



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 12, 2024 – 12:49 pm GMT

PDB ID : 8CH9
Title : Crystal structure of arsenite oxidase from *Alcaligenes faecalis* (Af Aio) bound to arsenic oxyanion
Authors : Engrola, F.; Correia, M.A.S.; Romao, M.J.; Santos-Silva, T.
Deposited on : 2023-02-07
Resolution : 1.43 Å (reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

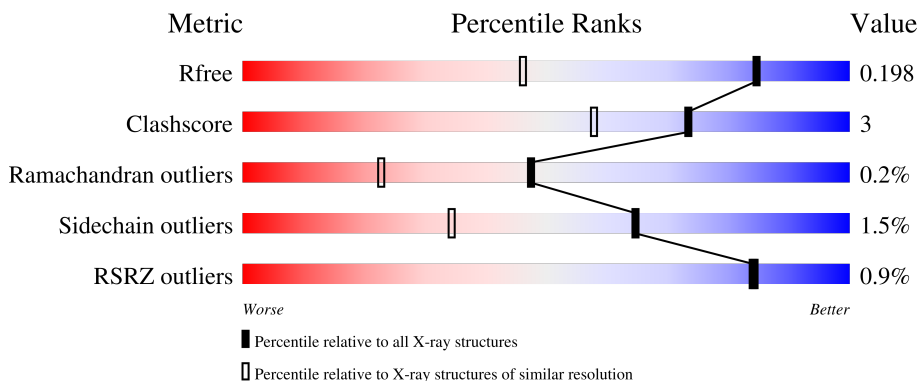
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.43 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2021 (1.46-1.42)
Clashscore	141614	2086 (1.46-1.42)
Ramachandran outliers	138981	2047 (1.46-1.42)
Sidechain outliers	138945	2047 (1.46-1.42)
RSRZ outliers	127900	1993 (1.46-1.42)

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

Mol	Chain	Length	Quality of chain
1	A	823	 91% 8% .
1	C	823	 92% 7% .
1	E	823	 90% 8% .
2	B	134	 92% 7% .
2	D	134	 92% 7% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	134	 % 94% 5% •
2	H	134	 % 91% 6% •••
3	G	823	 % 90% 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
9	GOL	E	909	-	-	X	-

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 35090 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	Total	C	N	O	S	5	18	0
			6626	4166	1176	1242	42			
1	C	823	Total	C	N	O	S	7	19	0
			6636	4170	1173	1250	43			
1	E	822	Total	C	N	O	S	12	24	0
			6647	4174	1176	1253	44			

- Molecule 2 is a protein called Arsenite oxidase subunit AioB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	134	Total	C	N	O	S	0	4	0
			1026	640	172	205	9			
2	D	133	Total	C	N	O	S	0	7	0
			1038	646	174	209	9			
2	F	134	Total	C	N	O	S	3	5	0
			1029	642	172	205	10			
2	H	133	Total	C	N	O	S	0	6	0
			1028	640	172	206	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	LEU	-	expression tag	UNP Q7SIF3
D	0	LEU	-	expression tag	UNP Q7SIF3
F	0	LEU	-	expression tag	UNP Q7SIF3
H	0	LEU	-	expression tag	UNP Q7SIF3

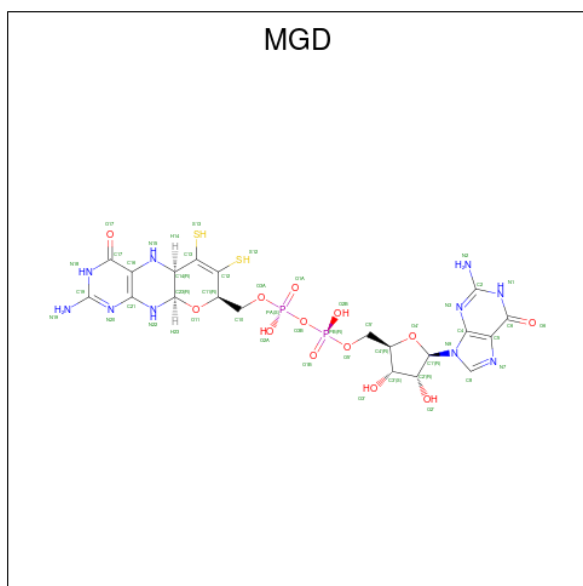
- Molecule 3 is a protein called Arsenite oxidase subunit AioA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	821	Total	C	N	O	S	0	20	0
			6619	4163	1173	1240	43			

There is a discrepancy between the modelled and reference sequences:

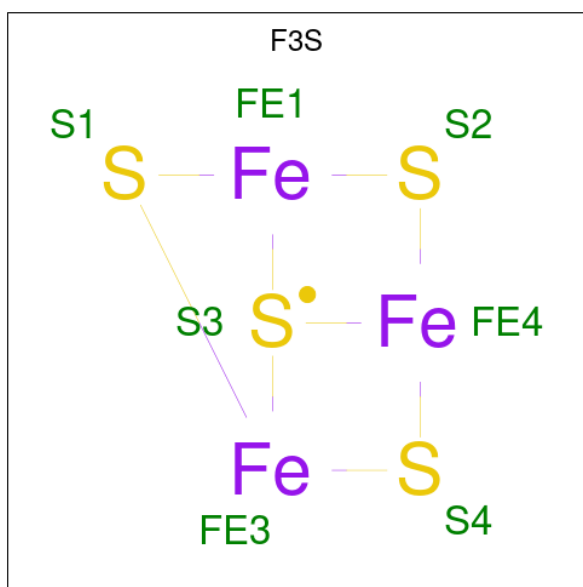
Chain	Residue	Modelled	Actual	Comment	Reference
G	280	ASP	GLU	conflict	UNP Q7SIF4

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



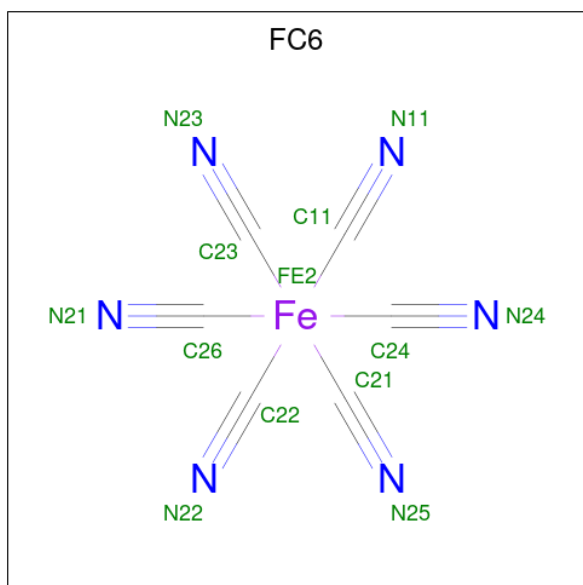
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
4	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2	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

- Molecule 5 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe₃S₄).



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

- Molecule 6 is HEXACYANOFERRATE(3-) (three-letter code: FC6) (formula: C_6FeN_6).

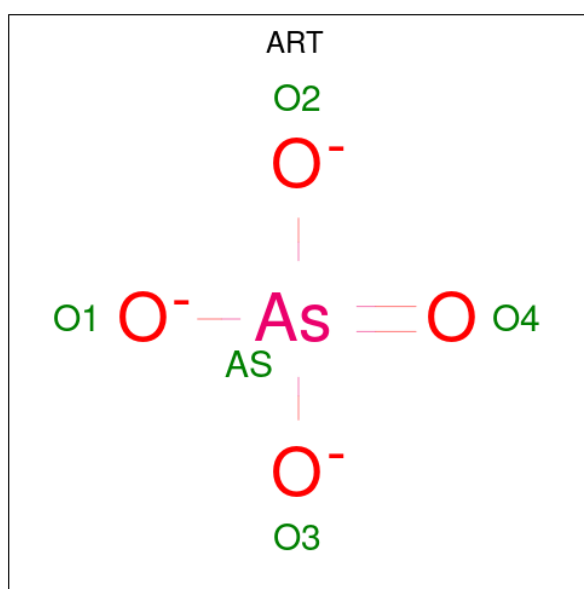


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	Fe	N	0	0
			13	6	1	6		
6	C	1	Total	C	Fe	N	0	0
			13	6	1	6		
6	E	1	Total	C	Fe	N	0	0
			13	6	1	6		
6	G	1	Total	C	Fe	N	0	0
			13	6	1	6		

- Molecule 7 is MOLYBDENUM(IV) ION (three-letter code: 4MO) (formula: Mo) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 8 is ARSENATE (three-letter code: ART) (formula: AsO₄) (labeled as "Ligand of Interest" by depositor).



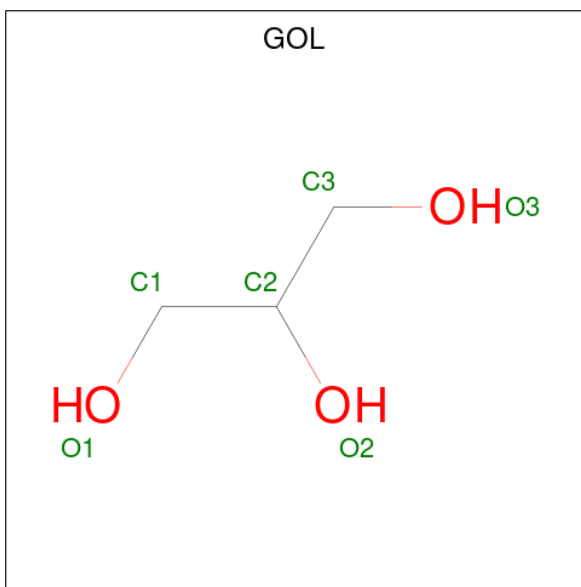
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	As	O	0	0
			5	1	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	As	O	0	0
			5	1	4		
8	E	1	Total	As	O	0	0
			5	1	4		
8	G	1	Total	As	O	0	0
			5	1	4		

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



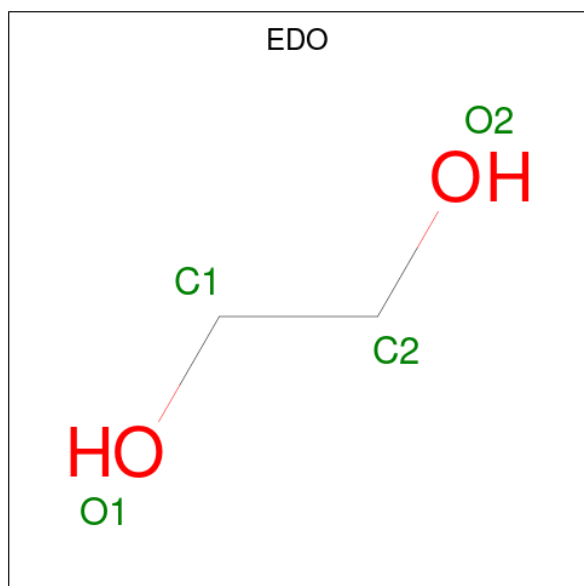
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			6	3	3		
9	A	1	Total	C	O	0	0
			6	3	3		
9	A	1	Total	C	O	0	0
			6	3	3		
9	B	1	Total	C	O	0	0
			6	3	3		
9	C	1	Total	C	O	0	0
			6	3	3		
9	C	1	Total	C	O	0	0
			6	3	3		
9	C	1	Total	C	O	0	0
			6	3	3		
9	E	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	E	1	Total	C	O	0	0
			6	3	3		
9	E	1	Total	C	O	0	0
			6	3	3		
9	E	1	Total	C	O	0	0
			6	3	3		
9	G	1	Total	C	O	0	0
			6	3	3		
9	G	1	Total	C	O	0	0
			6	3	3		
9	G	1	Total	C	O	0	0
			6	3	3		

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



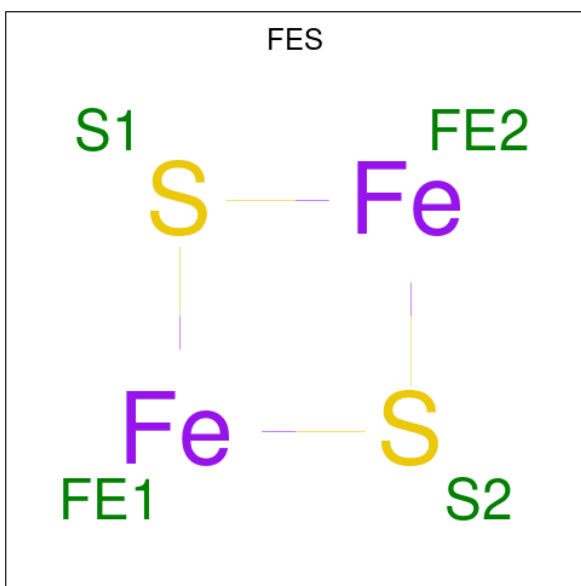
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	C	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

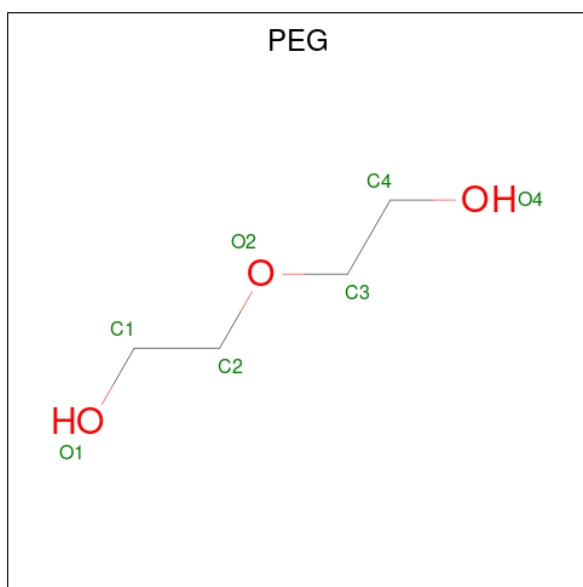
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	C	1	Total C O 4 2 2	0	0
10	C	1	Total C O 4 2 2	0	0
10	D	1	Total C O 4 2 2	0	0
10	E	1	Total C O 4 2 2	0	0
10	G	1	Total C O 4 2 2	0	0

- Molecule 11 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe₂S₂).



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

- Molecule 12 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	C	1	Total C O 7 4 3	0	0
12	H	1	Total C O 7 4 3	0	0

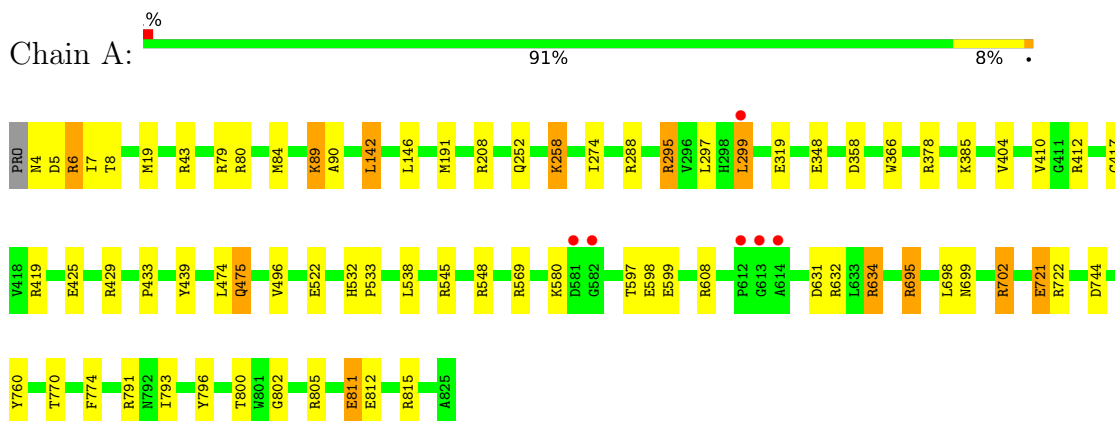
- Molecule 13 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	A	866	Total O 866 866	0	0
13	B	143	Total O 143 143	0	0
13	C	770	Total O 770 770	0	0
13	D	154	Total O 154 154	0	0
13	E	830	Total O 830 830	0	0
13	F	134	Total O 134 134	0	0
13	G	771	Total O 771 771	0	0
13	H	137	Total O 137 137	0	0

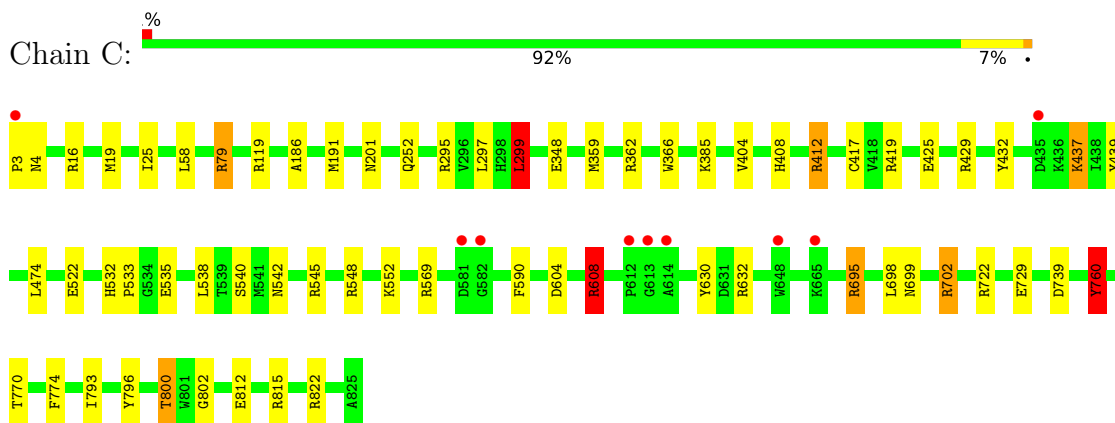
3 Residue-property plots [i](#)

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.

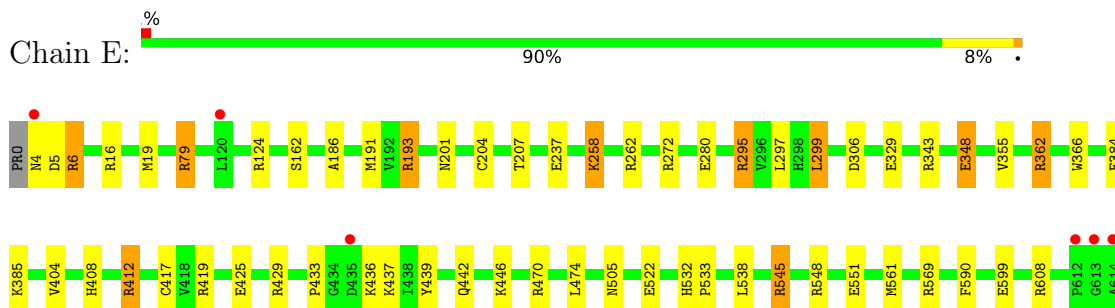
- Molecule 1: Arsenite oxidase subunit AioA



- Molecule 1: Arsenite oxidase subunit AioA

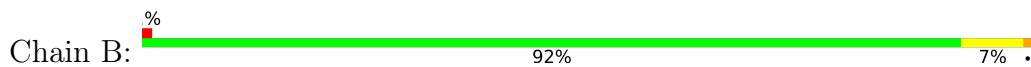


- Molecule 1: Arsenite oxidase subunit AioA

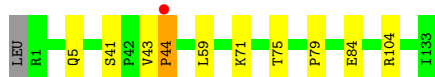
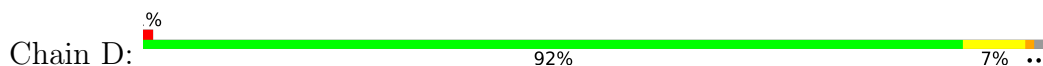




- Molecule 2: Arsenite oxidase subunit AioB



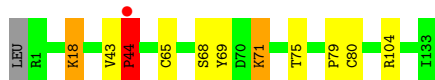
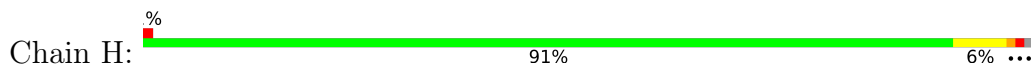
- Molecule 2: Arsenite oxidase subunit AioB



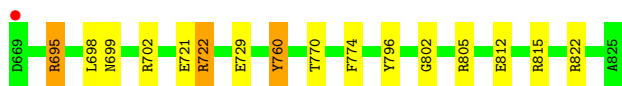
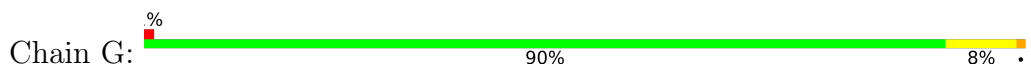
- Molecule 2: Arsenite oxidase subunit AioB



- Molecule 2: Arsenite oxidase subunit AioB



- Molecule 3: Arsenite oxidase subunit AioA



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	90.29Å 109.95Å 117.42Å 98.33° 89.97° 96.58°	Depositor
Resolution (Å)	65.47 – 1.43 65.39 – 1.43	Depositor EDS
% Data completeness (in resolution range)	73.0 (65.47-1.43) 73.0 (65.39-1.43)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.89 (at 1.43Å)	Xtrriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.167 , 0.195 0.172 , 0.198	Depositor DCC
R_{free} test set	29766 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	12.7	Xtrriage
Anisotropy	0.049	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 43.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	35090	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: MGD, 4MO, F3S, ART, EDO, FES, PEG, GOL, FC6

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	0.77	3/6790 (0.0%)	1.10	40/9199 (0.4%)
1	C	0.72	0/6803	1.10	36/9216 (0.4%)
1	E	0.75	6/6813 (0.1%)	1.11	37/9228 (0.4%)
2	B	0.67	0/1047	1.01	1/1426 (0.1%)
2	D	0.70	0/1060	1.07	3/1444 (0.2%)
2	F	0.69	0/1053	1.02	1/1435 (0.1%)
2	H	0.64	0/1052	1.02	2/1432 (0.1%)
3	G	0.73	4/6786 (0.1%)	1.11	35/9196 (0.4%)
All	All	0.73	13/31404 (0.0%)	1.09	155/42576 (0.4%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
1	C	0	6
1	E	0	6
3	G	0	5
All	All	0	23

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	348	GLU	C-O	-8.53	1.07	1.23
3	G	280	ASP	CG-OD1	7.40	1.42	1.25
1	A	348	GLU	CD-OE1	6.55	1.32	1.25
1	E	237	GLU	CD-OE2	6.33	1.32	1.25
3	G	272	ARG	NE-CZ	6.00	1.40	1.33

The worst 5 of 155 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	295	ARG	NE-CZ-NH1	13.45	127.03	120.30
3	G	295	ARG	NE-CZ-NH1	12.98	126.79	120.30
1	C	295	ARG	NE-CZ-NH1	11.58	126.09	120.30
2	H	44	PRO	N-CA-CB	-11.38	89.64	103.30
3	G	280	ASP	CB-CG-OD2	-10.69	108.68	118.30

There are no chirality outliers.

5 of 23 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	295	ARG	Mainchain
1	A	433	PRO	Peptide
1	A	545	ARG	Sidechain
1	A	695	ARG	Sidechain
1	A	702	ARG	Sidechain

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	6626	0	6406	45	0
1	C	6636	0	6405	39	0
1	E	6647	0	6409	36	0
2	B	1026	0	1000	9	0
2	D	1038	0	1005	7	0
2	F	1029	0	1006	4	0
2	H	1028	0	999	8	0
3	G	6619	0	6407	49	0
4	A	94	0	44	1	0
4	C	94	0	44	1	0
4	E	94	0	44	2	0
4	G	94	0	44	1	0
5	A	7	0	0	0	0
5	C	7	0	0	0	0
5	E	7	0	0	0	0
5	G	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	A	13	0	0	1	0
6	C	13	0	0	1	0
6	E	13	0	0	1	0
6	G	13	0	0	0	0
7	A	1	0	0	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
7	G	1	0	0	0	0
8	A	5	0	0	0	0
8	C	5	0	0	0	0
8	E	5	0	0	1	0
8	G	5	0	0	0	0
9	A	18	0	24	4	0
9	B	6	0	8	1	0
9	C	18	0	24	3	0
9	E	24	0	32	8	0
9	G	24	0	32	1	0
10	A	12	0	18	2	0
10	C	12	0	18	2	0
10	D	4	0	6	1	0
10	E	4	0	6	0	0
10	G	4	0	6	2	0
11	B	4	0	0	0	0
11	D	4	0	0	0	0
11	F	4	0	0	0	0
11	H	4	0	0	0	0
12	C	7	0	10	2	0
12	H	7	0	10	3	0
13	A	866	0	0	11	0
13	B	143	0	0	5	0
13	C	770	0	0	11	0
13	D	154	0	0	3	0
13	E	830	0	0	12	0
13	F	134	0	0	0	0
13	G	771	0	0	15	0
13	H	137	0	0	2	0
All	All	35090	0	30007	200	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 3.

The worst 5 of 200 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:425[B]:GLU:OE1	9:C:906:GOL:O1	1.70	1.08
1:A:4:ASN:O	1:A:6:ARG:N	1.88	1.07
1:A:425[B]:GLU:OE1	9:A:912:GOL:O3	1.77	1.03
1:A:569[B]:ARG:NH1	13:A:1001:HOH:O	1.95	0.99
3:G:273[A]:ILE:HD12	3:G:275:PHE:CE1	2.02	0.95

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	839/823 (102%)	811 (97%)	26 (3%)	2 (0%)	47	23
1	C	841/823 (102%)	811 (96%)	29 (3%)	1 (0%)	51	24
1	E	843/823 (102%)	817 (97%)	24 (3%)	2 (0%)	47	23
2	B	136/134 (102%)	129 (95%)	6 (4%)	1 (1%)	22	4
2	D	138/134 (103%)	129 (94%)	9 (6%)	0	100	100
2	F	137/134 (102%)	130 (95%)	7 (5%)	0	100	100
2	H	137/134 (102%)	130 (95%)	6 (4%)	1 (1%)	22	4
3	G	840/823 (102%)	814 (97%)	26 (3%)	0	100	100
All	All	3911/3828 (102%)	3771 (96%)	133 (3%)	7 (0%)	47	23

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	ASP
2	H	44	PRO
2	B	44	PRO
1	E	5	ASP
1	C	793	ILE

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	694/676 (103%)	684 (99%)	10 (1%)	67	37
1	C	696/676 (103%)	682 (98%)	14 (2%)	55	21
1	E	698/676 (103%)	689 (99%)	9 (1%)	69	39
2	B	116/112 (104%)	114 (98%)	2 (2%)	60	28
2	D	118/112 (105%)	115 (98%)	3 (2%)	47	13
2	F	117/112 (104%)	113 (97%)	4 (3%)	37	6
2	H	117/112 (104%)	114 (97%)	3 (3%)	46	12
3	G	695/676 (103%)	682 (98%)	13 (2%)	57	22
All	All	3251/3152 (103%)	3193 (98%)	58 (2%)	65	25

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	104	ARG
2	H	18	LYS
1	E	760	TYR
3	G	774	PHE
3	G	439	TYR

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

Mol	Chain	Res	Type
1	A	688	GLN
3	G	482	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 54 ligands modelled in this entry, 4 are monoatomic - leaving 50 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
4	MGD	A	902	7	41,52,52	1.15	3 (7%)	40,81,81	1.11	3 (7%)
9	GOL	C	907	-	5,5,5	0.22	0	5,5,5	0.68	0
4	MGD	E	901	7	41,52,52	1.69	3 (7%)	40,81,81	1.15	5 (12%)
5	F3S	A	903	1	0,9,9	-	-	-	-	-
9	GOL	E	908	-	5,5,5	0.35	0	5,5,5	1.02	0
8	ART	A	906	-	0,4,4	-	-	0,6,6	-	-
4	MGD	C	901	7	41,52,52	1.11	3 (7%)	40,81,81	1.19	3 (7%)
6	FC6	E	905	-	12,12,12	3.41	7 (58%)	-	-	-
9	GOL	A	912	-	5,5,5	0.24	0	5,5,5	0.92	0
9	GOL	C	908	-	5,5,5	0.51	0	5,5,5	1.42	1 (20%)
9	GOL	E	907	-	5,5,5	0.14	0	5,5,5	0.39	0
11	FES	B	201	2	0,4,4	-	-	-	-	-
11	FES	F	201	2	0,4,4	-	-	-	-	-
10	EDO	A	911	-	3,3,3	1.23	0	2,2,2	1.10	0
6	FC6	A	904	-	12,12,12	3.94	8 (66%)	-	-	-
12	PEG	C	911	-	6,6,6	0.17	0	5,5,5	0.29	0
4	MGD	E	902	7	41,52,52	1.19	3 (7%)	40,81,81	1.17	5 (12%)
6	FC6	C	905	-	12,12,12	3.68	8 (66%)	-	-	-
10	EDO	C	910	-	3,3,3	0.17	0	2,2,2	0.06	0
10	EDO	C	912	-	3,3,3	0.46	0	2,2,2	0.58	0
9	GOL	G	2409	-	5,5,5	0.31	0	5,5,5	1.10	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
9	GOL	E	910	-	5,5,5	0.21	0	5,5,5	0.68	0
4	MGD	G	2403	7	41,52,52	1.13	4 (9%)	40,81,81	1.30	5 (12%)
11	FES	H	201	2	0,4,4	-	-	-	-	-
9	GOL	B	202	-	5,5,5	0.15	0	5,5,5	0.35	0
10	EDO	A	909	-	3,3,3	0.18	0	2,2,2	0.77	0
12	PEG	H	202	-	6,6,6	0.68	0	5,5,5	0.30	0
8	ART	G	2407	-	0,4,4	-	-	0,6,6	-	-
11	FES	D	201	2	0,4,4	-	-	-	-	-
4	MGD	C	902	7	41,52,52	1.34	5 (12%)	40,81,81	1.38	5 (12%)
5	F3S	C	903	1	0,9,9	-	-	-	-	-
10	EDO	E	911	-	3,3,3	0.93	0	2,2,2	0.49	0
10	EDO	G	2401	-	3,3,3	0.30	0	2,2,2	0.30	0
9	GOL	G	2408	-	5,5,5	0.64	0	5,5,5	1.28	1 (20%)
9	GOL	E	909	-	5,5,5	0.17	0	5,5,5	0.65	0
9	GOL	G	2410	-	5,5,5	0.27	0	5,5,5	0.48	0
4	MGD	G	2402	7	41,52,52	1.55	6 (14%)	40,81,81	1.22	3 (7%)
5	F3S	G	2404	3	0,9,9	-	-	-	-	-
9	GOL	C	906	-	5,5,5	0.43	0	5,5,5	0.75	0
4	MGD	A	901	7	41,52,52	1.38	4 (9%)	40,81,81	1.14	2 (5%)
8	ART	C	904	-	0,4,4	-	-	0,6,6	-	-
9	GOL	A	908	-	5,5,5	0.28	0	5,5,5	0.55	0
5	F3S	E	903	1	0,9,9	-	-	-	-	-
6	FC6	G	2405	-	12,12,12	3.12	7 (58%)	-	-	-
10	EDO	A	910	-	3,3,3	0.37	0	2,2,2	0.63	0
9	GOL	A	907	-	5,5,5	0.24	0	5,5,5	0.40	0
10	EDO	C	913	-	3,3,3	0.59	0	2,2,2	0.40	0
8	ART	E	904	-	0,4,4	-	-	0,6,6	-	-
10	EDO	D	202	-	3,3,3	0.07	0	2,2,2	0.03	0
9	GOL	G	2411	-	5,5,5	0.52	0	5,5,5	0.94	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
4	MGD	A	902	7	-	4/18/66/66	0/6/6/6
9	GOL	C	907	-	-	0/4/4/4	-
4	MGD	E	901	7	-	4/18/66/66	0/6/6/6
9	GOL	E	908	-	-	2/4/4/4	-
5	F3S	A	903	1	-	-	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MGD	C	901	7	-	3/18/66/66	0/6/6/6
9	GOL	A	912	-	-	0/4/4/4	-
11	FES	B	201	2	-	-	0/1/1/1
9	GOL	C	908	-	-	2/4/4/4	-
9	GOL	E	907	-	-	0/4/4/4	-
11	FES	F	201	2	-	-	0/1/1/1
10	EDO	A	911	-	-	1/1/1/1	-
12	PEG	C	911	-	-	2/4/4/4	-
4	MGD	E	902	7	-	5/18/66/66	0/6/6/6
10	EDO	C	910	-	-	1/1/1/1	-
10	EDO	C	912	-	-	1/1/1/1	-
9	GOL	G	2409	-	-	2/4/4/4	-
9	GOL	E	910	-	-	2/4/4/4	-
4	MGD	G	2403	7	-	5/18/66/66	0/6/6/6
11	FES	H	201	2	-	-	0/1/1/1
9	GOL	B	202	-	-	1/4/4/4	-
10	EDO	A	909	-	-	0/1/1/1	-
12	PEG	H	202	-	-	4/4/4/4	-
11	FES	D	201	2	-	-	0/1/1/1
4	MGD	C	902	7	-	4/18/66/66	0/6/6/6
10	EDO	E	911	-	-	1/1/1/1	-
10	EDO	G	2401	-	-	0/1/1/1	-
5	F3S	C	903	1	-	-	0/3/3/3
9	GOL	G	2408	-	-	0/4/4/4	-
9	GOL	E	909	-	-	3/4/4/4	-
9	GOL	G	2410	-	-	2/4/4/4	-
4	MGD	G	2402	7	-	4/18/66/66	0/6/6/6
5	F3S	G	2404	3	-	-	0/3/3/3
9	GOL	C	906	-	-	3/4/4/4	-
4	MGD	A	901	7	-	4/18/66/66	0/6/6/6
9	GOL	A	908	-	-	0/4/4/4	-
10	EDO	A	910	-	-	1/1/1/1	-
5	F3S	E	903	1	-	-	0/3/3/3
9	GOL	A	907	-	-	0/4/4/4	-
10	EDO	C	913	-	-	1/1/1/1	-
10	EDO	D	202	-	-	1/1/1/1	-
9	GOL	G	2411	-	-	0/4/4/4	-

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	904	FC6	C23-FE2	-8.56	1.68	1.93
6	C	905	FC6	C21-FE2	-8.46	1.68	1.93
4	E	901	MGD	C23-C14	-8.13	1.47	1.53
6	E	905	FC6	C21-FE2	-7.14	1.72	1.93
4	G	2402	MGD	C23-C14	-6.39	1.48	1.53

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	902	MGD	O11-C23-C14	4.14	111.73	108.96
4	C	901	MGD	O11-C23-N22	-3.77	104.69	108.57
4	G	2402	MGD	O4'-C1'-C2'	-3.71	101.50	106.93
4	A	901	MGD	O11-C23-C14	3.65	111.40	108.96
4	A	902	MGD	O11-C23-C14	3.39	111.23	108.96

There are no chirality outliers.

5 of 63 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	901	MGD	PA-O3B-PB-O5'
4	A	901	MGD	C5'-O5'-PB-O1B
4	A	901	MGD	C5'-O5'-PB-O3B
4	A	902	MGD	PA-O3B-PB-O5'
4	A	902	MGD	C5'-O5'-PB-O1B

There are no ring outliers.

23 monomers are involved in 37 short contacts:

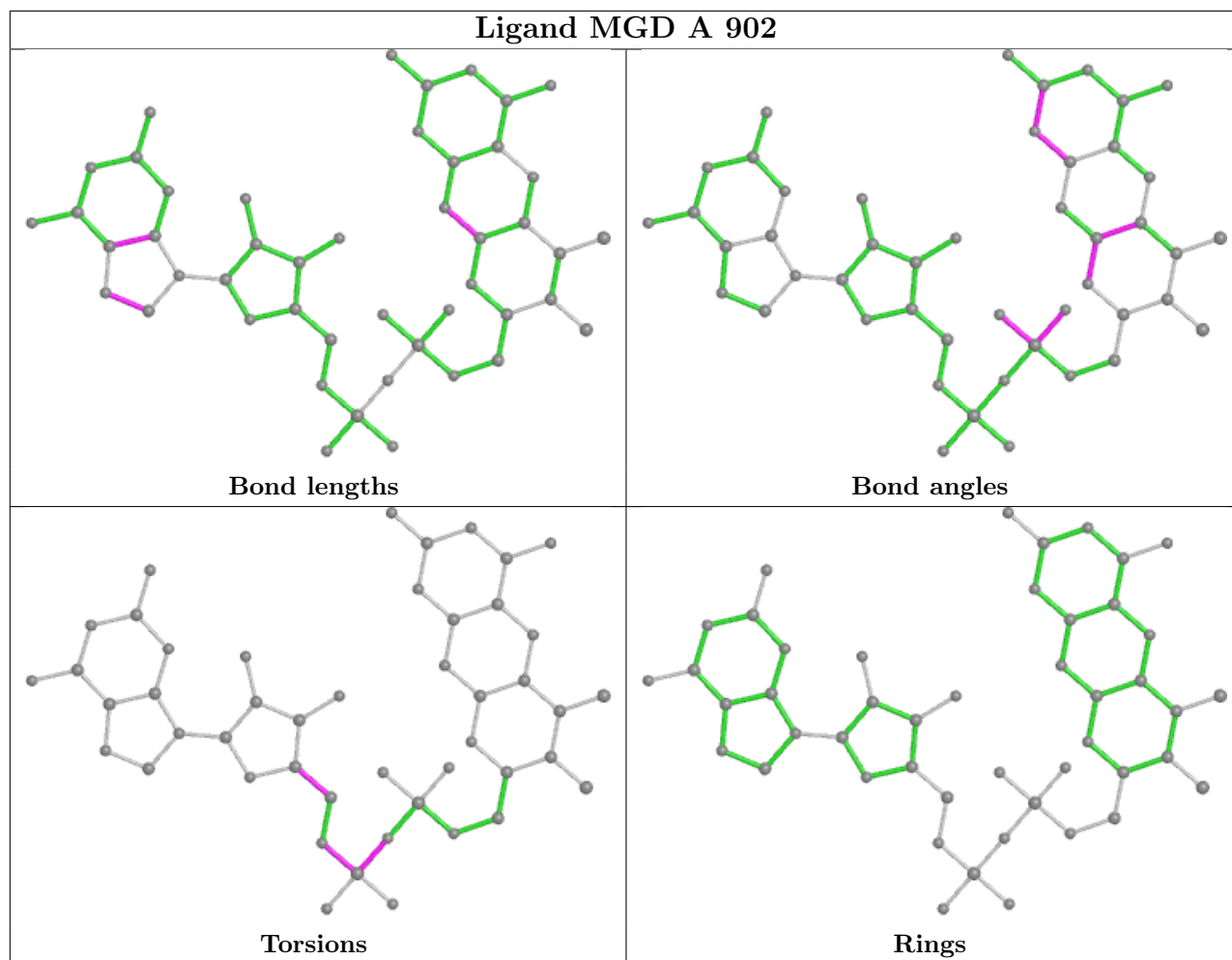
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	902	MGD	1	0
6	E	905	FC6	1	0
9	A	912	GOL	3	0
9	E	907	GOL	3	0
10	A	911	EDO	1	0
6	A	904	FC6	1	0
12	C	911	PEG	2	0
4	E	902	MGD	2	0
6	C	905	FC6	1	0
9	E	910	GOL	1	0
4	G	2403	MGD	1	0
9	B	202	GOL	1	0
12	H	202	PEG	3	0

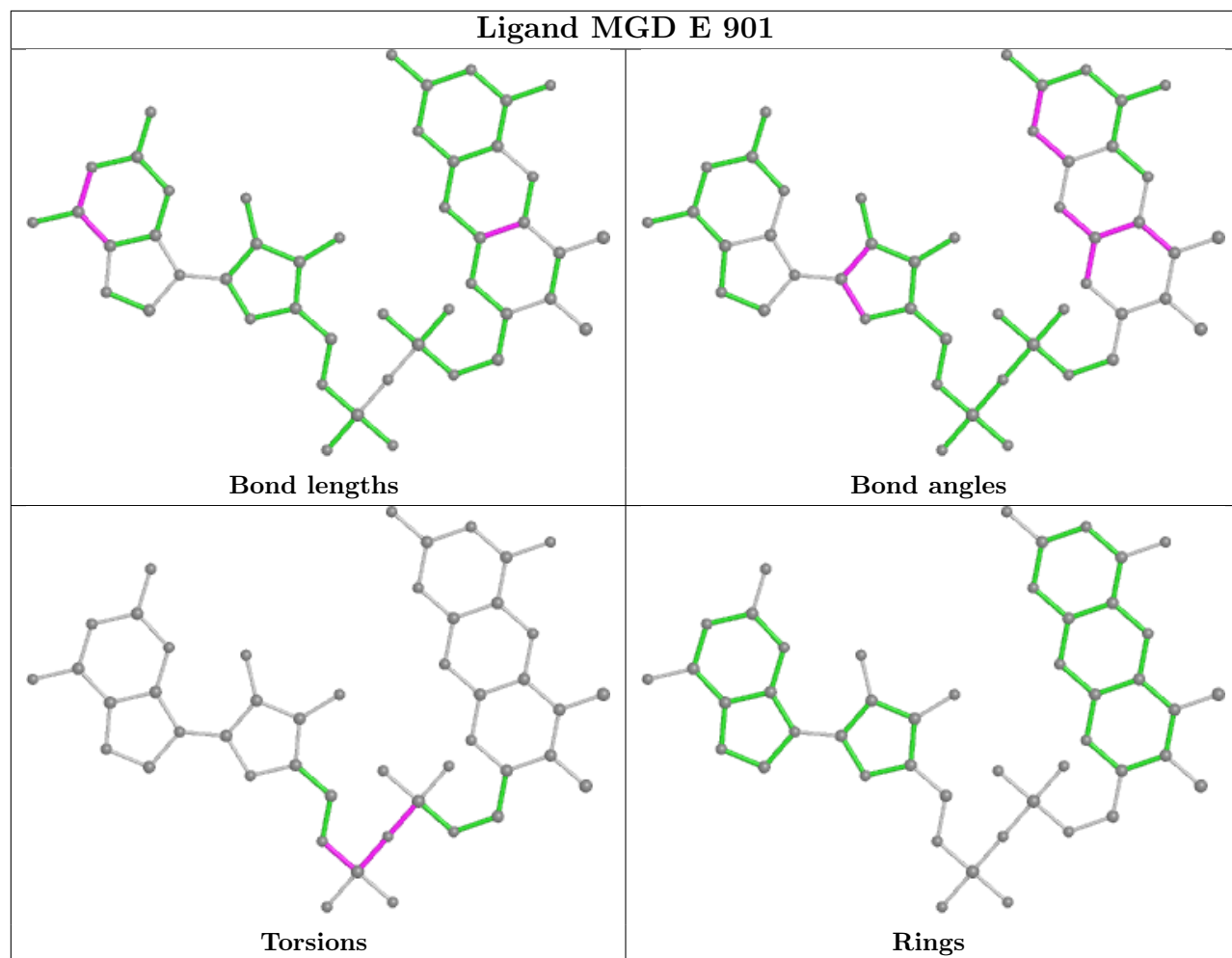
Continued on next page...

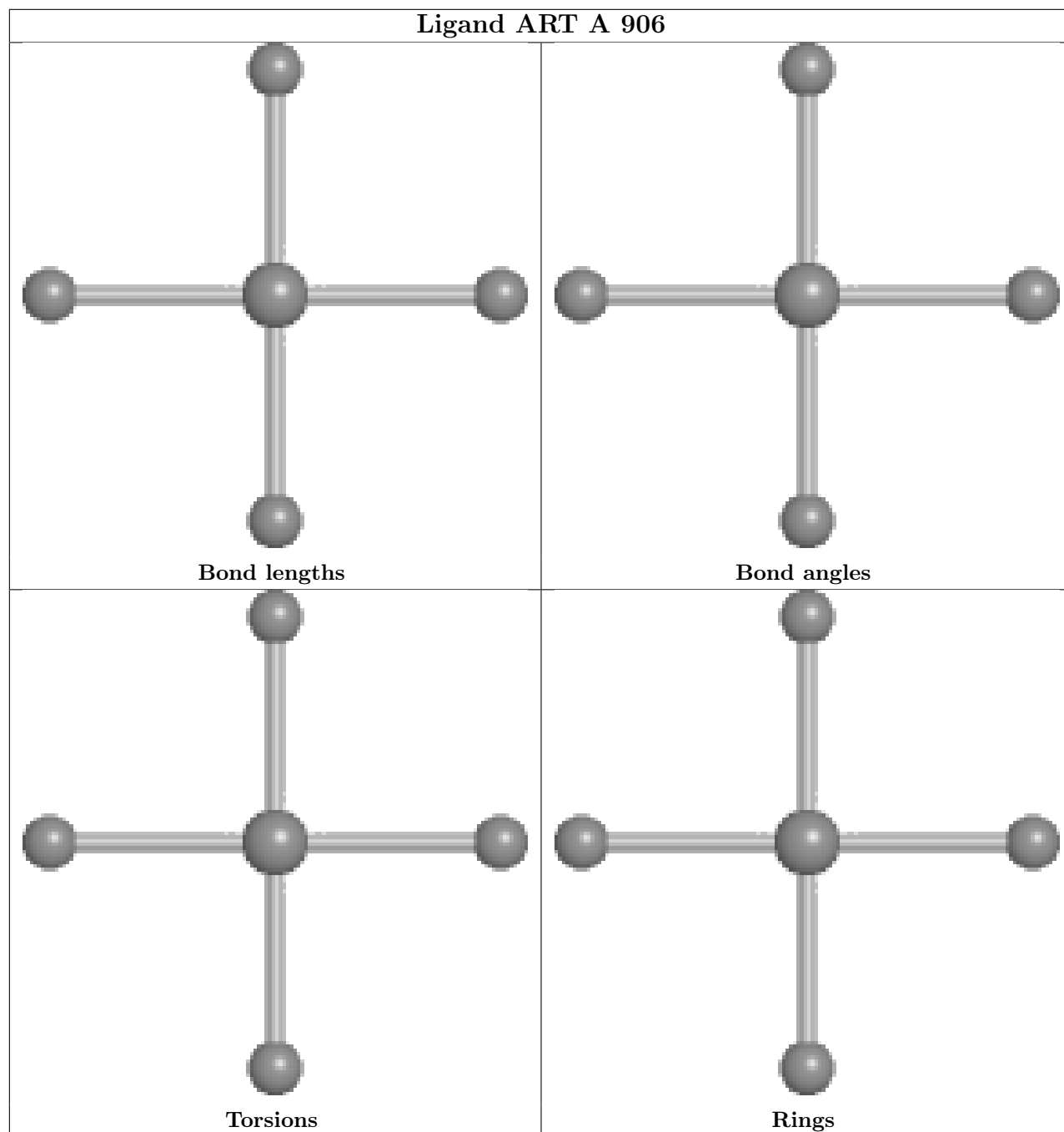
Continued from previous page...

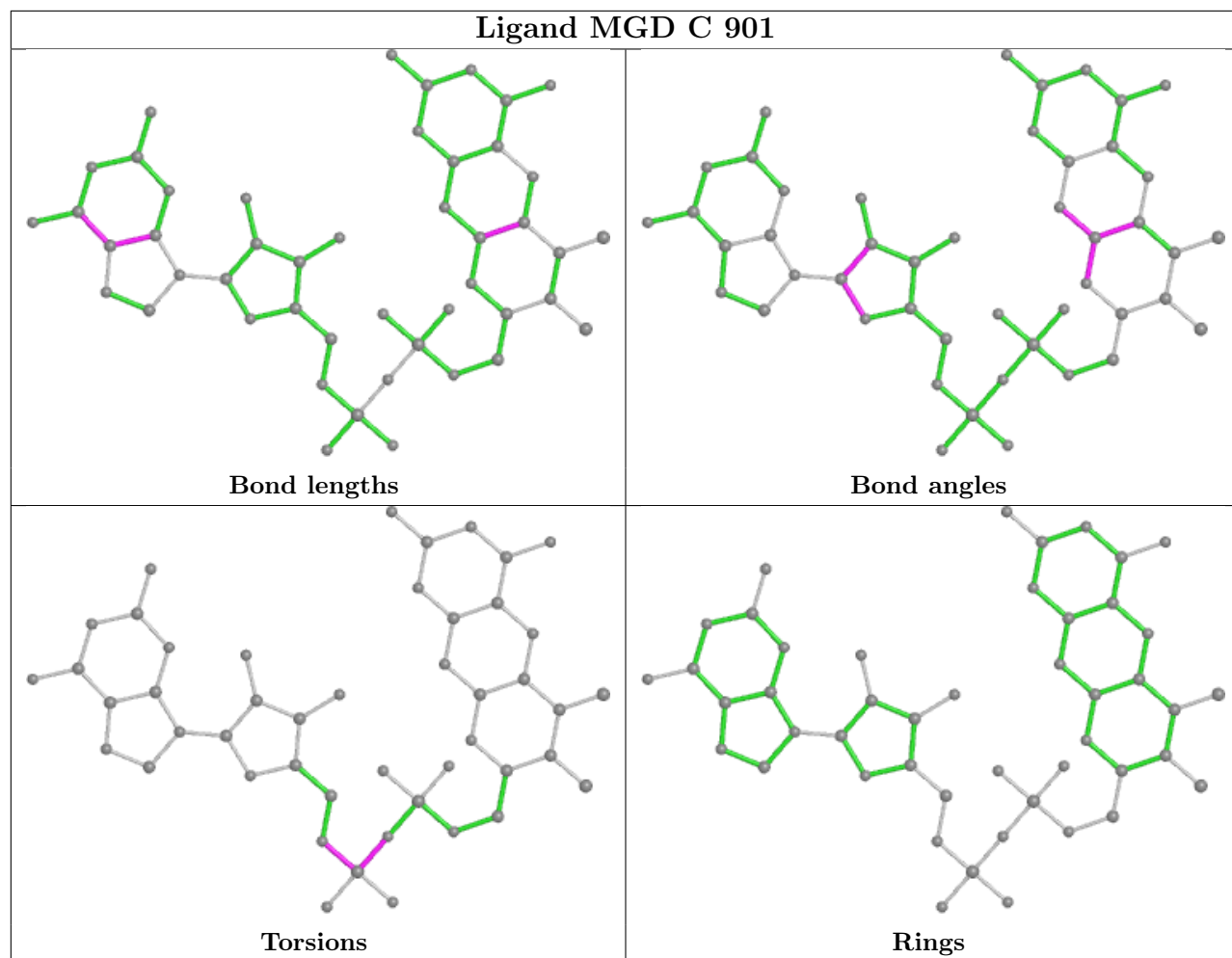
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	902	MGD	1	0
10	G	2401	EDO	2	0
9	G	2408	GOL	1	0
9	E	909	GOL	4	0
9	C	906	GOL	3	0
9	A	908	GOL	1	0
10	A	910	EDO	1	0
10	C	913	EDO	2	0
8	E	904	ART	1	0
10	D	202	EDO	1	0

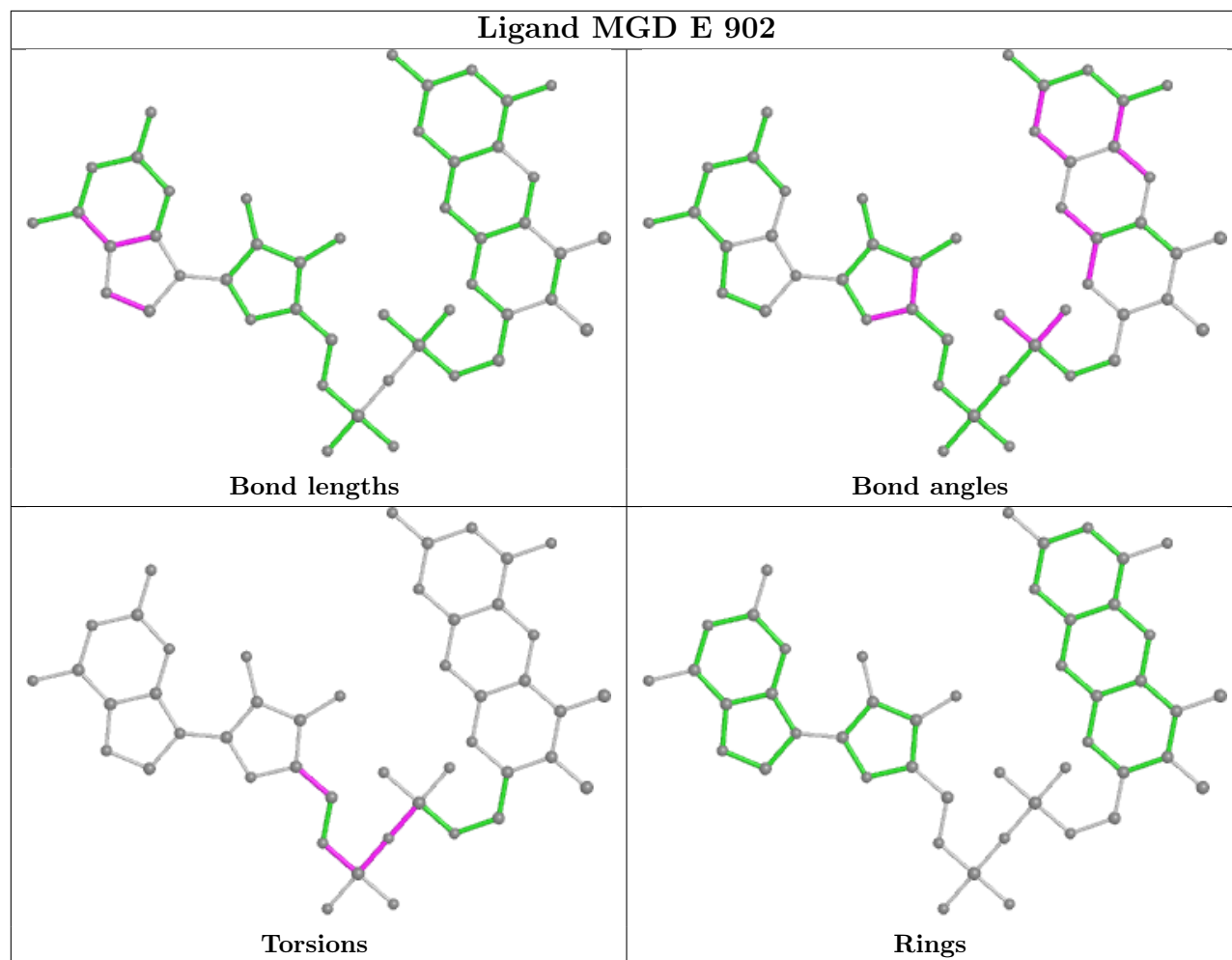
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

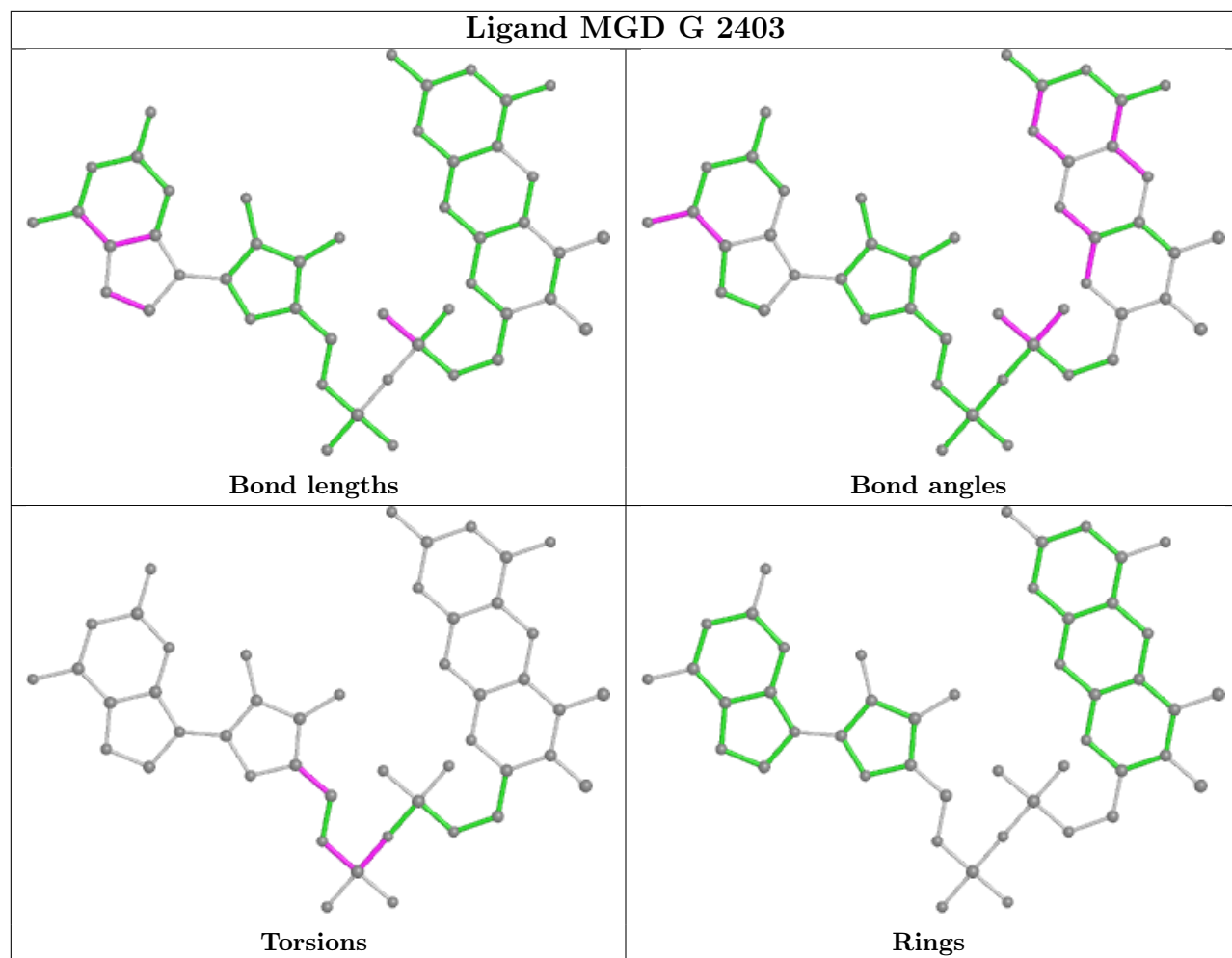


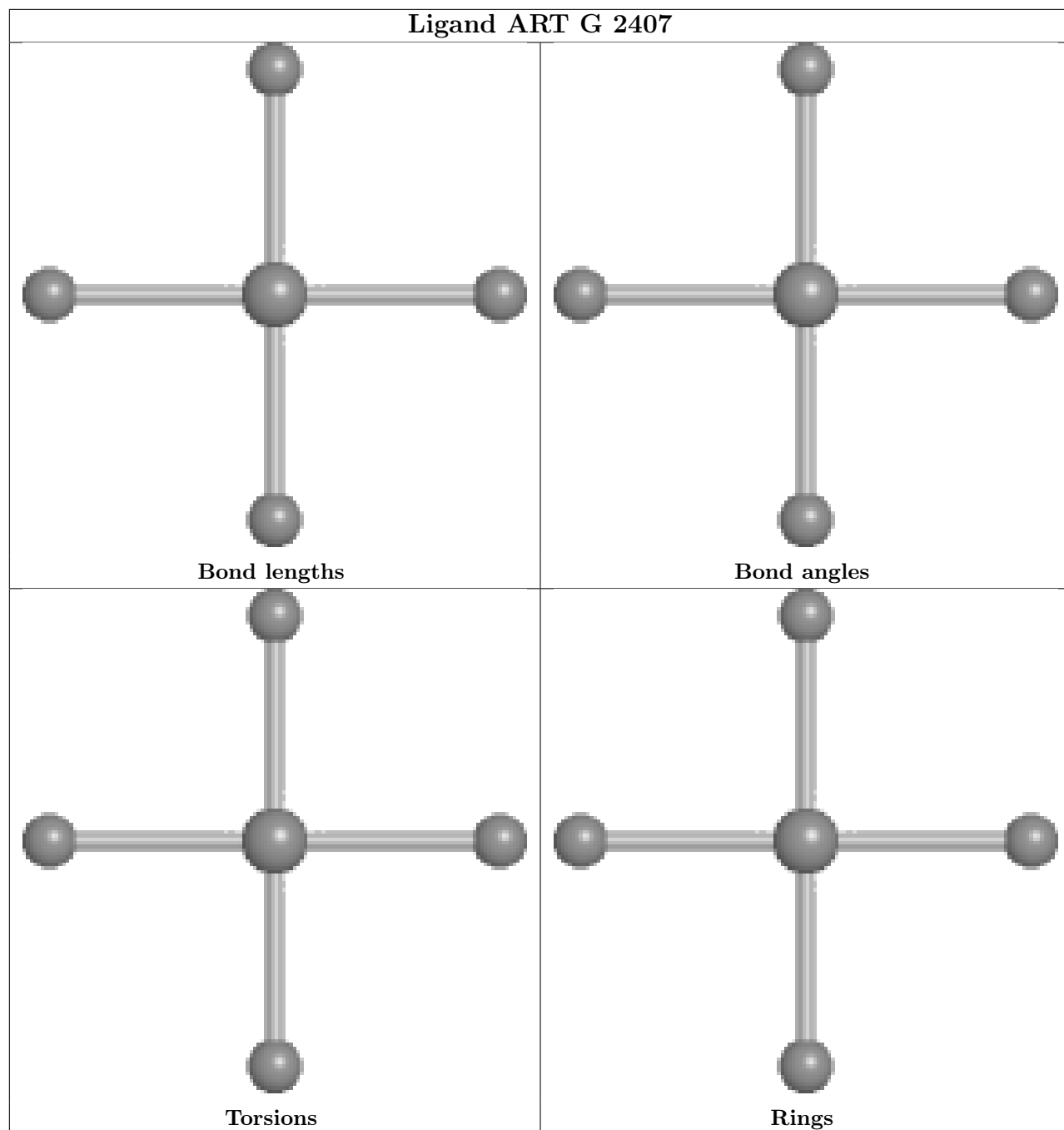


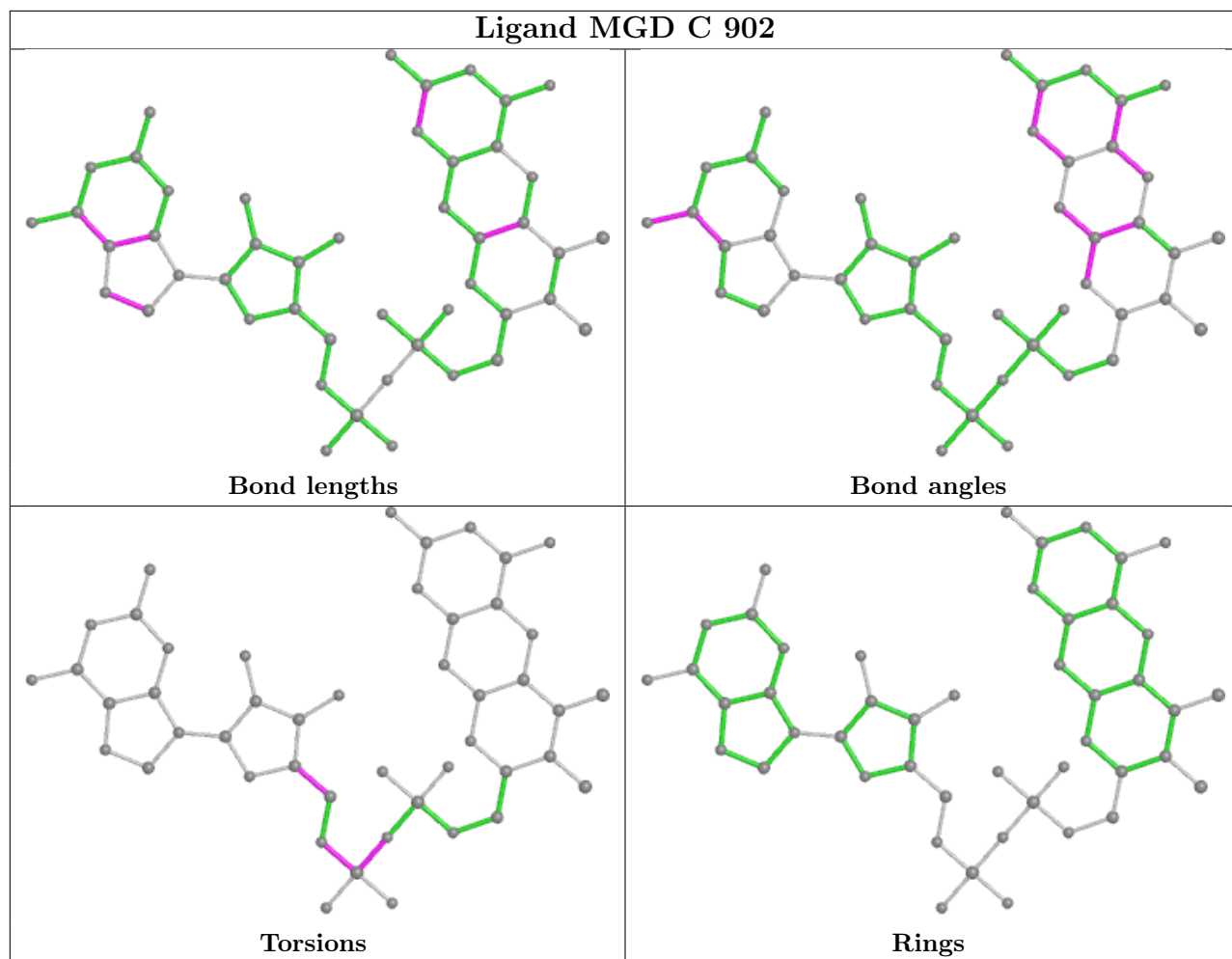


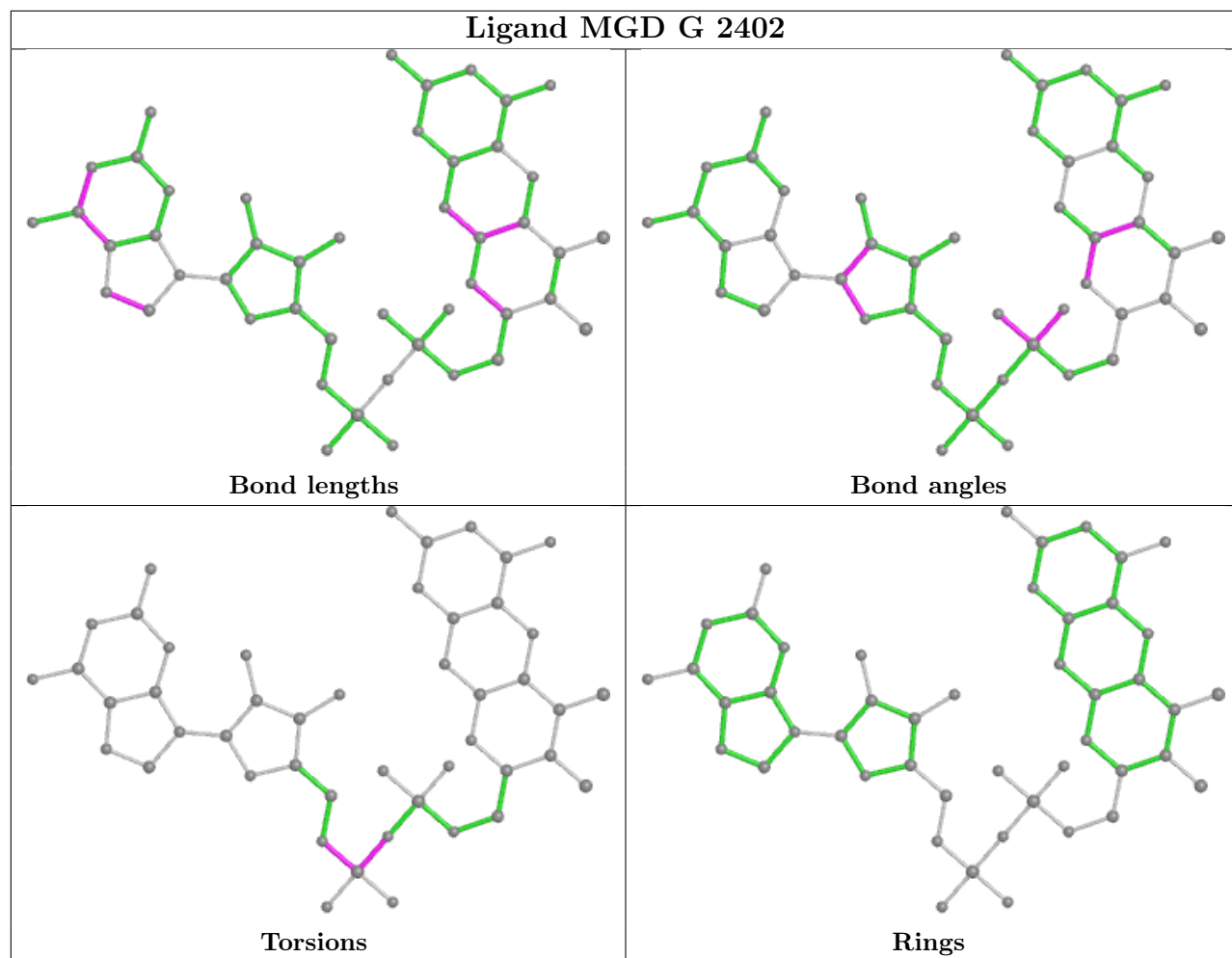


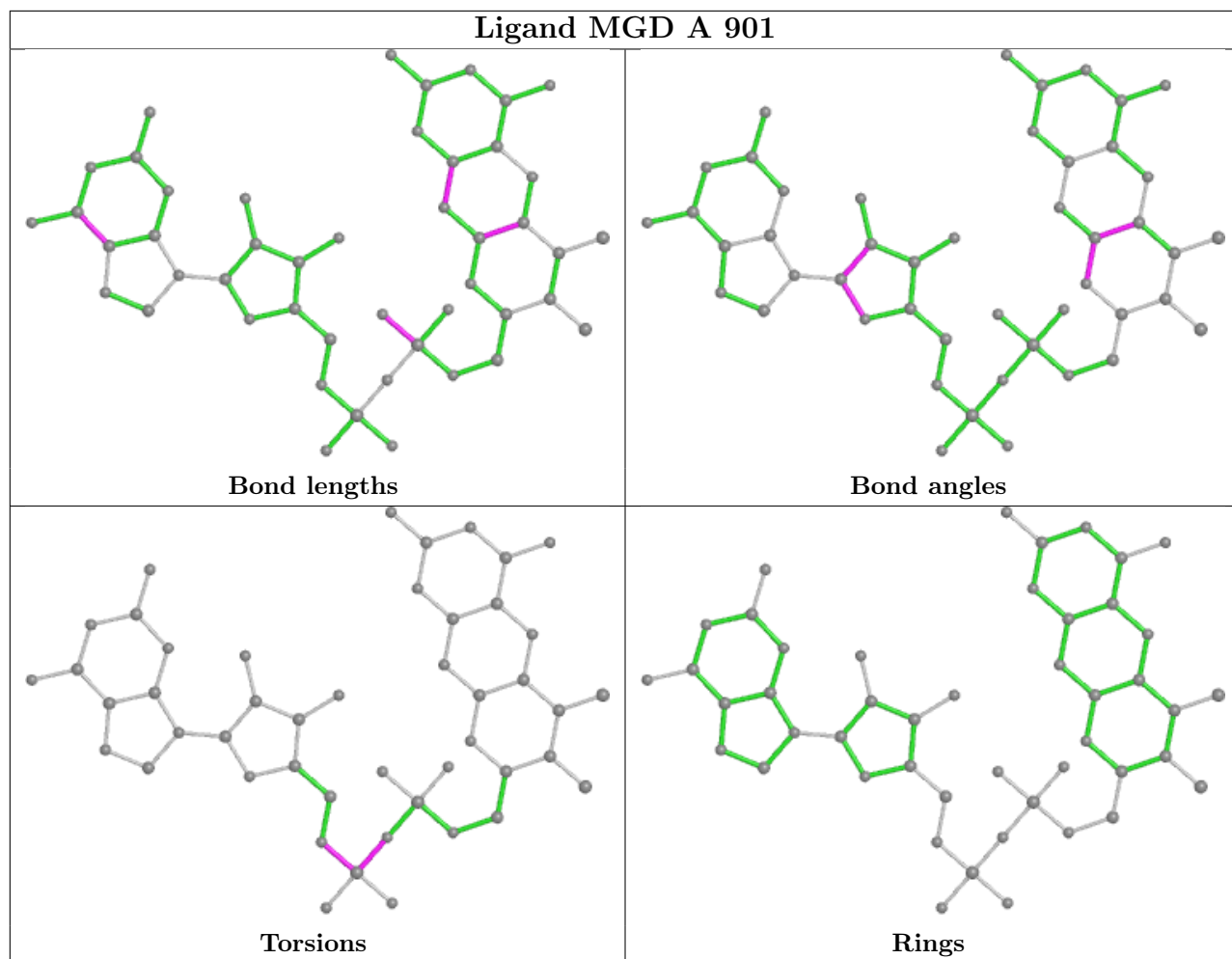


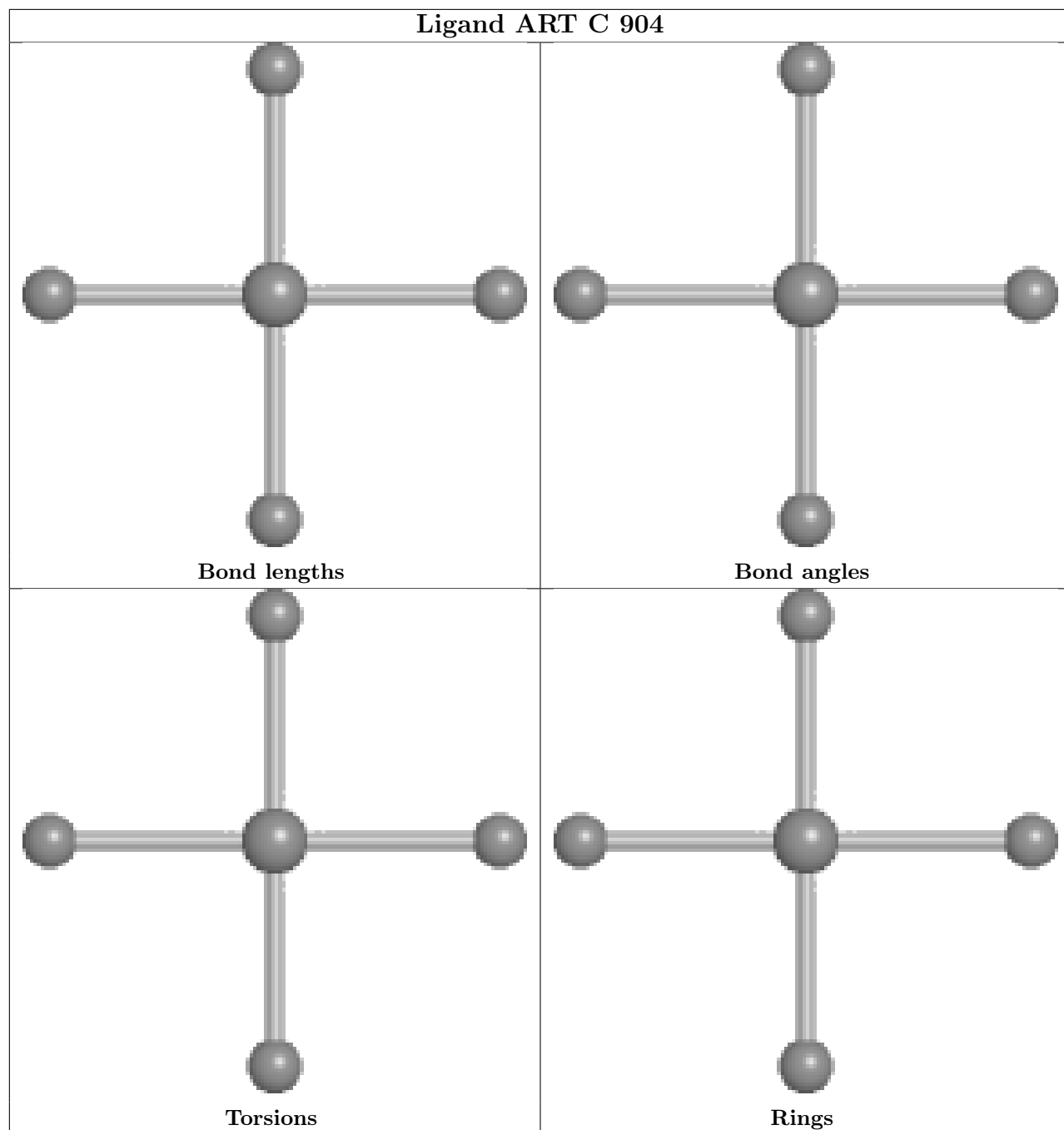


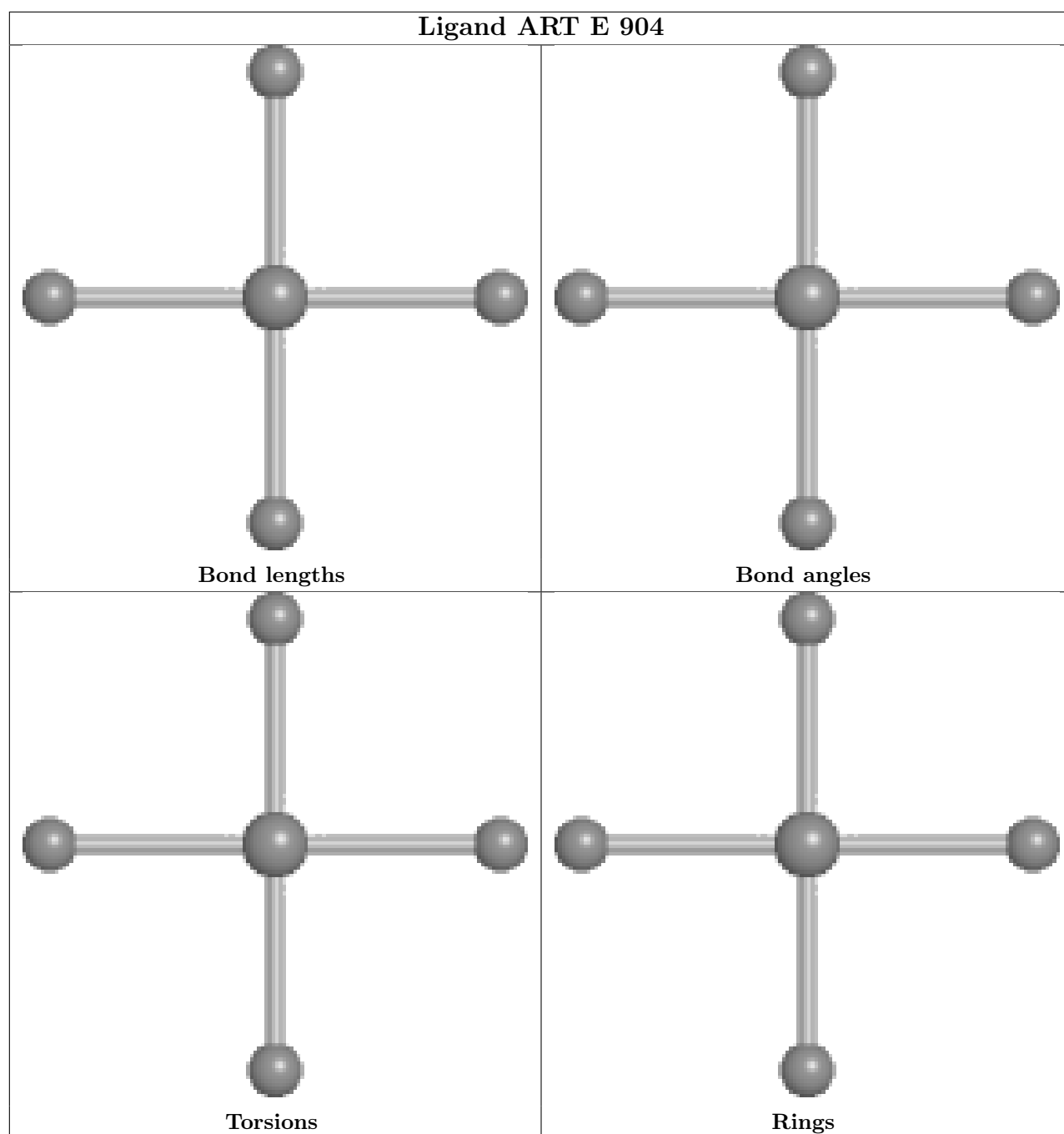












5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

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

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	822/823 (99%)	-0.28	6 (0%) 87 88	7, 12, 26, 46	4 (0%)
1	C	823/823 (100%)	-0.11	9 (1%) 80 81	8, 14, 28, 61	4 (0%)
1	E	822/823 (99%)	-0.16	6 (0%) 87 88	7, 14, 26, 59	4 (0%)
2	B	134/134 (100%)	-0.23	2 (1%) 73 73	8, 14, 28, 47	0
2	D	133/134 (99%)	-0.23	1 (0%) 86 86	9, 14, 28, 49	1 (0%)
2	F	134/134 (100%)	-0.19	2 (1%) 73 73	9, 15, 29, 49	1 (0%)
2	H	133/134 (99%)	-0.23	1 (0%) 86 86	10, 14, 27, 42	0
3	G	821/823 (99%)	-0.10	6 (0%) 87 88	8, 14, 27, 51	1 (0%)
All	All	3822/3828 (99%)	-0.17	33 (0%) 84 84	7, 14, 28, 61	15 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	3	PRO	4.3
1	C	613	GLY	4.0
1	A	582	GLY	3.5
1	A	612	PRO	3.4
1	E	435	ASP	3.3

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands

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.

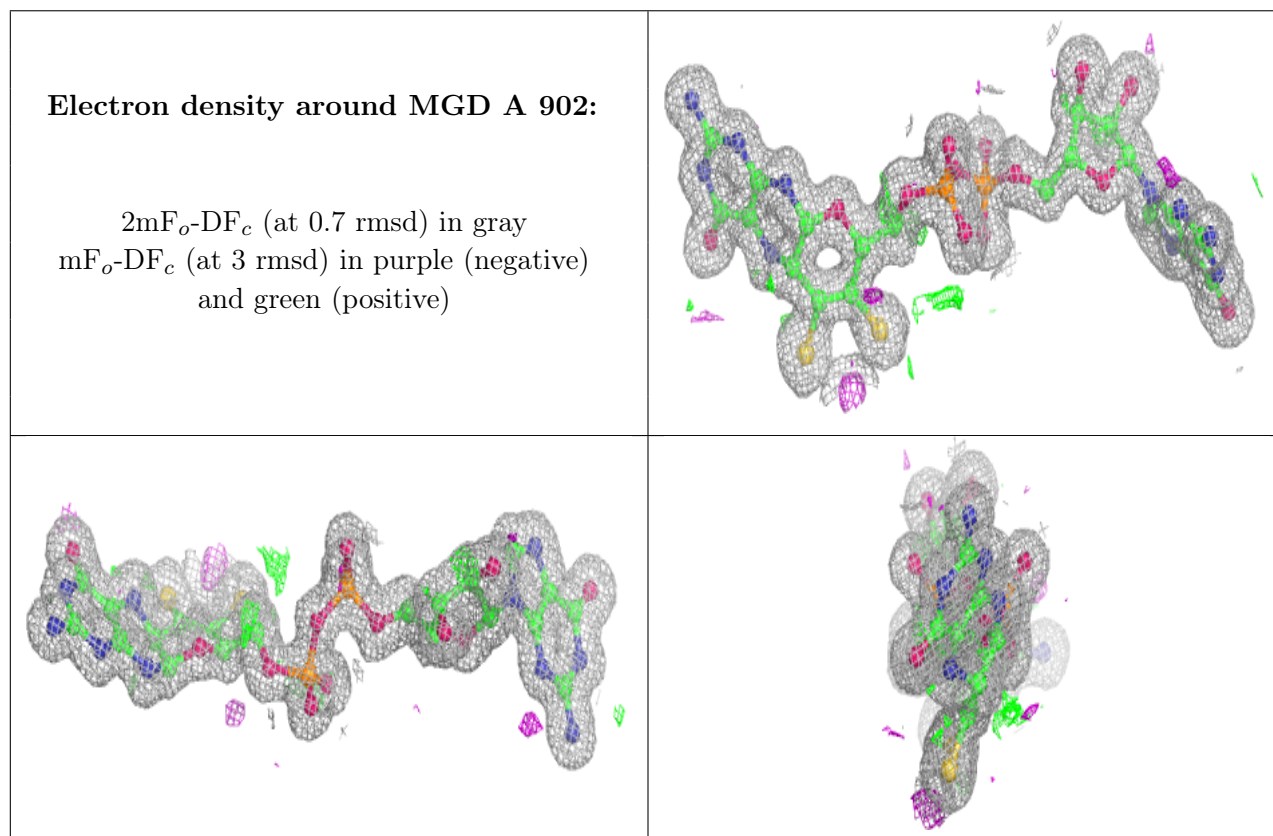
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
10	EDO	A	909	4/4	0.64	0.13	33,39,39,44	0
10	EDO	E	911	4/4	0.73	0.14	31,39,40,40	0
12	PEG	H	202	7/7	0.80	0.15	32,37,42,43	0
10	EDO	C	912	4/4	0.82	0.12	38,49,52,55	0
9	GOL	B	202	6/6	0.83	0.12	48,50,54,55	0
9	GOL	E	910	6/6	0.83	0.16	28,40,47,49	0
10	EDO	D	202	4/4	0.84	0.21	42,42,42,44	0
10	EDO	A	911	4/4	0.85	0.14	32,34,38,40	0
10	EDO	G	2401	4/4	0.85	0.18	33,33,37,38	0
10	EDO	C	910	4/4	0.85	0.09	39,42,42,42	0
9	GOL	E	909	6/6	0.86	0.23	38,42,46,50	0
12	PEG	C	911	7/7	0.86	0.22	35,53,64,64	0
9	GOL	G	2408	6/6	0.86	0.16	20,27,28,46	0
9	GOL	G	2410	6/6	0.89	0.15	27,37,46,47	0
9	GOL	G	2411	6/6	0.90	0.19	20,36,37,39	0
9	GOL	A	907	6/6	0.90	0.11	20,23,24,24	0
9	GOL	E	908	6/6	0.91	0.11	22,25,30,33	0
10	EDO	C	913	4/4	0.92	0.23	21,23,24,44	0
9	GOL	C	906	6/6	0.92	0.12	20,31,34,37	0
9	GOL	C	907	6/6	0.94	0.13	19,23,25,25	0
9	GOL	C	908	6/6	0.94	0.10	14,21,29,34	0
9	GOL	A	912	6/6	0.94	0.10	16,25,28,32	0
10	EDO	A	910	4/4	0.95	0.15	21,27,38,38	0
9	GOL	A	908	6/6	0.95	0.09	13,17,26,33	0
9	GOL	G	2409	6/6	0.96	0.09	12,20,31,33	0
9	GOL	E	907	6/6	0.97	0.10	19,33,37,38	0
6	FC6	G	2405	13/13	0.98	0.09	13,20,27,30	1
4	MGD	A	902	47/47	0.98	0.07	7,9,10,11	0
4	MGD	C	901	47/47	0.98	0.06	7,9,10,11	0
4	MGD	C	902	47/47	0.98	0.07	8,10,12,12	0
4	MGD	E	901	47/47	0.98	0.07	7,9,10,11	0
4	MGD	E	902	47/47	0.98	0.07	8,10,11,13	0
4	MGD	G	2402	47/47	0.98	0.07	7,9,11,11	0
4	MGD	G	2403	47/47	0.98	0.08	8,10,12,12	0
6	FC6	E	905	13/13	0.98	0.11	14,22,28,29	1
4	MGD	A	901	47/47	0.99	0.07	6,8,9,9	0
6	FC6	A	904	13/13	0.99	0.09	13,19,26,26	1

Continued on next page...

Continued from previous page...

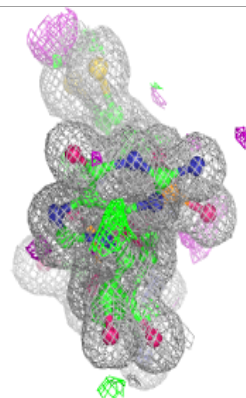
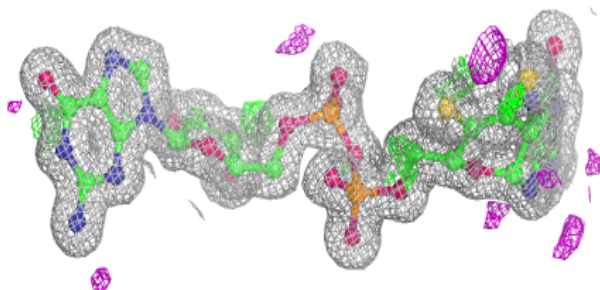
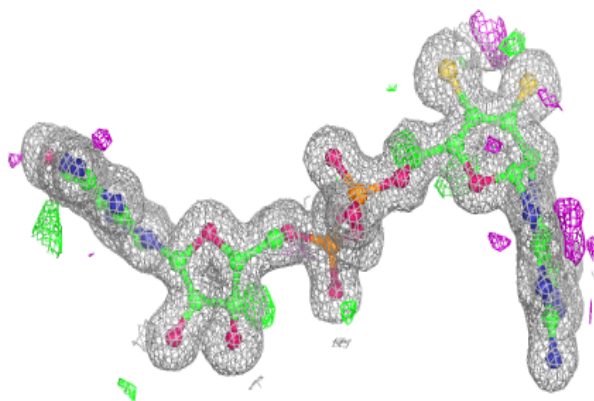
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	ART	E	904	5/5	0.99	0.08	7,12,16,18	1
8	ART	G	2407	5/5	0.99	0.07	8,12,16,18	1
6	FC6	C	905	13/13	0.99	0.07	15,20,30,31	1
8	ART	A	906	5/5	1.00	0.07	9,11,13,13	2
8	ART	C	904	5/5	1.00	0.06	10,12,17,18	1
5	F3S	A	903	7/7	1.00	0.07	8,8,8,8	0
5	F3S	C	903	7/7	1.00	0.06	9,9,9,9	0
5	F3S	E	903	7/7	1.00	0.05	8,8,9,9	0
5	F3S	G	2404	7/7	1.00	0.07	8,8,9,9	0
7	4MO	A	905	1/1	1.00	0.04	8,8,8,8	0
7	4MO	C	909	1/1	1.00	0.04	11,11,11,11	0
11	FES	B	201	4/4	1.00	0.05	10,11,11,12	0
11	FES	D	201	4/4	1.00	0.05	10,11,12,13	0
11	FES	F	201	4/4	1.00	0.05	10,11,12,13	0
11	FES	H	201	4/4	1.00	0.05	11,11,11,13	0
7	4MO	E	906	1/1	1.00	0.04	8,8,8,8	0
7	4MO	G	2406	1/1	1.00	0.04	11,11,11,11	0

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

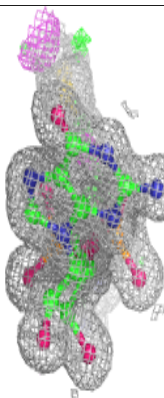
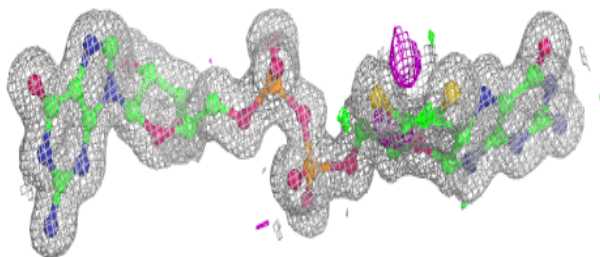
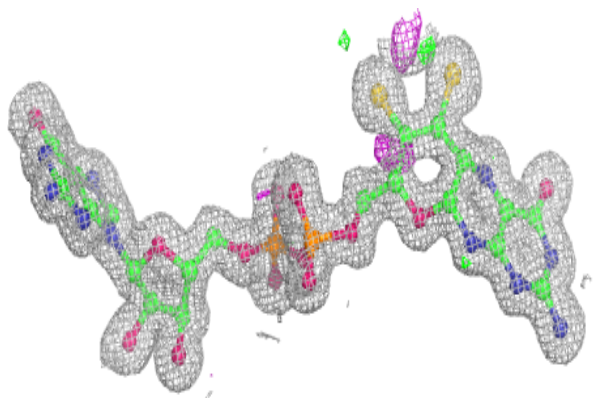


Electron density around MGD C 901:

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

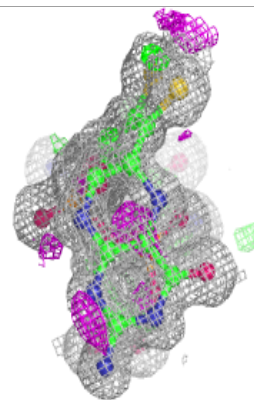
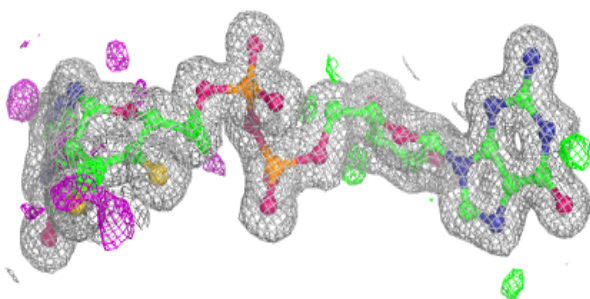
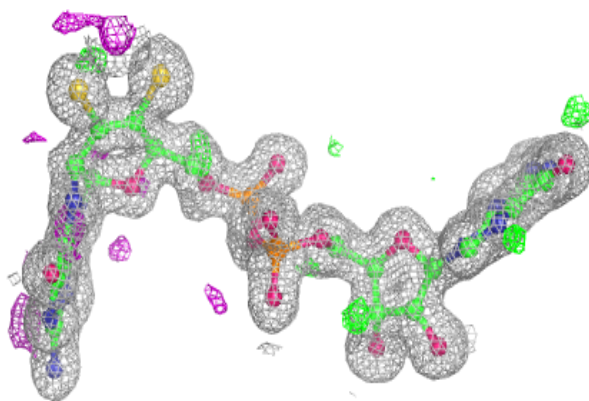
**Electron density around MGD C 902:**

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

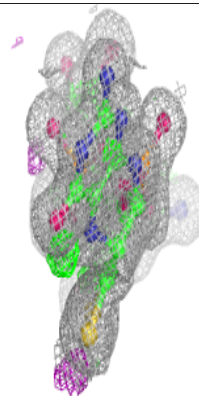
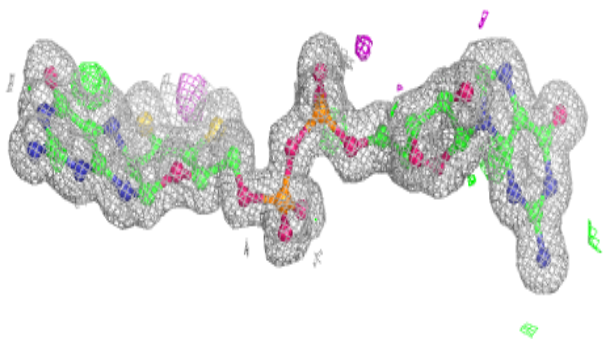
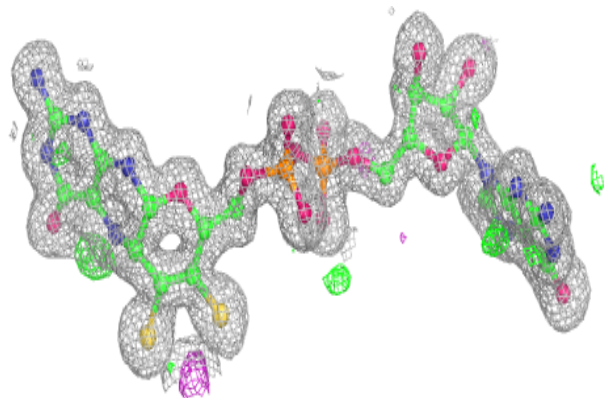


Electron density around MGD E 901:

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

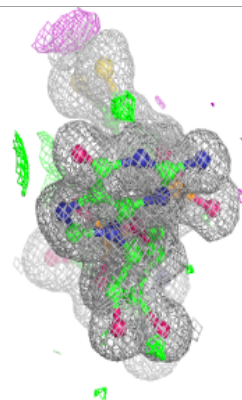
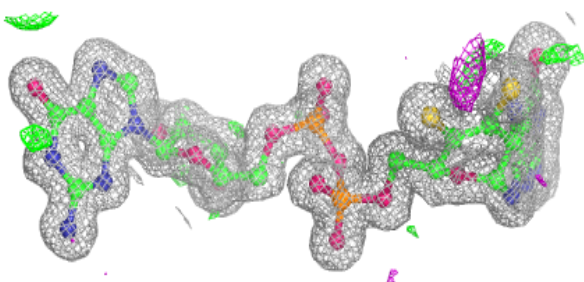
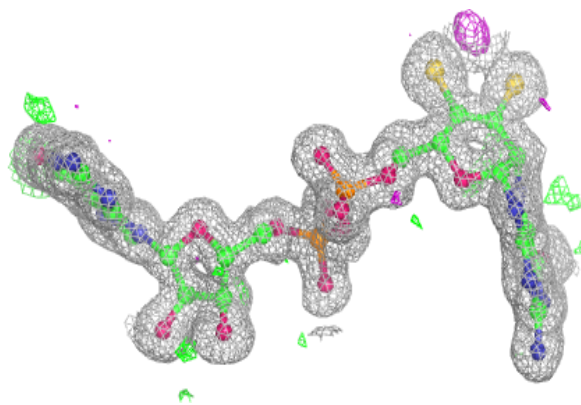
**Electron density around MGD E 902:**

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

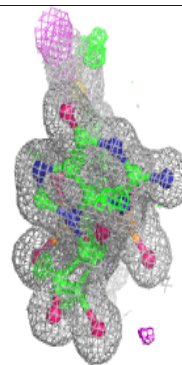
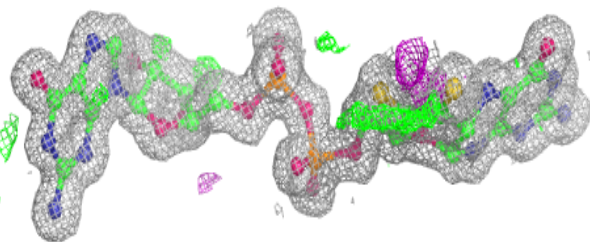
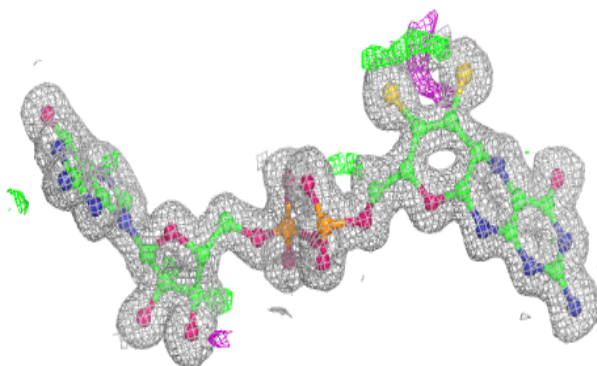


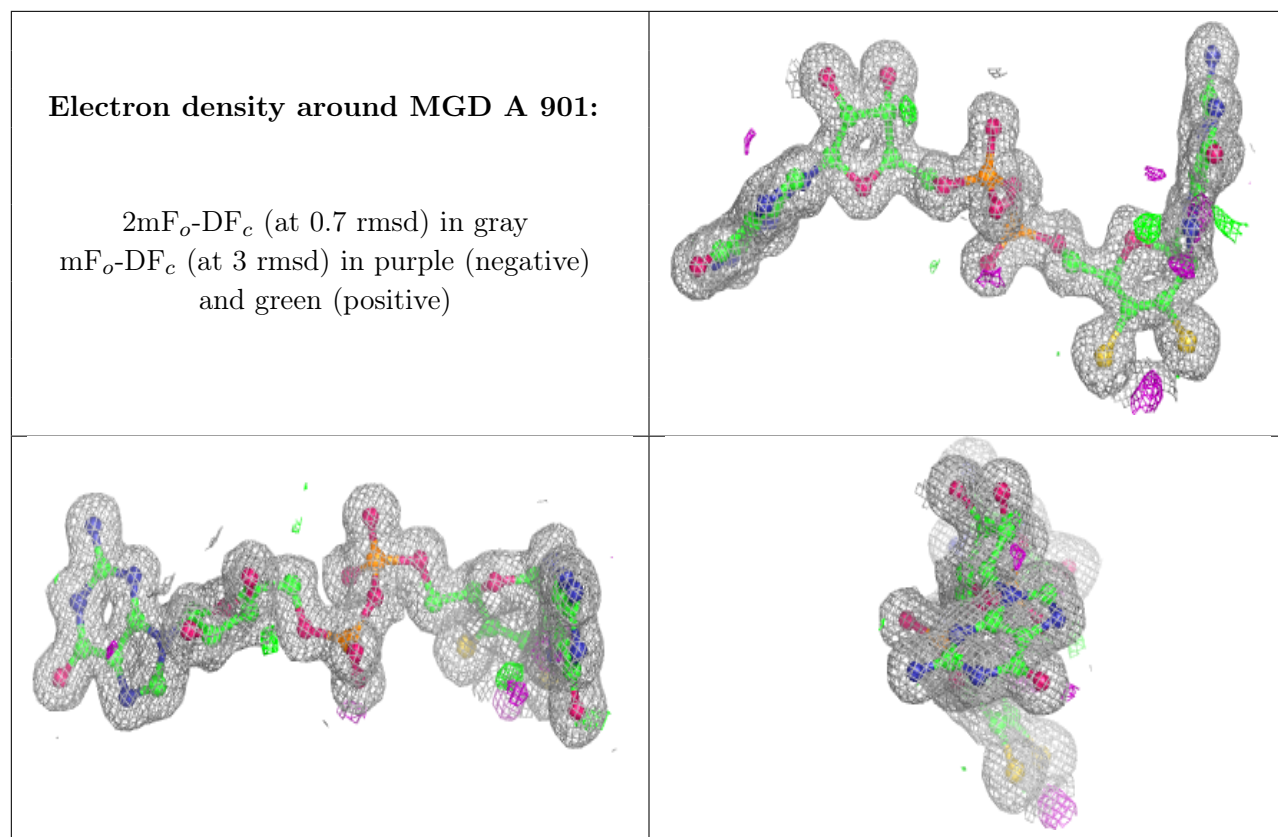
Electron density around MGD G 2402:

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

**Electron density around MGD G 2403:**

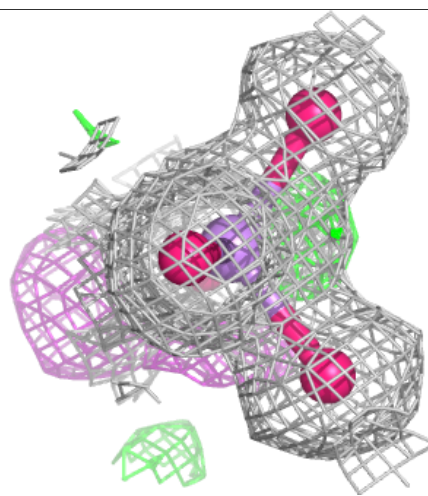
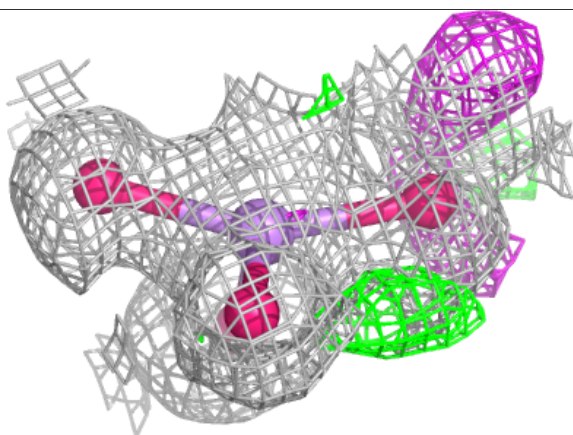
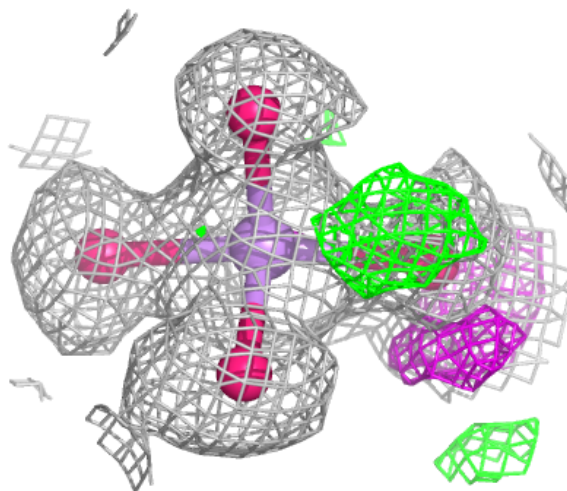
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





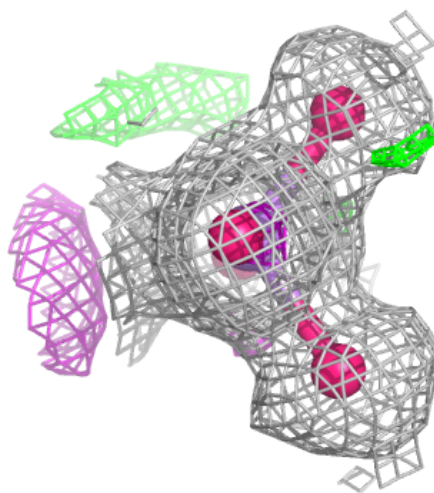
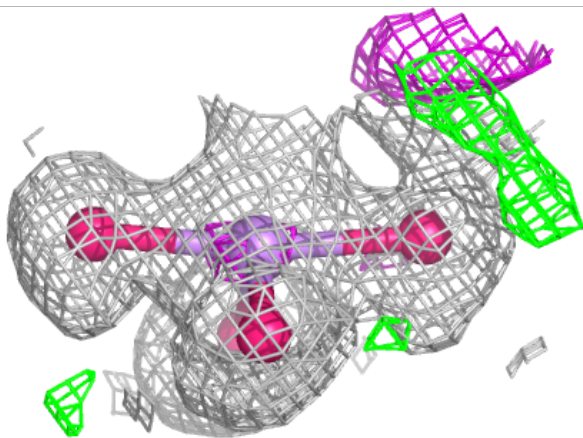
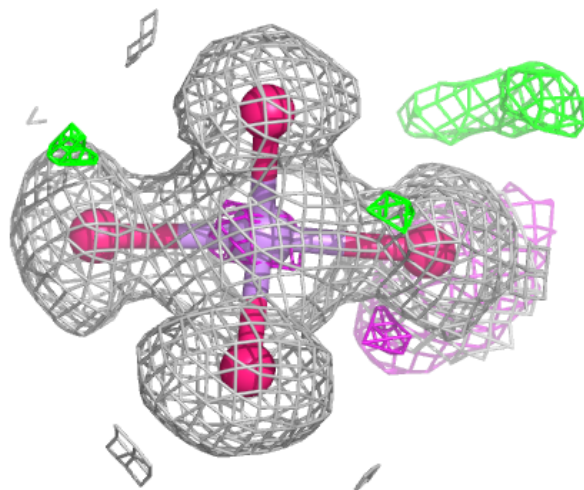
Electron density around ART E 904:

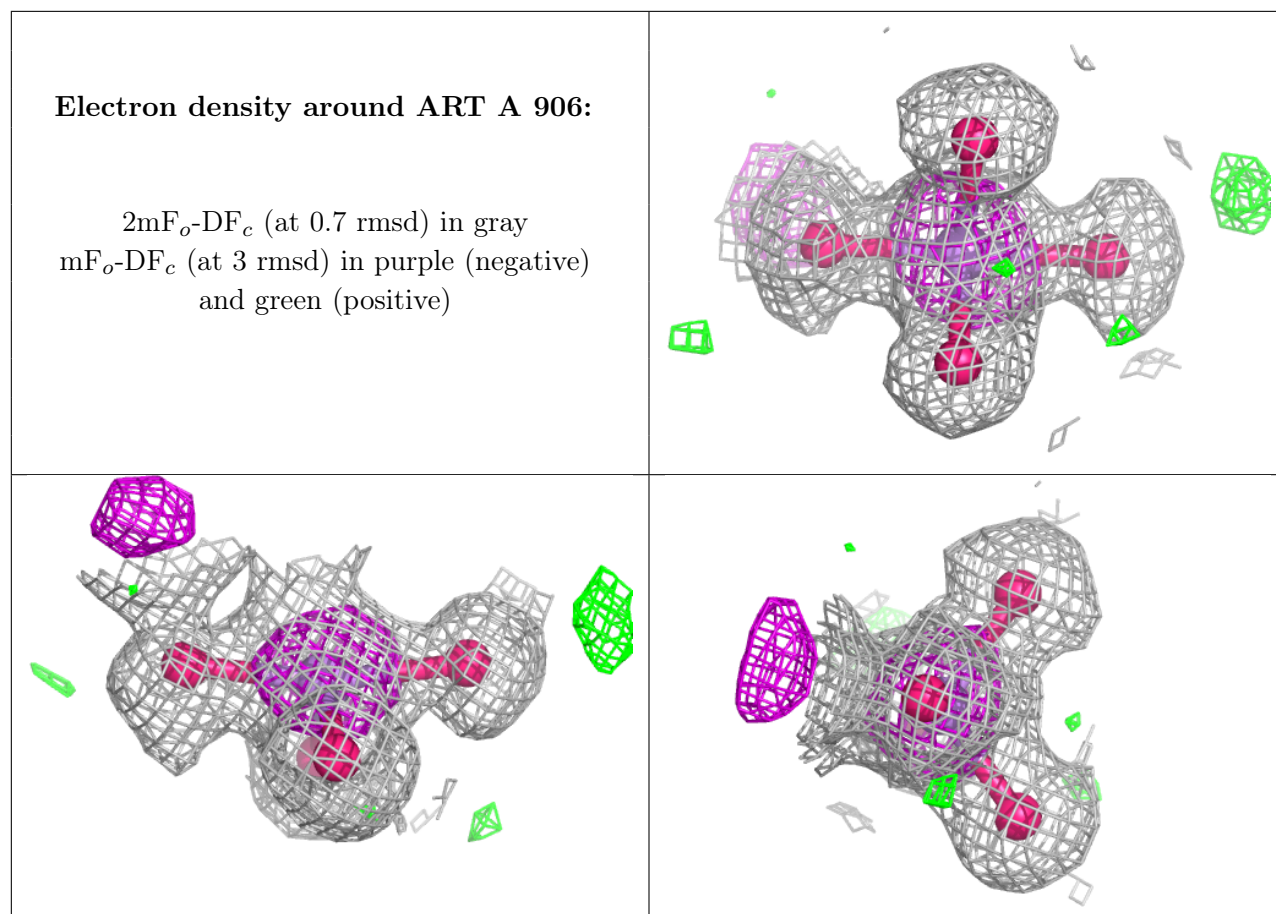
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around ART G 2407:

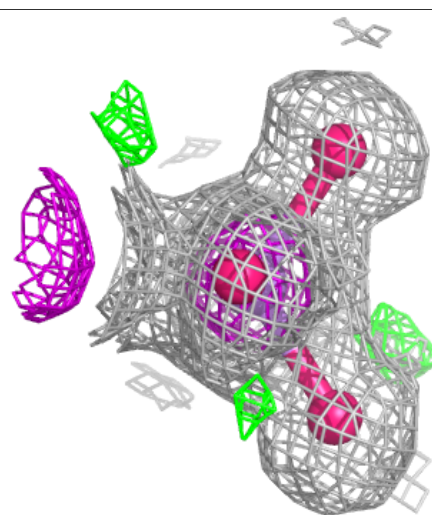
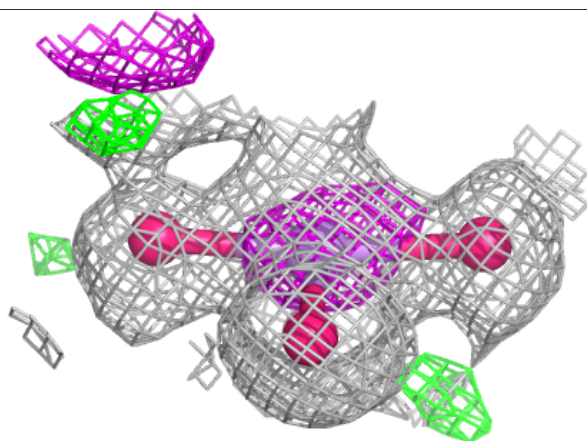
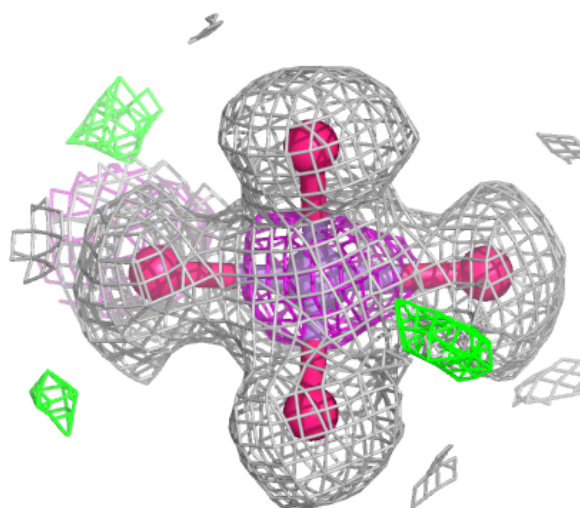
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





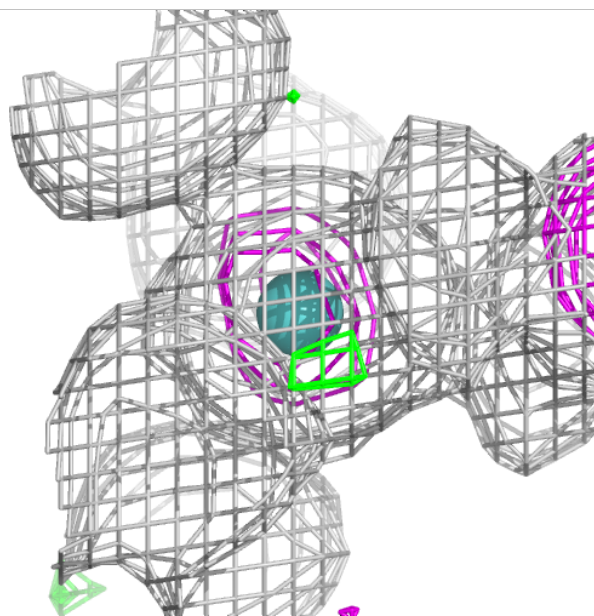
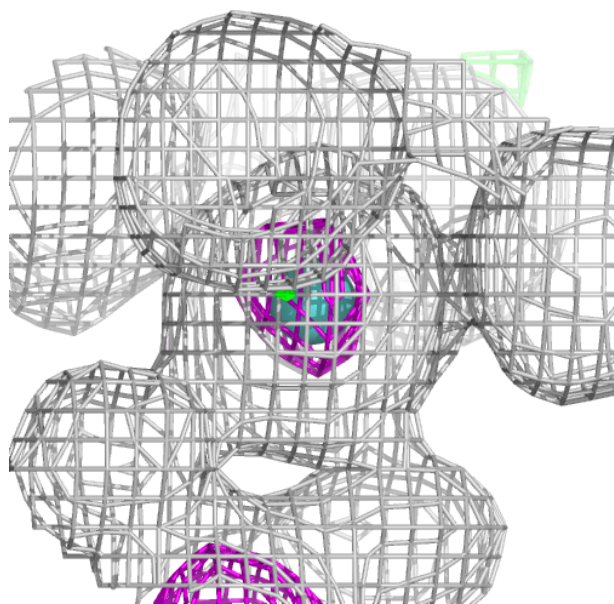
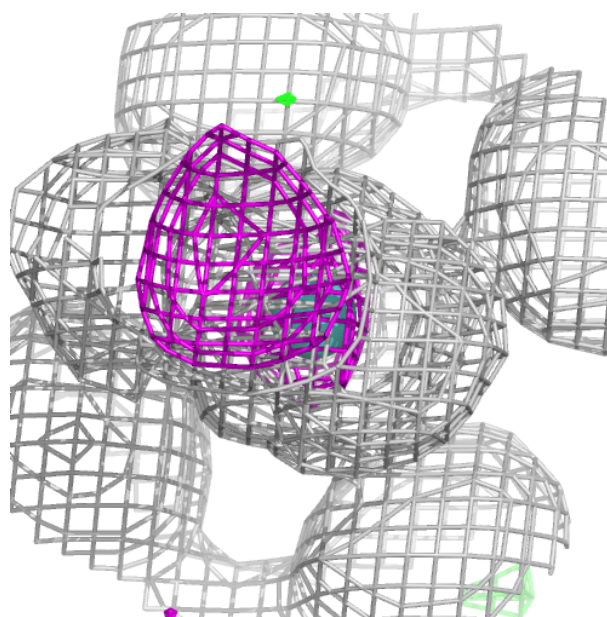
Electron density around ART C 904:

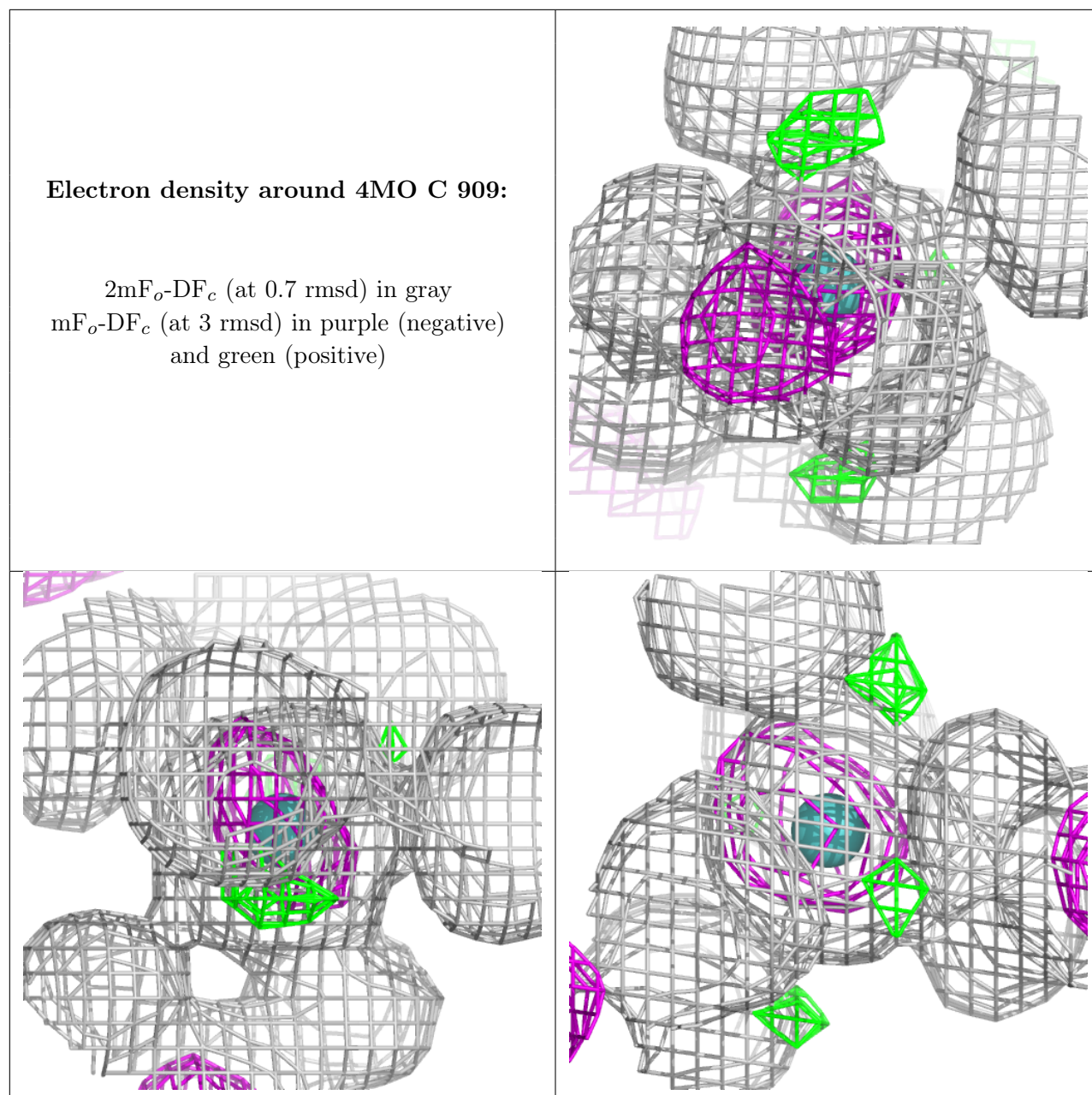
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

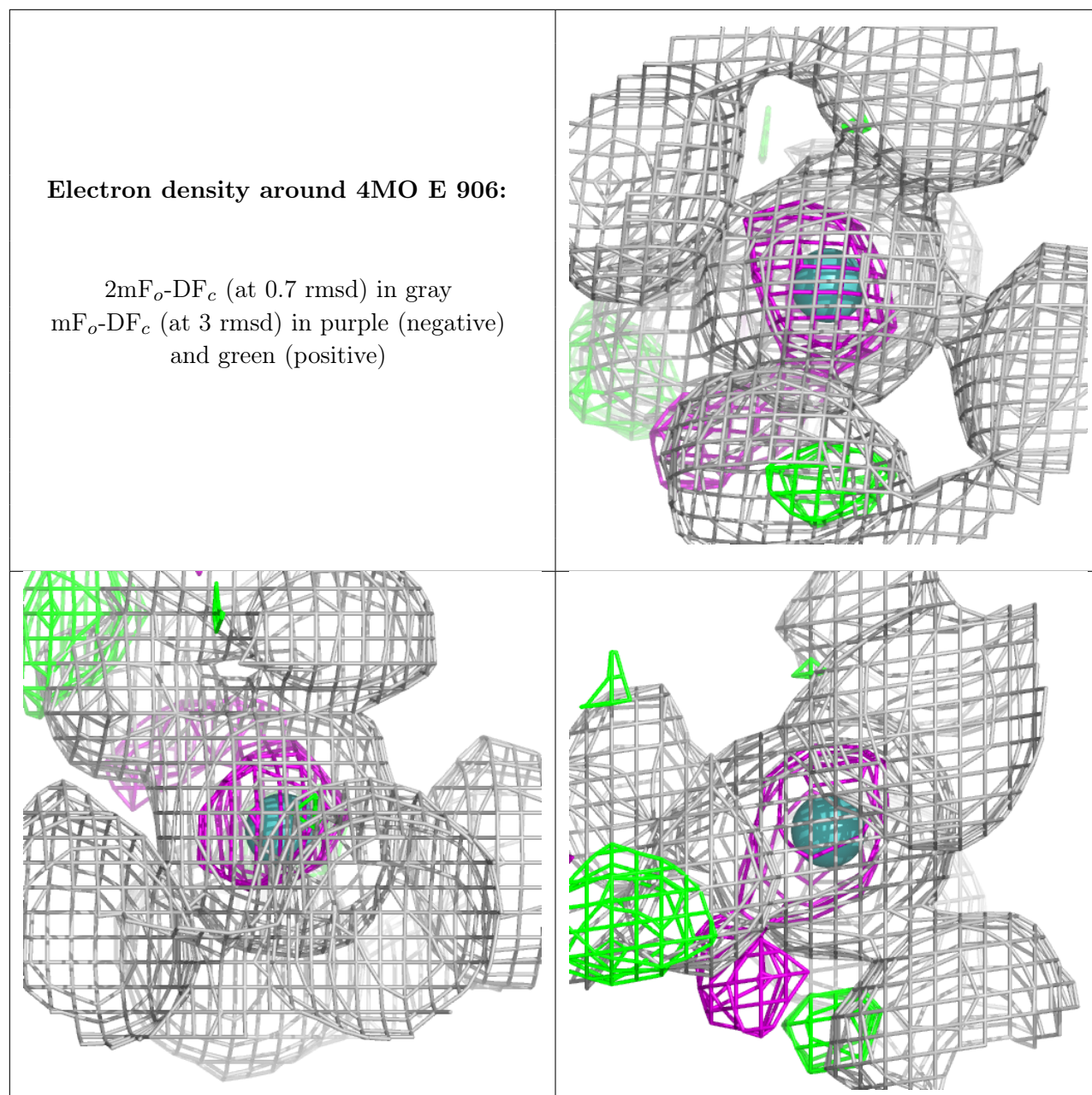


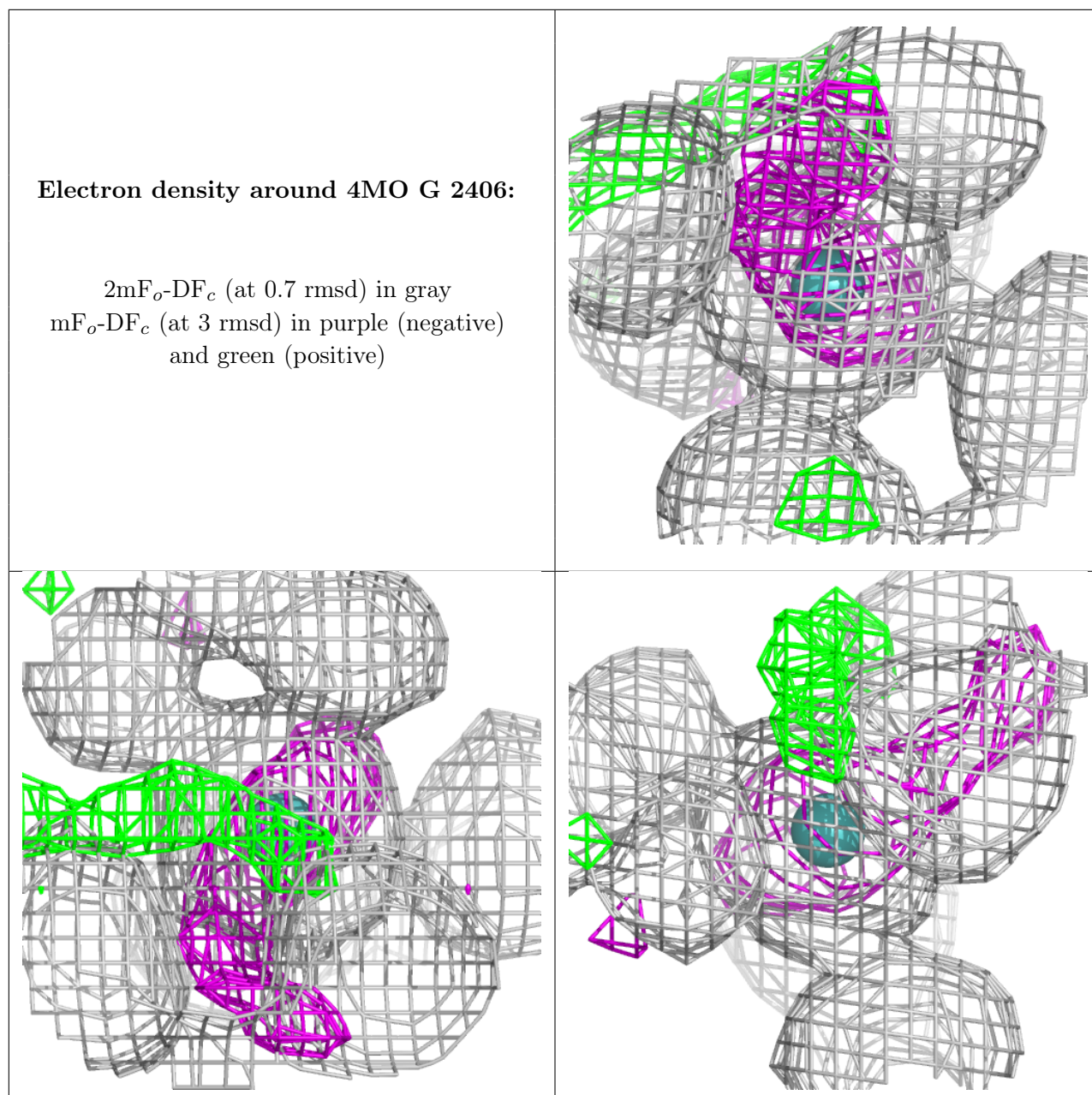
Electron density around 4MO A 905:

$2mF_o-DF_c$ (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.