XCC B 705



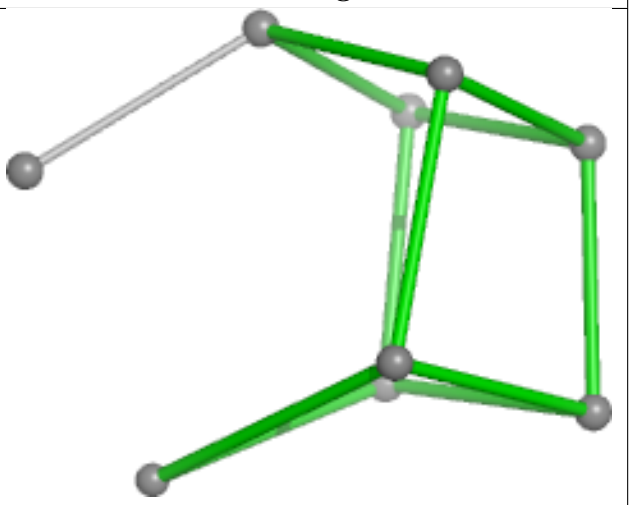
Bond lengths



Bond angles

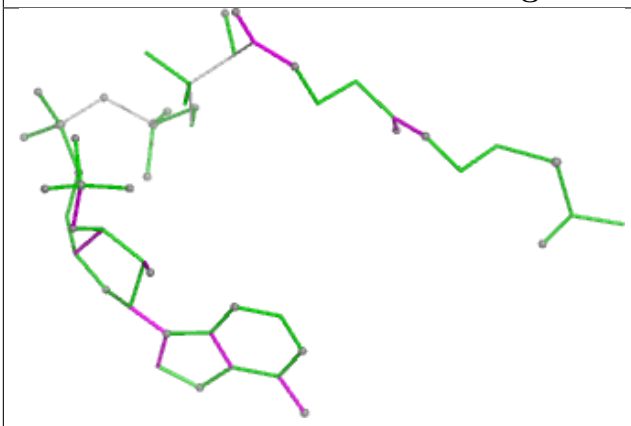


Torsions

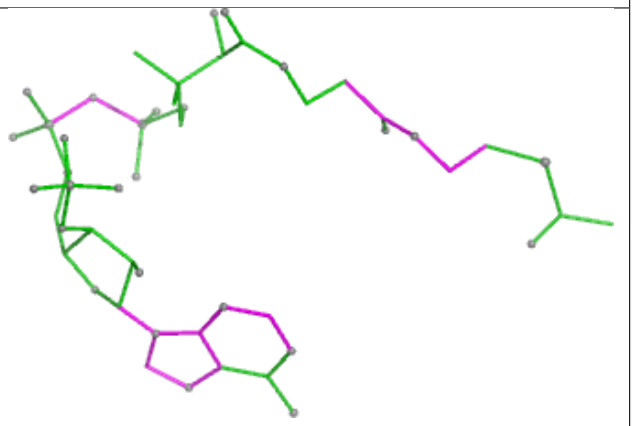


Rings

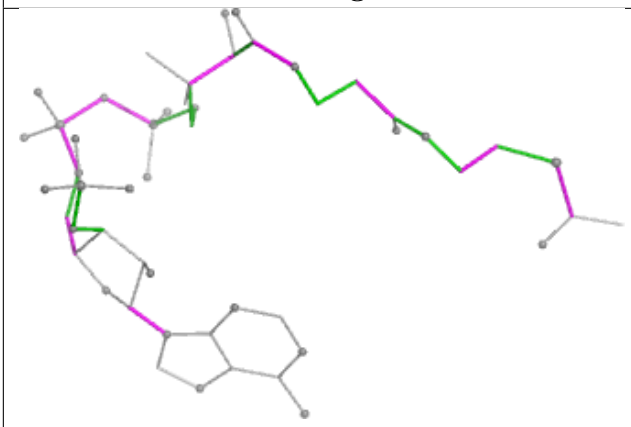
Ligand ACO D 805



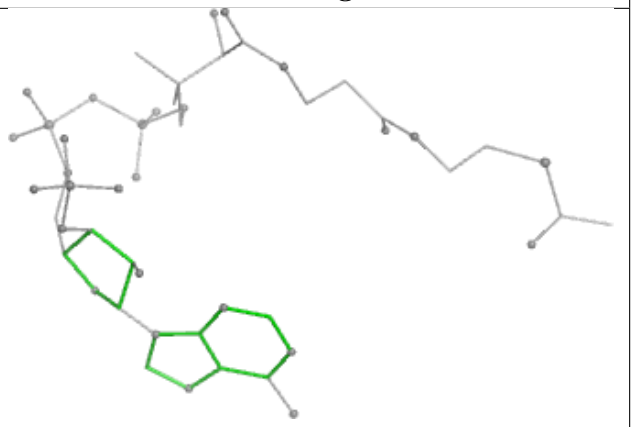
Bond lengths



Bond angles

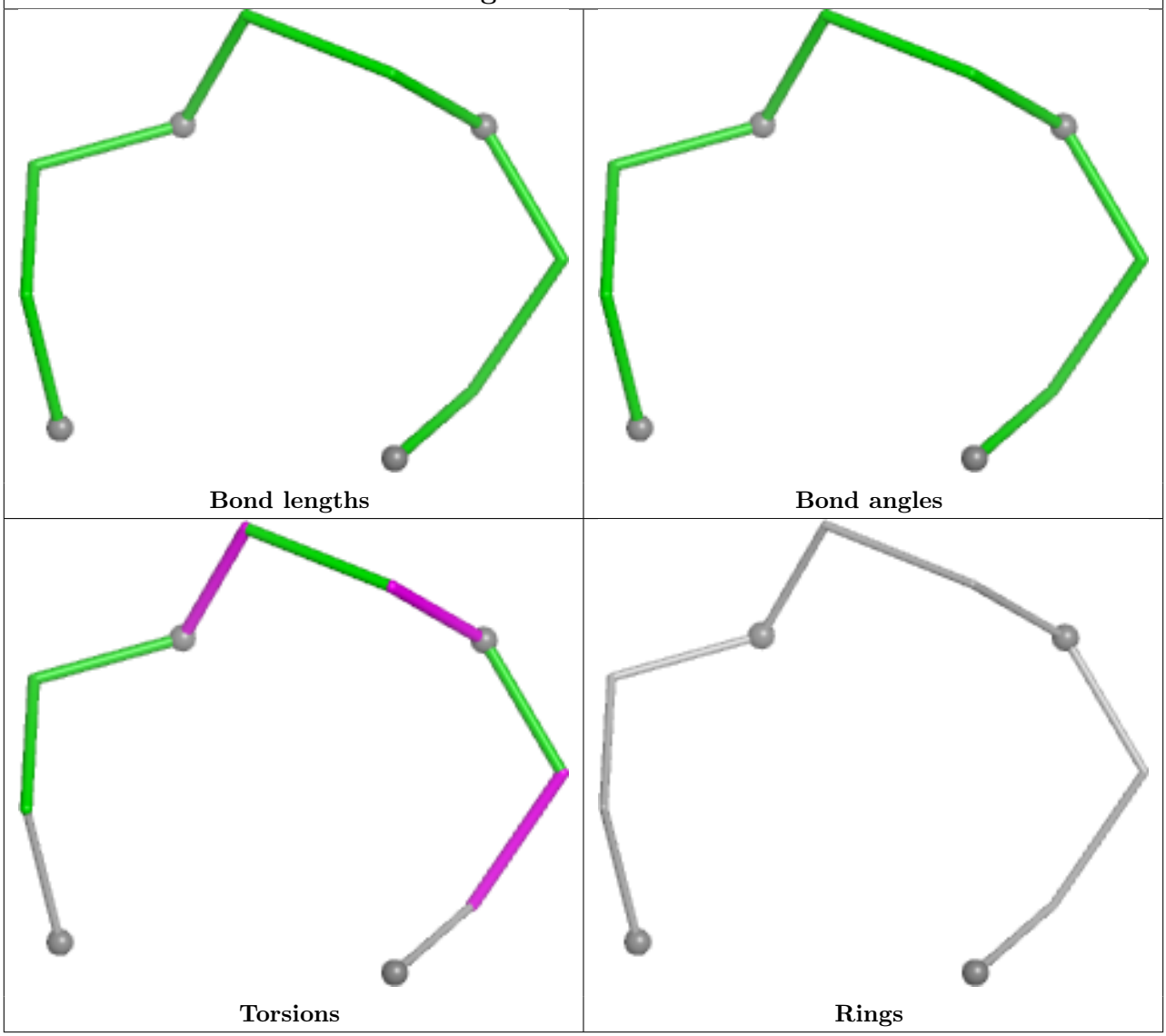


Torsions



Rings

Ligand PE4 A 805



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	694/708 (98%)	0.58	104 (14%) 5 5	42, 112, 196, 245	1 (0%)
1	D	708/708 (100%)	0.11	24 (3%) 48 40	63, 93, 159, 220	0
2	B	629/631 (99%)	-0.47	3 (0%) 87 83	39, 57, 82, 133	0
2	C	630/631 (99%)	-0.27	6 (0%) 79 74	39, 72, 104, 157	1 (0%)
All	All	2661/2678 (99%)	0.01	137 (5%) 33 27	39, 76, 169, 245	2 (0%)

The worst 5 of 137 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	371	THR	8.1
1	A	438	VAL	6.4
1	A	431	LEU	6.4
1	A	427	TYR	5.5
1	A	368	VAL	5.4

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

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
5	EDO	A	815	4/4	0.52	0.20	99,104,118,120	0
5	EDO	B	723	4/4	0.52	0.26	97,98,106,108	0
5	EDO	B	717	4/4	0.53	0.43	73,78,86,94	0
7	GOL	A	807	6/6	0.54	0.12	121,132,134,137	0
12	ACO	D	805	51/51	0.54	0.22	53,160,214,232	0
7	GOL	B	707	6/6	0.56	0.23	79,103,116,118	0
5	EDO	A	816	4/4	0.60	0.36	72,74,85,100	0
5	EDO	C	711	4/4	0.60	0.27	74,80,81,85	0
5	EDO	D	813	4/4	0.60	0.38	87,101,117,121	0
5	EDO	D	811	4/4	0.61	0.15	101,109,109,120	0
5	EDO	D	812	4/4	0.63	0.13	99,111,116,119	0
10	CL	B	712	1/1	0.65	0.16	122,122,122,122	0
5	EDO	B	724	4/4	0.67	0.26	89,96,100,112	0
9	CA	D	808	1/1	0.67	0.12	126,126,126,126	0
9	CA	A	810	1/1	0.68	0.11	136,136,136,136	0
5	EDO	B	721	4/4	0.68	0.30	99,102,105,122	0
10	CL	B	713	1/1	0.71	0.17	98,98,98,98	0
5	EDO	B	718	4/4	0.71	0.15	92,94,100,108	0
5	EDO	C	710	4/4	0.72	0.23	105,108,111,121	0
5	EDO	A	806	4/4	0.72	0.20	91,96,102,106	0
8	PEG	D	804	7/7	0.72	0.29	94,111,120,121	0
5	EDO	C	713	4/4	0.72	0.14	100,101,110,111	0
9	CA	D	809	1/1	0.74	0.15	163,163,163,163	0
9	CA	C	707	1/1	0.74	0.32	123,123,123,123	0
13	TRS	D	807	8/8	0.74	0.24	60,85,102,103	0
7	GOL	B	701	6/6	0.75	0.21	79,87,110,114	0
7	GOL	B	706	6/6	0.75	0.12	88,96,103,107	0
5	EDO	B	725	4/4	0.75	0.19	82,89,105,110	0
5	EDO	A	820	4/4	0.76	0.30	92,99,107,117	0
5	EDO	A	804	4/4	0.76	0.19	80,82,92,94	0
5	EDO	B	714	4/4	0.78	0.19	87,88,93,120	0
5	EDO	B	727	4/4	0.78	0.30	66,91,92,101	0
7	GOL	D	806	6/6	0.78	0.24	93,115,117,119	0
5	EDO	C	709	4/4	0.79	0.22	84,90,98,104	0
5	EDO	D	810	4/4	0.80	0.20	80,100,104,110	0
6	PE4	A	805	10/24	0.80	0.19	94,110,120,121	0
8	PEG	B	703	7/7	0.81	0.19	83,87,108,116	0
5	EDO	C	712	4/4	0.82	0.29	98,100,108,111	0
7	GOL	B	704	6/6	0.82	0.19	90,96,98,104	0
5	EDO	B	715	4/4	0.82	0.22	74,91,94,100	0
9	CA	A	809	1/1	0.83	0.14	125,125,125,125	0
9	CA	C	705	1/1	0.84	0.14	103,103,103,103	0
5	EDO	A	819	4/4	0.84	0.17	83,96,97,100	0

Continued on next page...

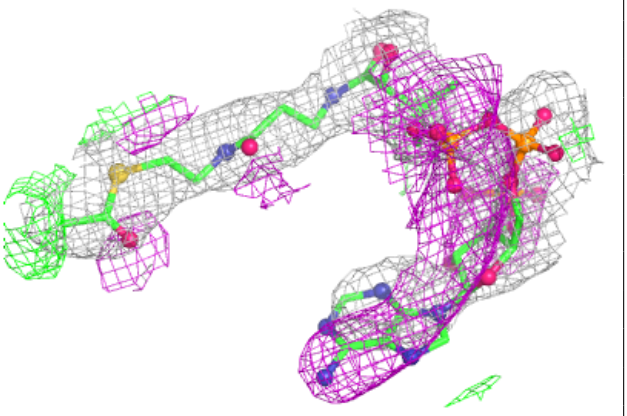
Continued from previous page...

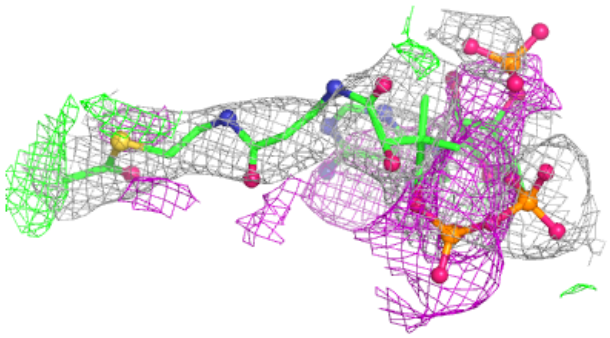
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	EDO	B	719	4/4	0.85	0.21	97,104,106,108	0
5	EDO	B	722	4/4	0.85	0.31	71,75,75,93	0
8	PEG	A	808	7/7	0.86	0.19	83,99,110,130	0
5	EDO	B	716	4/4	0.87	0.15	63,72,77,80	0
10	CL	A	814	1/1	0.87	0.22	105,105,105,105	0
5	EDO	C	703	4/4	0.87	0.26	91,93,94,96	0
5	EDO	B	728	4/4	0.88	0.28	72,73,83,89	0
9	CA	A	812	1/1	0.89	0.12	124,124,124,124	0
10	CL	A	813	1/1	0.89	0.15	94,94,94,94	0
5	EDO	A	817	4/4	0.89	0.28	88,91,108,124	0
10	CL	B	710	1/1	0.89	0.19	78,78,78,78	0
9	CA	B	708	1/1	0.91	0.13	122,122,122,122	0
5	EDO	C	714	4/4	0.91	0.18	81,82,83,83	0
10	CL	B	709	1/1	0.91	0.12	68,68,68,68	0
9	CA	C	706	1/1	0.91	0.11	123,123,123,123	0
9	CA	A	811	1/1	0.92	0.08	118,118,118,118	0
5	EDO	B	720	4/4	0.92	0.14	56,72,80,88	0
5	EDO	B	726	4/4	0.93	0.15	89,93,93,113	0
5	EDO	A	818	4/4	0.93	0.13	86,93,97,100	0
10	CL	B	711	1/1	0.94	0.12	96,96,96,96	0
11	XCC	C	704	9/9	0.94	0.08	55,81,126,128	0
10	CL	C	708	1/1	0.95	0.10	83,83,83,83	0
3	SF4	A	801	8/8	0.97	0.06	99,149,172,181	0
4	NI	A	803	1/1	0.98	0.06	183,183,183,183	0
3	SF4	B	702	8/8	0.99	0.04	53,55,72,74	0
11	XCC	B	705	9/9	0.99	0.04	45,60,75,88	0
4	NI	D	802	1/1	0.99	0.03	75,75,75,75	0
4	NI	D	803	1/1	0.99	0.02	53,53,53,53	0
4	NI	A	802	1/1	0.99	0.04	117,117,117,117	0
3	SF4	C	702	8/8	1.00	0.06	43,52,66,71	0
3	SF4	D	801	8/8	1.00	0.02	70,73,93,94	0
3	SF4	C	701	8/8	1.00	0.04	44,53,59,60	0

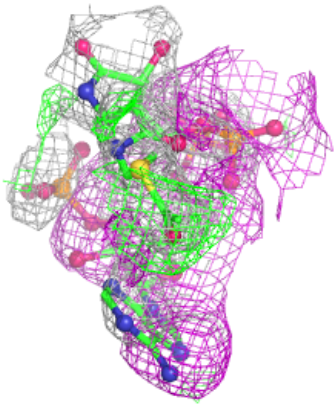
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around ACO D 805:

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

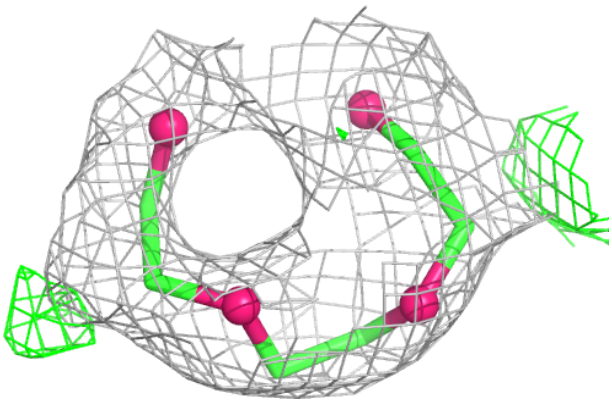


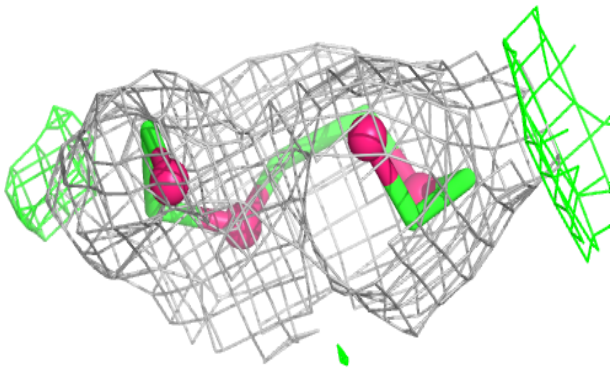


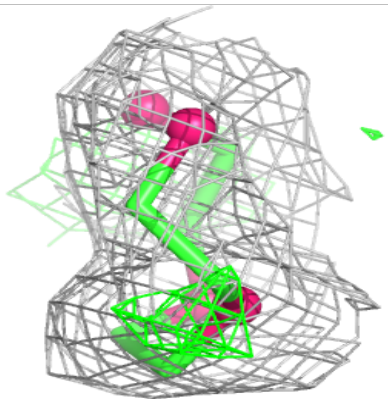


Electron density around PE4 A 805:

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

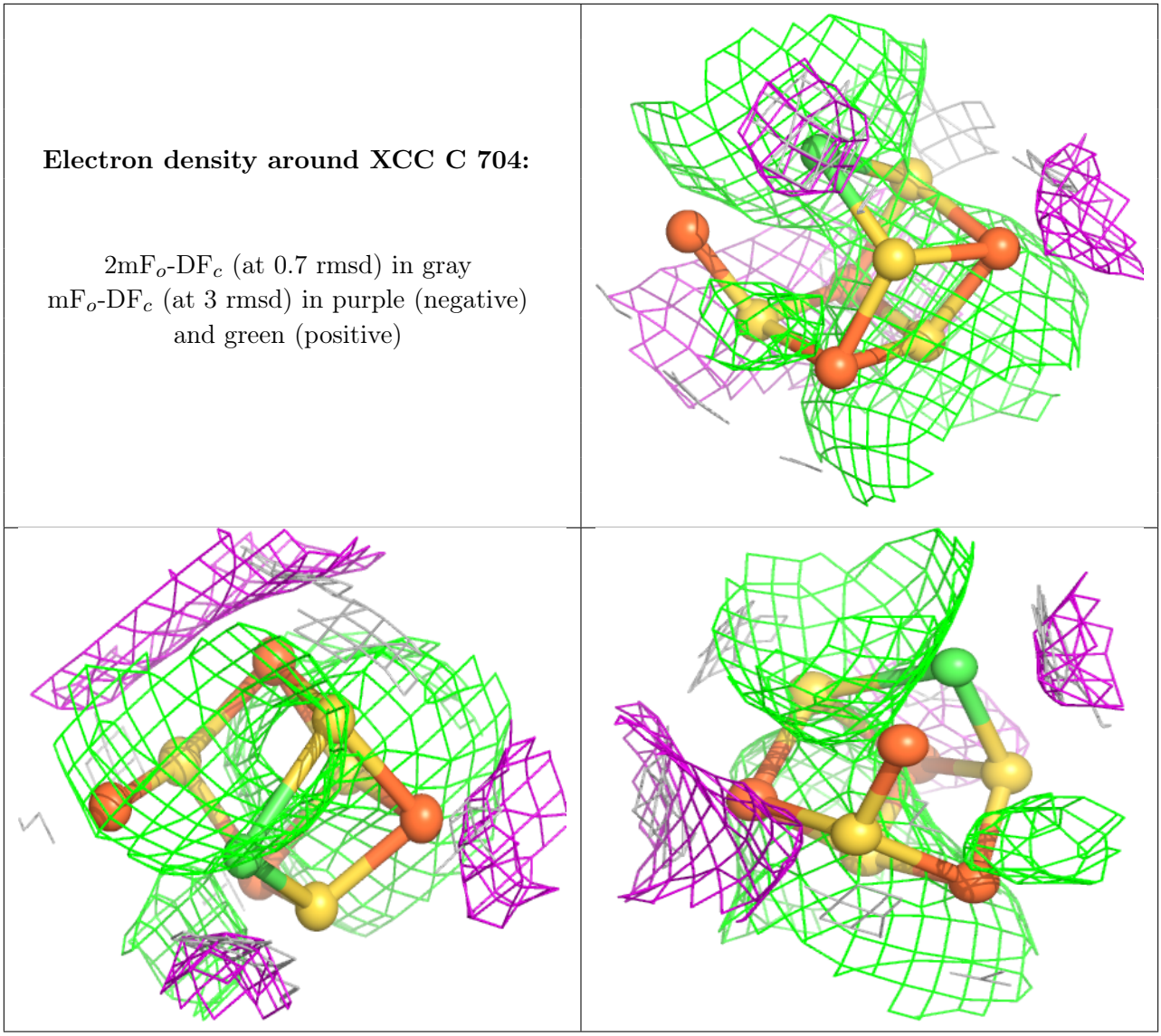






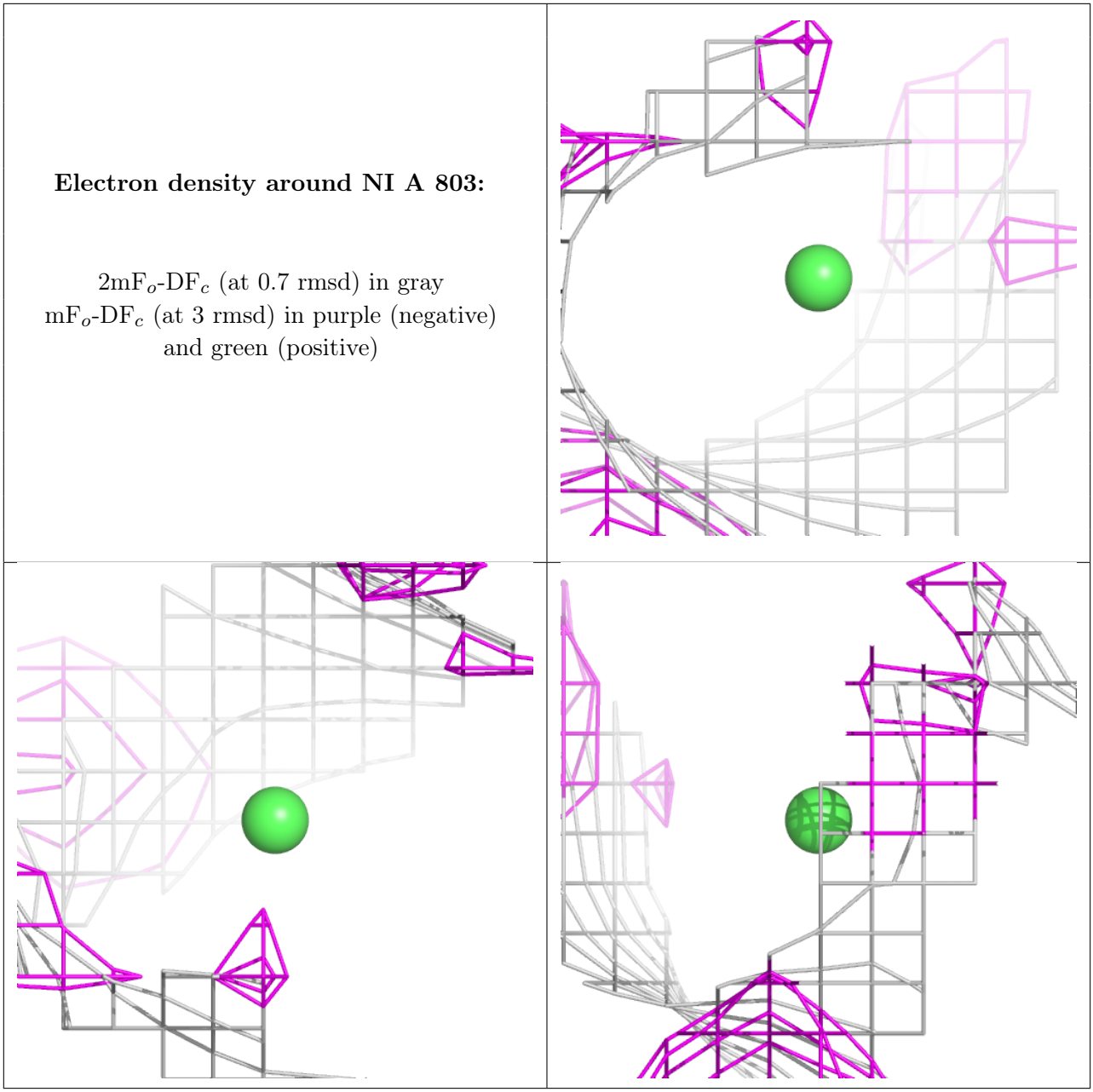
Electron density around XCC C 704:

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



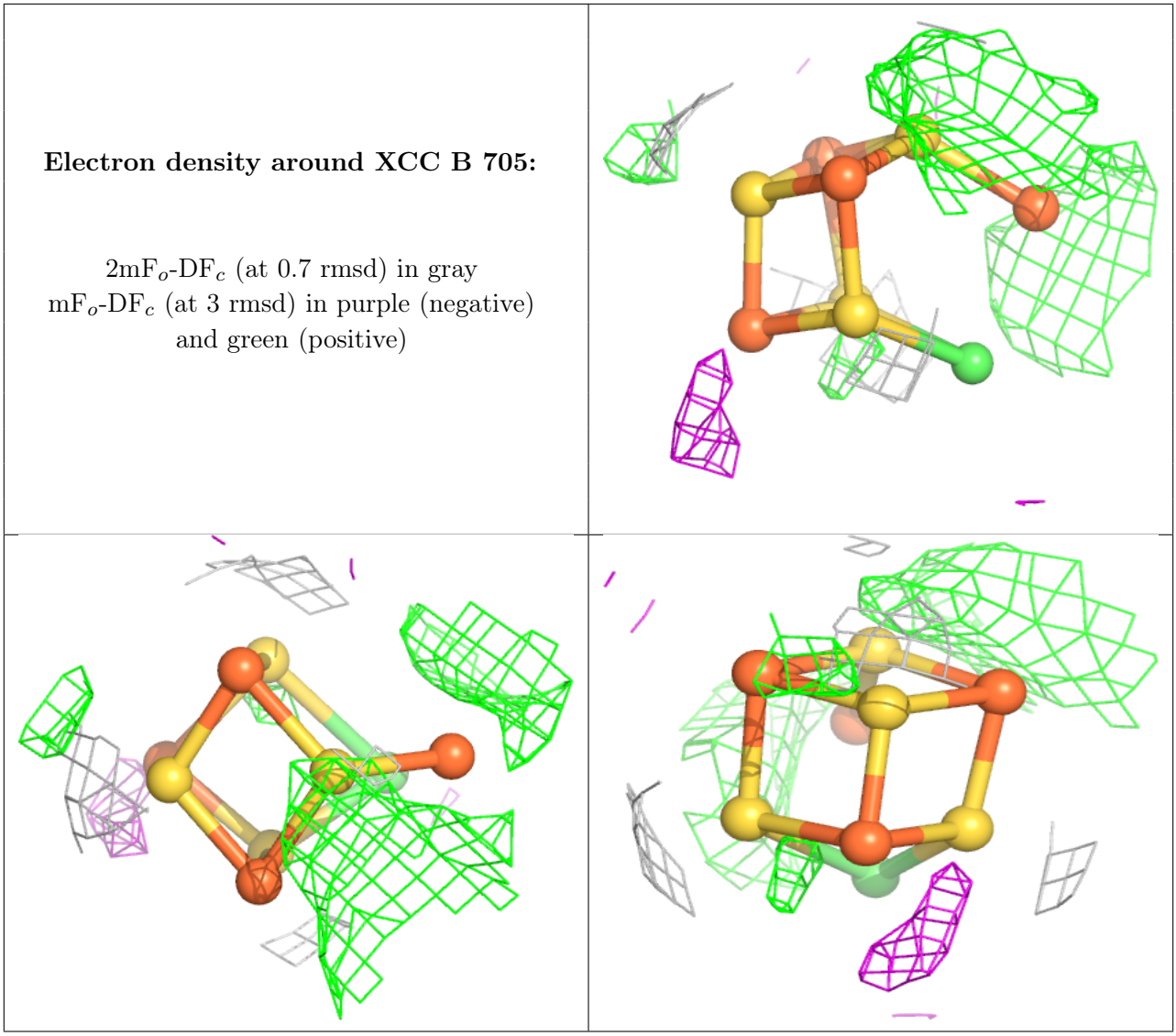
Electron density around NI A 803:

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



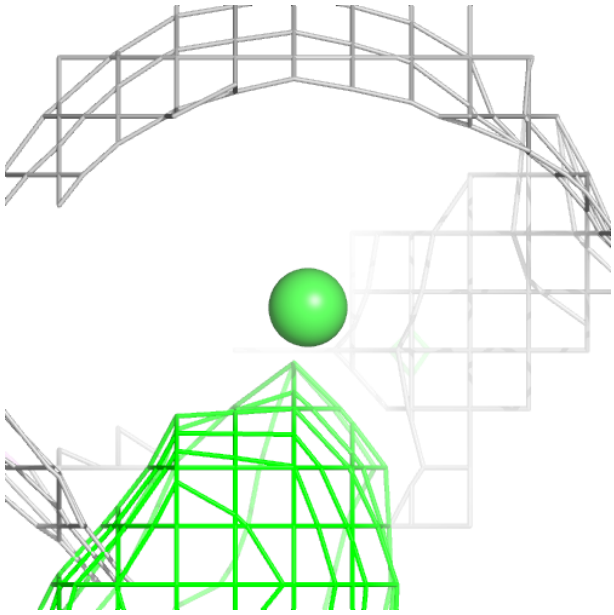
Electron density around XCC B 705:

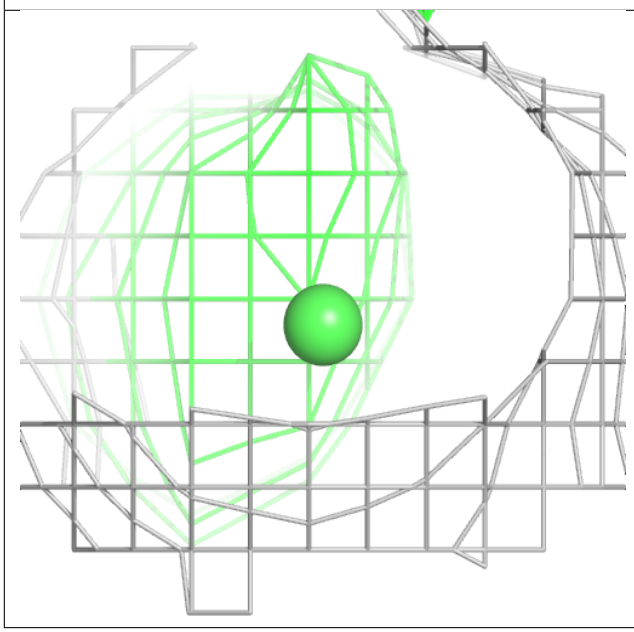
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

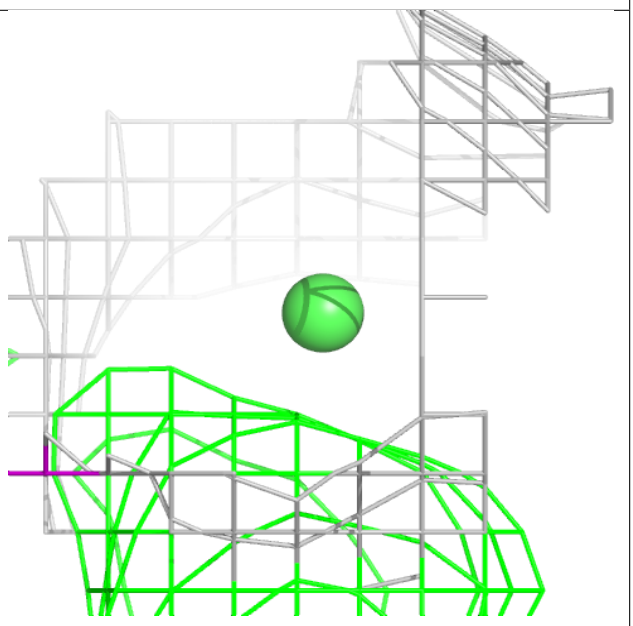


Electron density around NI D 802:

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

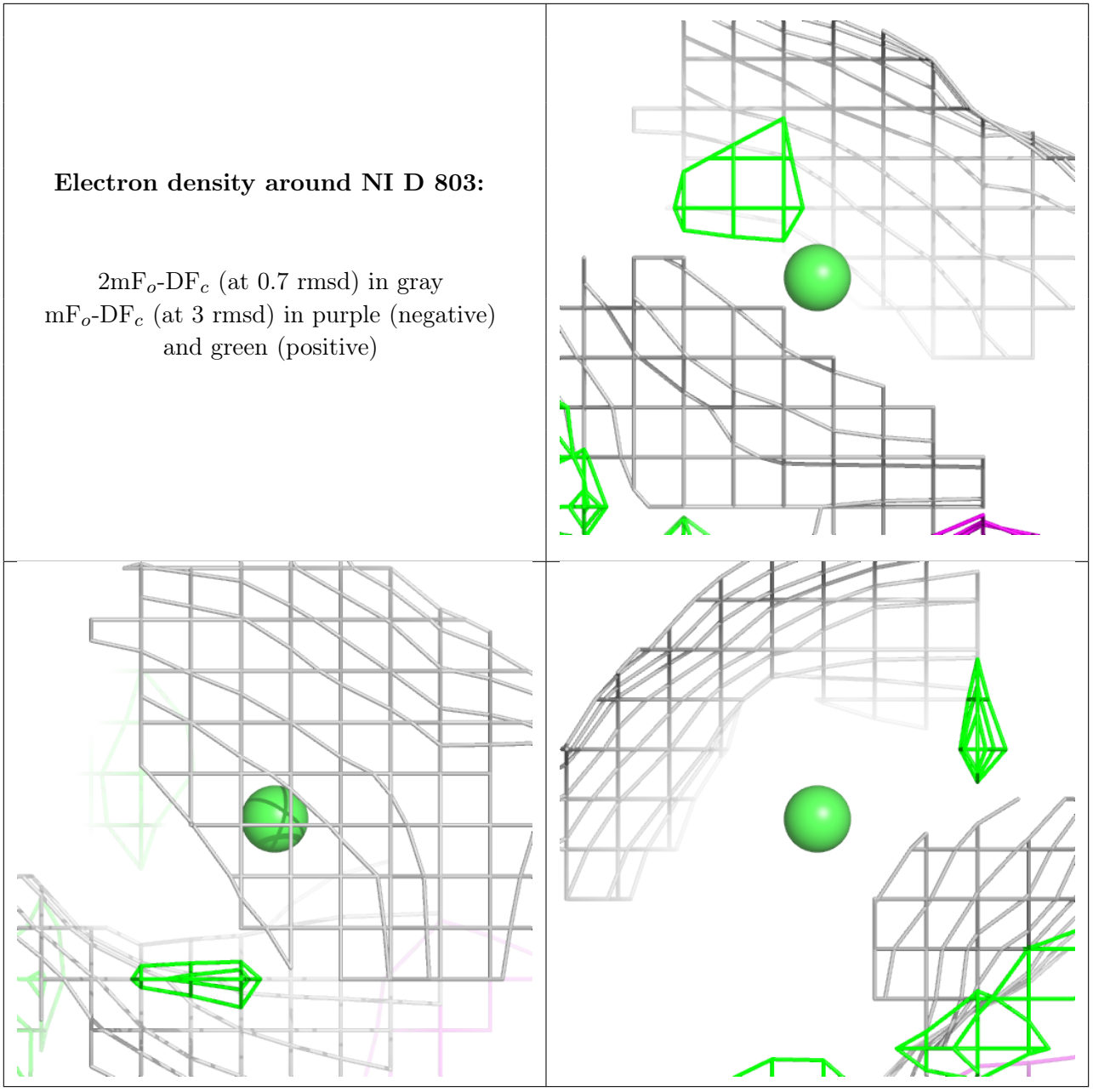






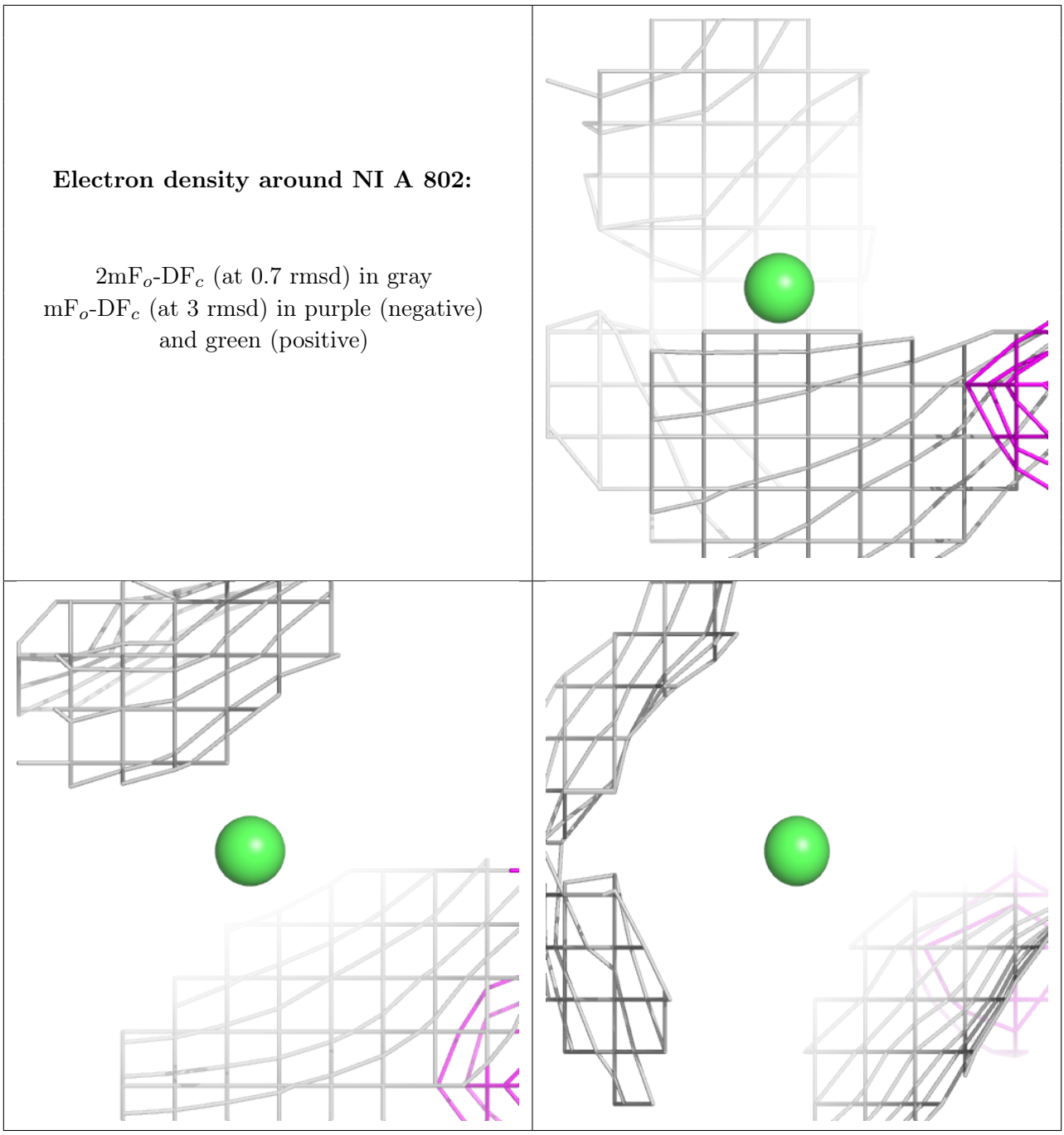
Electron density around NI D 803:

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



Electron density around NI A 802:

$2mF_o$-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.