wwPDB X-ray Structure Validation Summary Report (i)

Mar 30, 2025 – 03:55 am BST

PDB ID	:	9ERS / pdb_00009ers
Title	:	Hydrogenase-2 Ni-C state
Authors	:	Wong, K.L.; Carr, S.B.; Ash, P.A.; Vincent, K.A.
Deposited on	:	2024-03-25
Resolution	:	1.59 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:

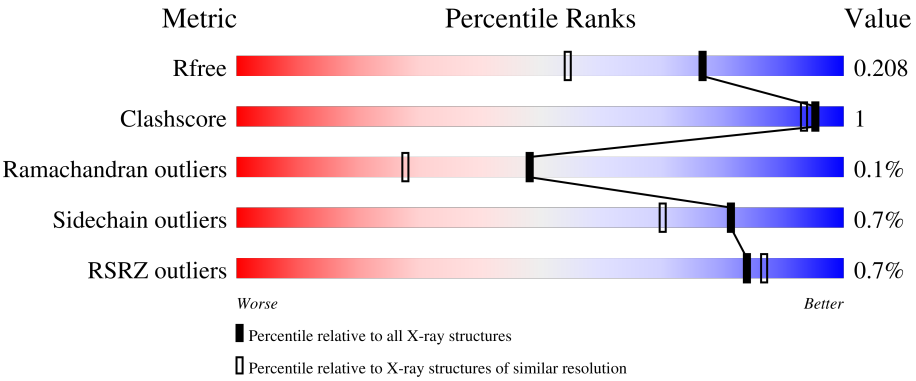
MolProbity	:	4.02b-467
Mogul	:	?.? (???), CSD ??CSD?? (????)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 1.59 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}	164625	4274 (1.60-1.60)
Clashscore	180529	4682 (1.60-1.60)
Ramachandran outliers	177936	4583 (1.60-1.60)
Sidechain outliers	177891	4582 (1.60-1.60)
RSRZ outliers	164620	4272 (1.60-1.60)

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	S	298	87% 10%
1	T	298	88% 10%
2	L	567	94%
2	M	567	94%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 14239 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 Hydrogenase-2 small chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	S	268	Total	C	N	O	S	0	4	0
			2066	1308	363	382	13			
1	T	267	Total	C	N	O	S	0	7	0
			2079	1317	363	386	13			

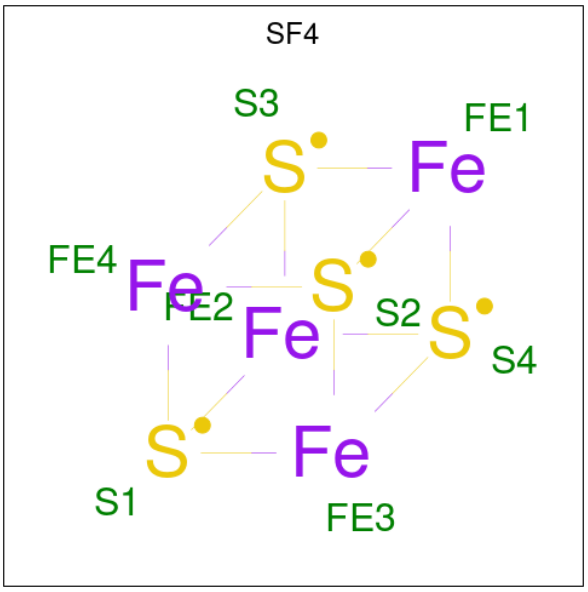
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	294	HIS	-	expression tag	UNP P69741
S	295	HIS	-	expression tag	UNP P69741
S	296	HIS	-	expression tag	UNP P69741
S	297	HIS	-	expression tag	UNP P69741
S	298	HIS	-	expression tag	UNP P69741
S	299	HIS	-	expression tag	UNP P69741
T	294	HIS	-	expression tag	UNP P69741
T	295	HIS	-	expression tag	UNP P69741
T	296	HIS	-	expression tag	UNP P69741
T	297	HIS	-	expression tag	UNP P69741
T	298	HIS	-	expression tag	UNP P69741
T	299	HIS	-	expression tag	UNP P69741

- Molecule 2 is a protein called Hydrogenase-2 large chain.

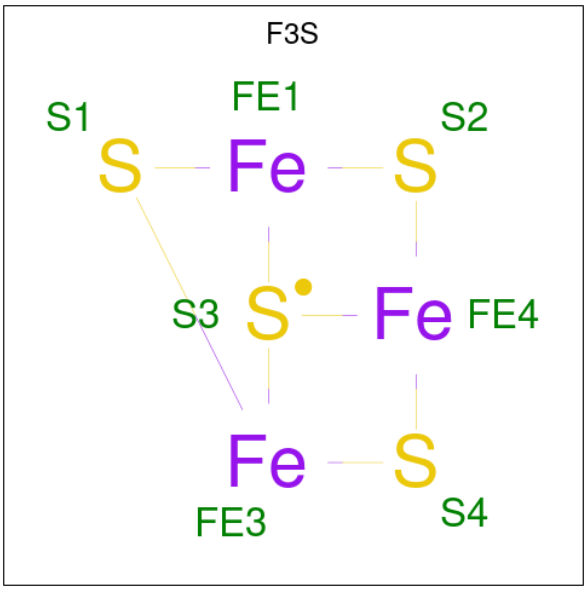
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	551	Total	C	N	O	S	0	6	0
			4320	2748	743	811	18			
2	M	551	Total	C	N	O	S	0	5	0
			4320	2747	744	811	18			

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



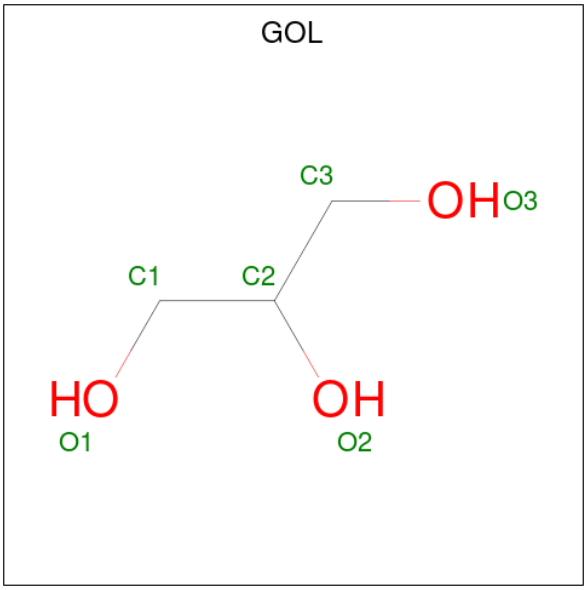
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	S	1	Total Fe S 8 4 4	0	0
3	S	1	Total Fe S 8 4 4	0	0
3	T	1	Total Fe S 8 4 4	0	0
3	T	1	Total Fe S 8 4 4	0	0

- Molecule 4 is FE3-S4 CLUSTER (CCD ID: F3S) (formula: Fe_3S_4).



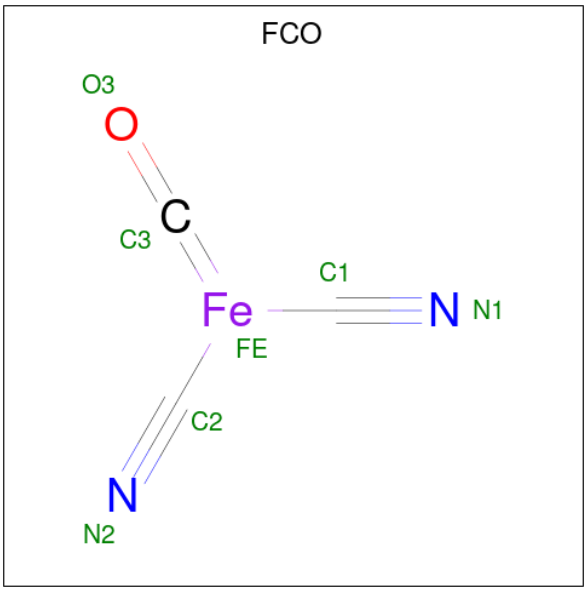
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	S	1	Total Fe S 7 3 4	0	0
4	T	1	Total Fe S 7 3 4	0	0

- Molecule 5 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	S	1	Total C O 6 3 3	0	0
5	S	1	Total C O 6 3 3	0	0
5	T	1	Total C O 6 3 3	0	0
5	M	1	Total C O 6 3 3	0	0
5	M	1	Total C O 6 3 3	0	0
5	M	1	Total C O 6 3 3	0	0

- Molecule 6 is CARBONMONOXIDE-(DICYANO) IRON (CCD ID: FCO) (formula: C_3FeN_2O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	L	1	Total C Fe N O 7 3 1 2 1	0	0
6	M	1	Total C Fe N O 7 3 1 2 1	0	0

- Molecule 7 is NICKEL (II) ION (CCD ID: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	L	1	Total Ni 1 1	0	0
7	M	1	Total Ni 1 1	0	0

- Molecule 8 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	L	1	Total Mg 1 1	0	0
8	M	2	Total Mg 2 2	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	S	225	Total O 225 225	0	0

Continued on next page...

Continued from previous page...

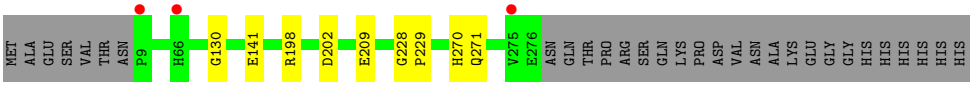
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	L	500	Total O 500 500	0	0
9	T	202	Total O 202 202	0	0
9	M	426	Total O 426 426	0	0

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.

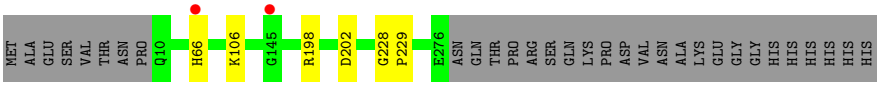
• Molecule 1: Hydrogenase-2 small chain

Chain S:



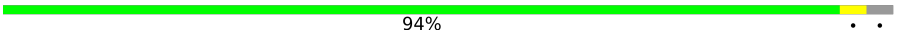
• Molecule 1: Hydrogenase-2 small chain

Chain T:



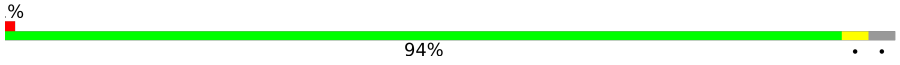
• Molecule 2: Hydrogenase-2 large chain

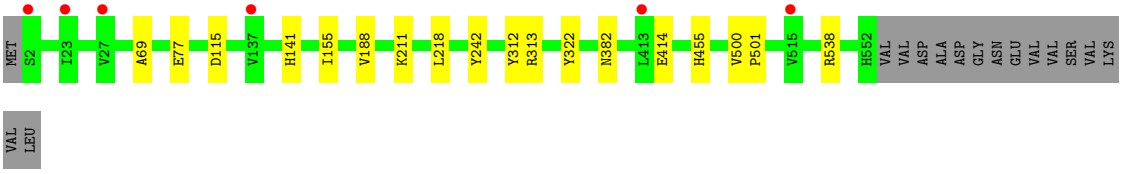
Chain L:



• Molecule 2: Hydrogenase-2 large chain

Chain M:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.86Å 100.06Å 168.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.73 – 1.59 44.73 – 1.59	Depositor EDS
% Data completeness (in resolution range)	99.3 (44.73-1.59) 99.3 (44.73-1.59)	Depositor EDS
R_{merge}	0.31	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.01 (at 1.59Å)	Xtriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.173 , 0.197 0.186 , 0.208	Depositor DCC
R_{free} test set	11119 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	18.8	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.010 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	14239	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% 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: F3S, GOL, NI, FCO, SF4, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	S	0.39	0/2127	0.64	0/2896
1	T	0.38	0/2145	0.63	0/2922
2	L	0.38	0/4439	0.67	0/6050
2	M	0.38	0/4433	0.66	0/6041
All	All	0.38	0/13144	0.66	0/17909

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	S	0	1
1	T	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	S	198	ARG	Sidechain
1	T	198	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	S	2066	0	1985	7	0
1	T	2079	0	2001	4	0
2	L	4320	0	4261	10	0
2	M	4320	0	4255	7	0
3	S	16	0	0	0	0
3	T	16	0	0	0	0
4	S	7	0	0	0	0
4	T	7	0	0	0	0
5	M	18	0	24	1	0
5	S	12	0	16	3	0
5	T	6	0	8	0	0
6	L	7	0	0	0	0
6	M	7	0	0	0	0
7	L	1	0	0	0	0
7	M	1	0	0	0	0
8	L	1	0	0	0	0
8	M	2	0	0	0	0
9	L	500	0	0	4	3
9	M	426	0	0	1	2
9	S	225	0	0	2	1
9	T	202	0	0	2	0
All	All	14239	0	12550	27	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 1.

The worst 5 of 27 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:S:130:GLY:HA3	9:S:490:HOH:O	1.89	0.71
2:M:313:ARG:HD3	2:M:322:TYR:CE2	2.34	0.63
1:S:141:GLU:HG3	9:S:401:HOH:O	2.04	0.56
1:S:202:ASP:OD2	1:T:202:ASP:OD2	2.23	0.56
2:L:313:ARG:HD3	2:L:322:TYR:CE2	2.42	0.54

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)
9:L:1143:HOH:O	9:M:892:HOH:O[3_544]	1.97	0.23
9:S:490:HOH:O	9:L:737:HOH:O[4_445]	2.04	0.16
9:L:1032:HOH:O	9:M:1093:HOH:O[3_544]	2.17	0.03

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	S	270/298 (91%)	261 (97%)	9 (3%)	0	100	100
1	T	272/298 (91%)	265 (97%)	7 (3%)	0	100	100
2	L	554/567 (98%)	531 (96%)	22 (4%)	1 (0%)	44	25
2	M	553/567 (98%)	530 (96%)	22 (4%)	1 (0%)	44	25
All	All	1649/1730 (95%)	1587 (96%)	60 (4%)	2 (0%)	48	28

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	211	LYS
2	M	211	LYS

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	S	217/239 (91%)	217 (100%)	0	100	100
1	T	219/239 (92%)	219 (100%)	0	100	100
2	L	471/479 (98%)	465 (99%)	6 (1%)	65	46
2	M	470/479 (98%)	465 (99%)	5 (1%)	70	53
All	All	1377/1436 (96%)	1366 (99%)	11 (1%)	81	66

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	242	TYR
2	M	312	TYR
2	M	414	GLU
2	M	382	ASN
2	L	382[A]	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 19 ligands modelled in this entry, 5 are monoatomic - leaving 14 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

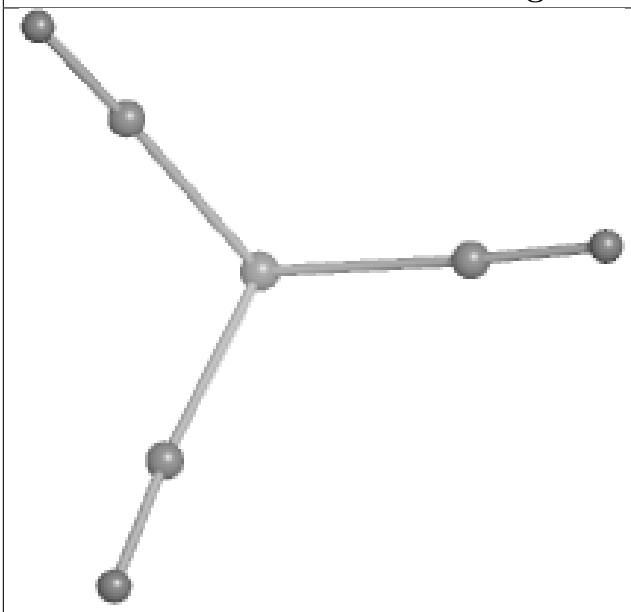
There are no torsion outliers.

There are no ring outliers.

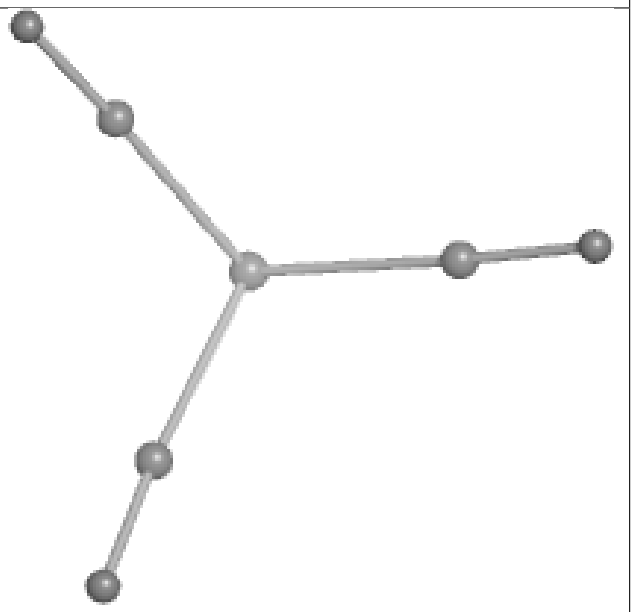
No monomer is involved in short contacts.

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.

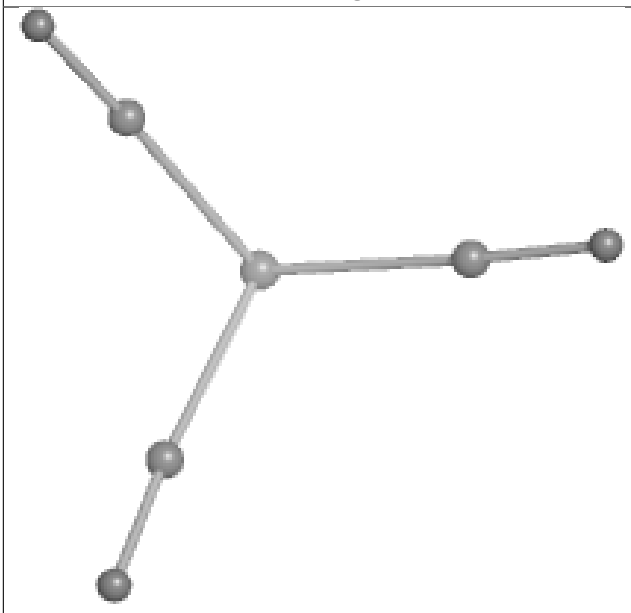
Ligand FCO L 601



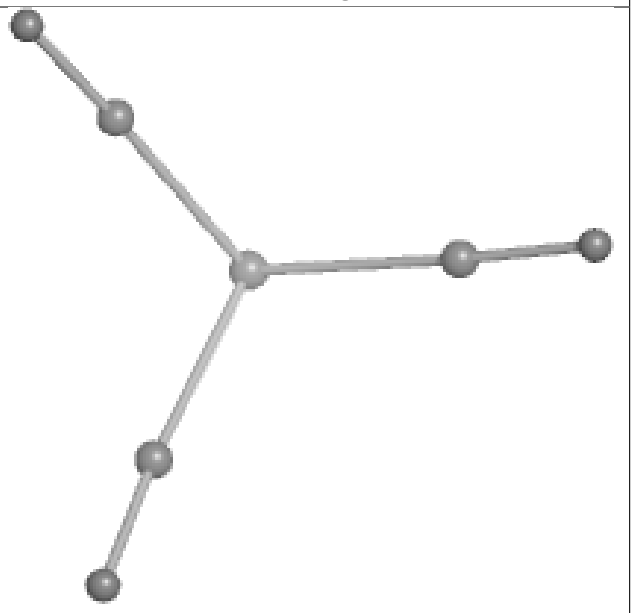
Bond lengths



Bond angles

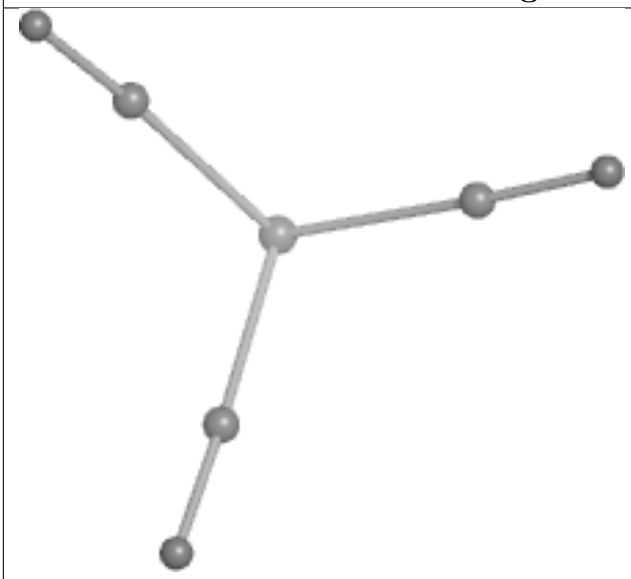


Torsions

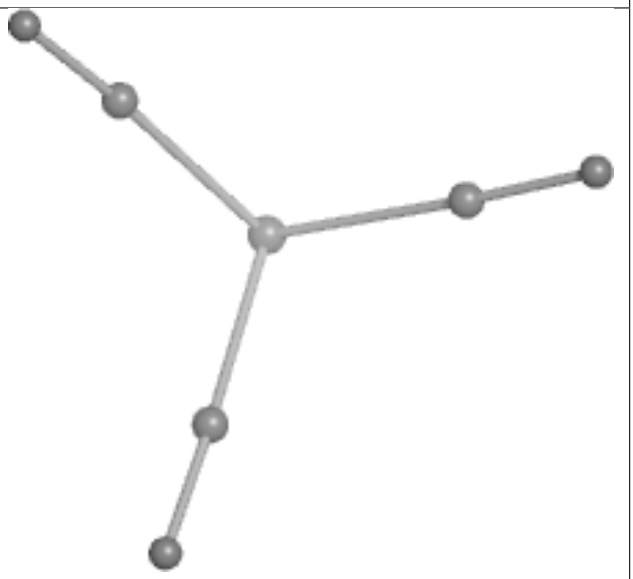


Rings

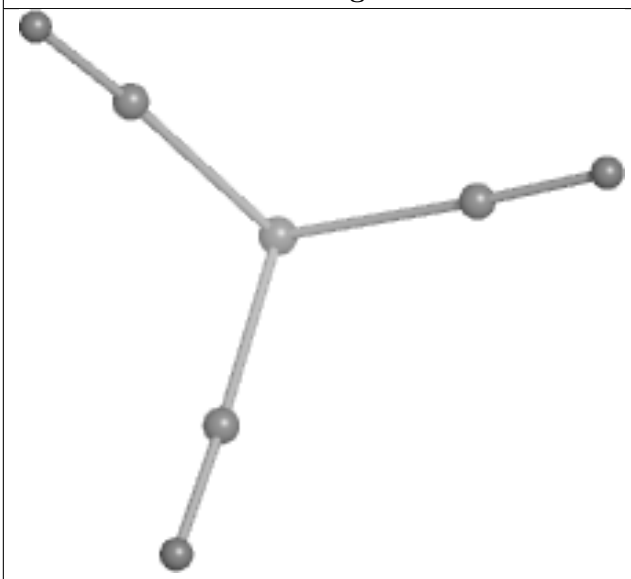
Ligand FCO M 602



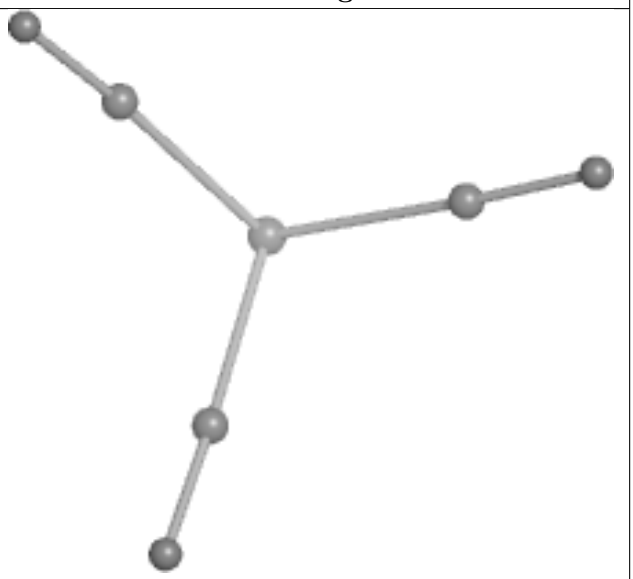
Bond lengths



Bond angles



Torsions



Rings

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, 95^{th} 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	S	268/298 (89%)	-0.16	3 (1%) 77 80	10, 21, 37, 58	4 (1%)
1	T	267/298 (89%)	-0.02	2 (0%) 84 87	11, 23, 39, 55	7 (2%)
2	L	551/567 (97%)	-0.24	0 100 100	8, 21, 33, 44	6 (1%)
2	M	551/567 (97%)	-0.00	6 (1%) 77 80	8, 23, 40, 63	5 (0%)
All	All	1637/1730 (94%)	-0.11	11 (0%) 84 87	8, 22, 38, 63	22 (1%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	M	27	VAL	3.2
1	S	9	PRO	3.1
1	S	66	HIS	2.9
2	M	137	VAL	2.9
2	M	515	VAL	2.7

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 monosaccharides in this entry.

6.4 Ligands

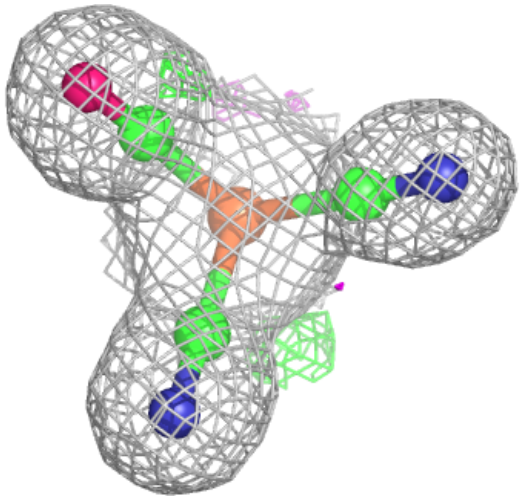
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

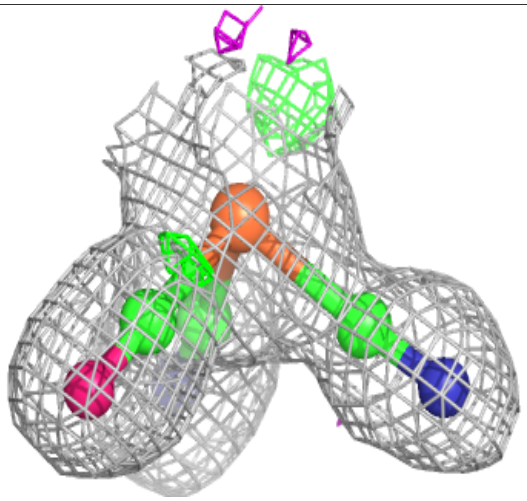
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	GOL	M	606	6/6	0.85	0.12	49,50,53,54	0
5	GOL	S	305	6/6	0.87	0.13	44,46,47,49	0
5	GOL	M	601	6/6	0.90	0.12	27,32,38,43	0
8	MG	M	607	1/1	0.92	0.10	74,74,74,74	0
5	GOL	M	605	6/6	0.93	0.10	25,37,40,43	0
5	GOL	S	304	6/6	0.94	0.08	24,25,28,28	0
5	GOL	T	304	6/6	0.97	0.06	24,26,27,29	0
6	FCO	M	602	7/7	0.99	0.05	14,17,19,19	0
3	SF4	T	301	8/8	1.00	0.02	18,20,21,21	0
3	SF4	T	303	8/8	1.00	0.03	18,19,20,21	0
4	F3S	S	302	7/7	1.00	0.03	16,16,17,18	0
4	F3S	T	302	7/7	1.00	0.02	16,17,18,18	0
6	FCO	L	601	7/7	1.00	0.04	15,16,17,18	0
3	SF4	S	301	8/8	1.00	0.02	18,19,21,21	0
7	NI	L	602	1/1	1.00	0.01	19,19,19,19	0
7	NI	M	603	1/1	1.00	0.02	21,21,21,21	0
8	MG	L	603	1/1	1.00	0.03	12,12,12,12	0
8	MG	M	604	1/1	1.00	0.04	15,15,15,15	0
3	SF4	S	303	8/8	1.00	0.03	16,17,19,19	0

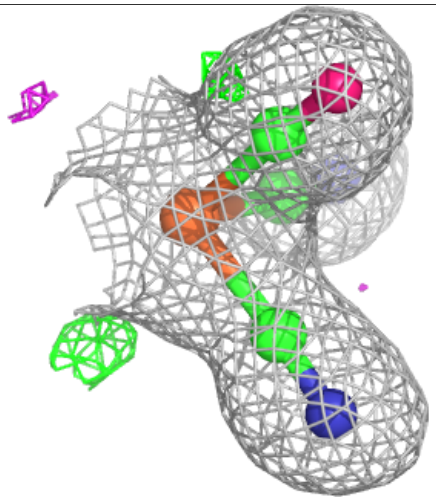
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around FCO M 602:

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

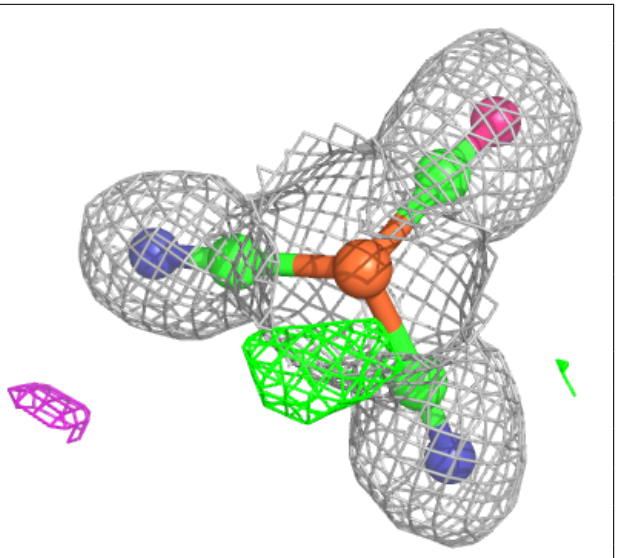


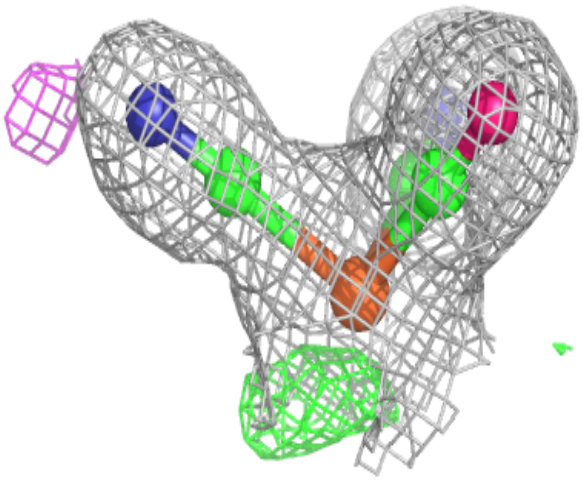


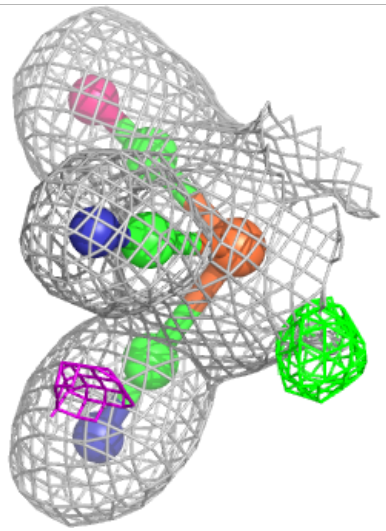


Electron density around FCO L 601:

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

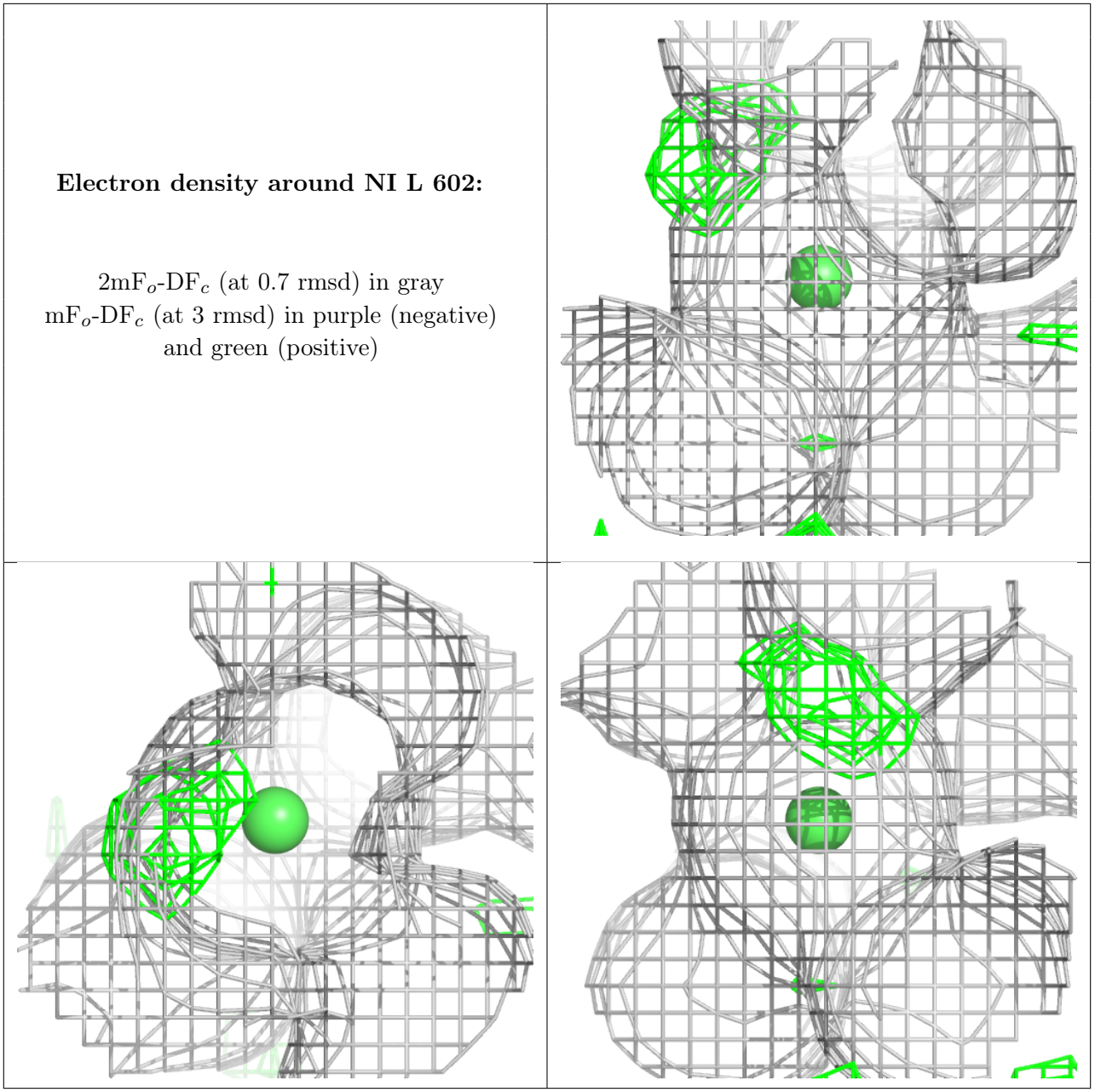






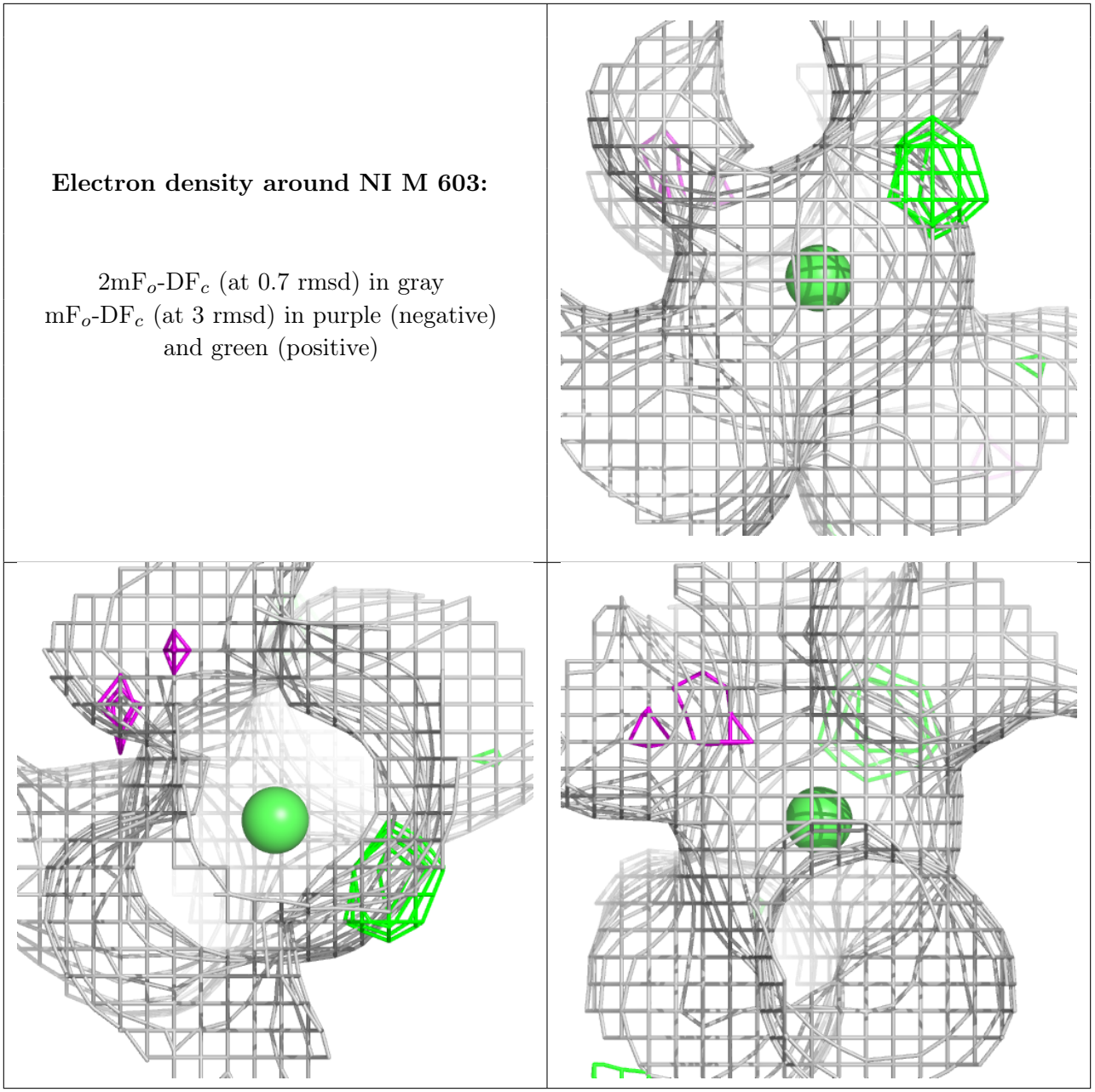
Electron density around NI L 602:

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



Electron density around NI M 603:

$2mF_o$-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.