



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2024 – 07:04 pm GMT

PDB ID : 8CJB
Title : A268M variant of the CODH/ACS complex of *C. hydrogenoformans*
Authors : Ruickoldt, J.; Jeoung, J.; Lennartz, F.; Dobbek, H.
Deposited on : 2023-02-13
Resolution : 2.50 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

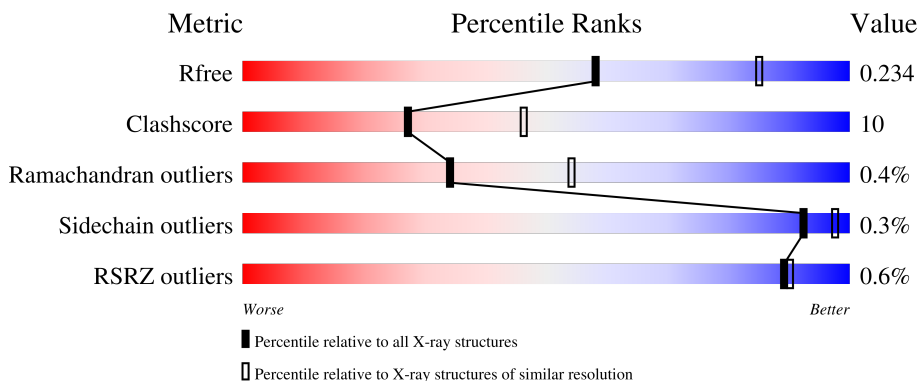
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	669	
2	B	731	

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
6	OH	A	705	-	-	X	-

2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 11267 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 Carbon monoxide dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	669	5147	3264	887	961	35	0	4	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	ASP	GLU	conflict	UNP A0A1L8D0M5
A	29	ILE	THR	conflict	UNP A0A1L8D0M5
A	73	GLN	MET	conflict	UNP A0A1L8D0M5
A	120	ALA	THR	conflict	UNP A0A1L8D0M5
A	153	THR	ILE	conflict	UNP A0A1L8D0M5
A	159	MET	LEU	conflict	UNP A0A1L8D0M5
A	199	GLU	ASP	conflict	UNP A0A1L8D0M5
A	205	SER	ALA	conflict	UNP A0A1L8D0M5
A	220	ILE	MET	conflict	UNP A0A1L8D0M5
A	389	ILE	VAL	conflict	UNP A0A1L8D0M5
A	393	LEU	PHE	conflict	UNP A0A1L8D0M5
A	494	THR	ALA	conflict	UNP A0A1L8D0M5
A	602	THR	SER	conflict	UNP A0A1L8D0M5

- Molecule 2 is a protein called CO-methylating acetyl-CoA synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	730	5837	3744	981	1082	30	8	8	0

There are 5 discrepancies between the modelled and reference sequences:

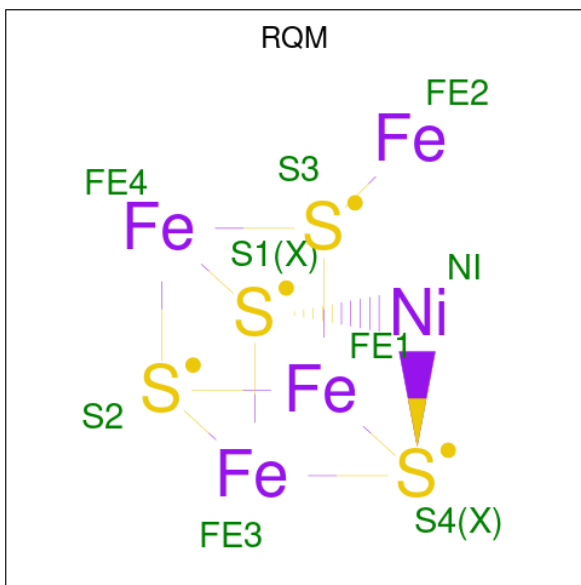
Chain	Residue	Modelled	Actual	Comment	Reference
B	225	LEU	ALA	conflict	UNP Q3ACS4
B	268	MET	ALA	engineered mutation	UNP Q3ACS4
B	733	ARG	-	expression tag	UNP Q3ACS4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	734	SER	-	expression tag	UNP Q3ACS4
B	735	HIS	-	expression tag	UNP Q3ACS4

- Molecule 3 is Fe(3)-Ni(1)-S(4) cluster (three-letter code: RQM) (formula: Fe_4NiS_4) (labeled as "Ligand of Interest" by depositor).



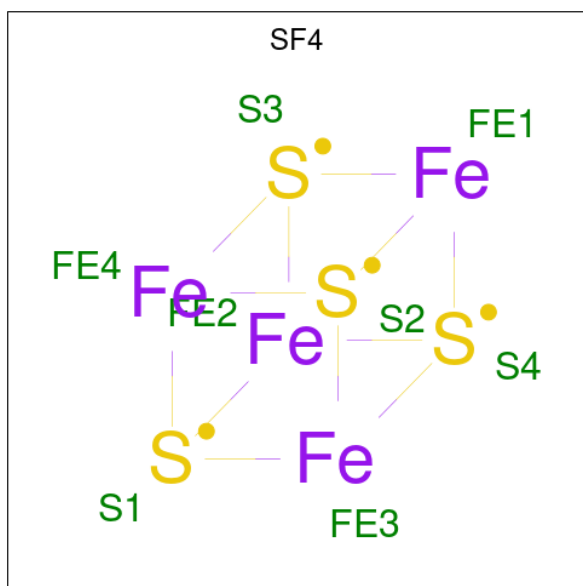
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	Fe	Ni	S		
3	A	1	9	4	1	4	0	0

- Molecule 4 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
4	A	1	4	2	2	0	0

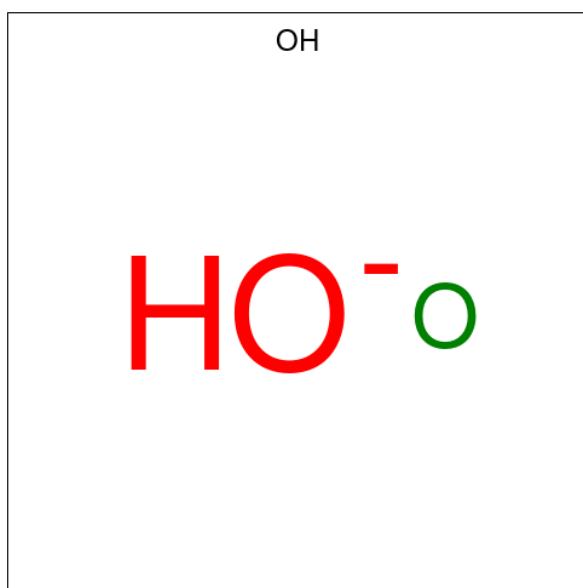
- Molecule 5 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
5	A	1	8	4	4	0	0
5	B	1	8	4	4	0	0

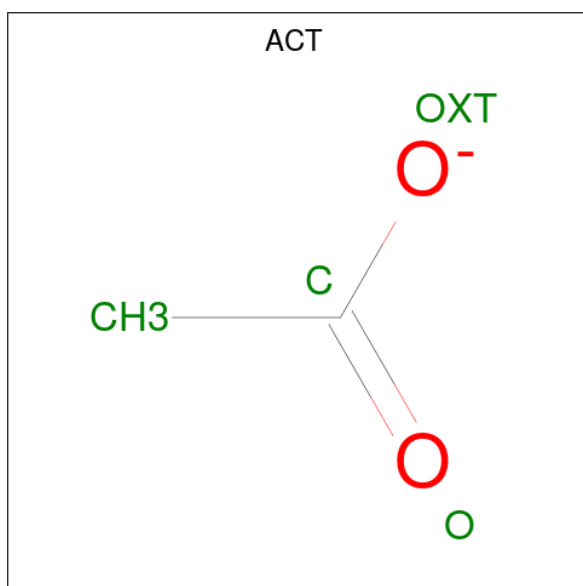
- Molecule 6 is HYDROXIDE ION (three-letter code: OH) (formula: HO) (labeled as "Ligand

of Interest" by depositor).



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

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



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

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

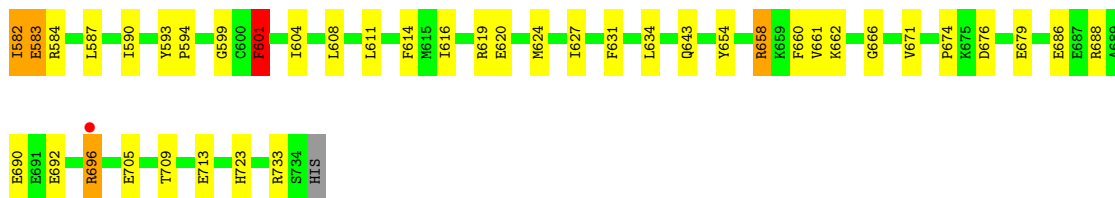
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Na	0	0
			1	1		

- Molecule 9 is NICKEL (II) ION (three-letter code: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	2	Total	Ni	0	1
			2	2		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	96	Total	O	0	0
			96	96		
10	B	149	Total	O	0	0
			149	149		



4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	142.12Å 142.12Å 290.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.02 – 2.50 48.37 – 2.31	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.02-2.50) 99.9 (48.37-2.31)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 2.32Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.206 , 0.235 0.206 , 0.234	Depositor DCC
R_{free} test set	1117 reflections (1.46%)	wwPDB-VP
Wilson B-factor (Å ²)	49.6	Xtrriage
Anisotropy	0.266	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 24.1	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	0.95	EDS
Total number of atoms	11267	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% 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: FES, OH, NA, RQM, SF4, ACT, 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	0.71	5/5259 (0.1%)	1.07	31/7118 (0.4%)
2	B	0.71	7/5995 (0.1%)	1.23	54/8111 (0.7%)
All	All	0.71	12/11254 (0.1%)	1.16	85/15229 (0.6%)

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	10
2	B	0	11
All	All	0	21

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	89	CYS	CB-SG	-17.25	1.52	1.82
1	A	184	TRP	CG-CD1	-12.67	1.19	1.36
2	B	384	ARG	CG-CD	11.32	1.80	1.51
2	B	384	ARG	CB-CG	-9.89	1.25	1.52
2	B	509	CYS	CB-SG	-6.53	1.71	1.82

The worst 5 of 85 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	322	ASP	CB-CG-OD1	33.81	148.73	118.30
2	B	601	PHE	CB-CG-CD2	-33.03	97.68	120.80
1	A	299	GLU	OE1-CD-OE2	-31.11	85.97	123.30
2	B	322	ASP	CB-CG-OD2	-24.26	96.47	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	384	ARG	CA-CB-CG	-21.73	65.59	113.40

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	148	ARG	Sidechain
1	A	198	ASP	Peptide
1	A	199	GLU	Sidechain
1	A	299	GLU	Sidechain
1	A	466	GLY	Peptide

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5147	0	5168	92	0
2	B	5837	0	5827	133	0
3	A	9	0	0	1	0
4	A	4	0	0	0	0
5	A	8	0	0	0	0
5	B	8	0	0	0	0
6	A	2	0	0	2	0
7	B	4	0	3	1	0
8	B	1	0	0	0	0
9	B	2	0	0	0	0
10	A	96	0	0	12	3
10	B	149	0	0	20	3
All	All	11267	0	10998	223	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 10.

The worst 5 of 223 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:B:384:ARG:CD	2:B:384:ARG:CG	1.80	1.60
2:B:384:ARG:CD	2:B:384:ARG:CB	2.18	1.22
2:B:384:ARG:CD	2:B:384:ARG:N	2.16	1.08
2:B:384:ARG:CG	2:B:384:ARG:NE	2.32	0.93
2:B:384:ARG:CD	2:B:384:ARG:HB2	2.00	0.91

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)
10:A:886:HOH:O	10:B:1040:HOH:O[5_554]	1.86	0.34
10:A:862:HOH:O	10:B:953:HOH:O[5_554]	1.88	0.32
10:A:858:HOH:O	10:B:939:HOH:O[5_554]	2.03	0.17

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	671/669 (100%)	638 (95%)	30 (4%)	3 (0%)	34 54
2	B	737/731 (101%)	717 (97%)	17 (2%)	3 (0%)	34 54
All	All	1408/1400 (101%)	1355 (96%)	47 (3%)	6 (0%)	34 54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	199	GLU
2	B	599	GLY
1	A	266	ASN
1	A	353	GLN
2	B	424	ALA

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	552/548 (101%)	551 (100%)	1 (0%)	93 98
2	B	622/615 (101%)	618 (99%)	4 (1%)	86 95
All	All	1174/1163 (101%)	1169 (100%)	5 (0%)	92 97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	572	PHE
2	B	80	LYS
2	B	384	ARG
2	B	401[A]	ASP
2	B	401[B]	ASP

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	280	HIS
2	B	467	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 10 ligands modelled in this entry, 2 are modelled with single atom and 3 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	RQM	A	701	6,1,10	0,12,12	-	-	-		
7	ACT	B	801	9	3,3,3	1.52	1 (33%)	3,3,3	1.40	0
5	SF4	B	804	2	0,12,12	-	-	-		
5	SF4	A	703	1	0,12,12	-	-	-		
4	FES	A	702	1	0,4,4	-	-	-		

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	RQM	A	701	6,1,10	-	-	0/4/4/4
5	SF4	A	703	1	-	-	0/6/5/5
4	FES	A	702	1	-	-	0/1/1/1
5	SF4	B	804	2	-	-	0/6/5/5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	801	ACT	CH3-C	2.45	1.59	1.49

There are no bond angle outliers.

There are no chirality outliers.

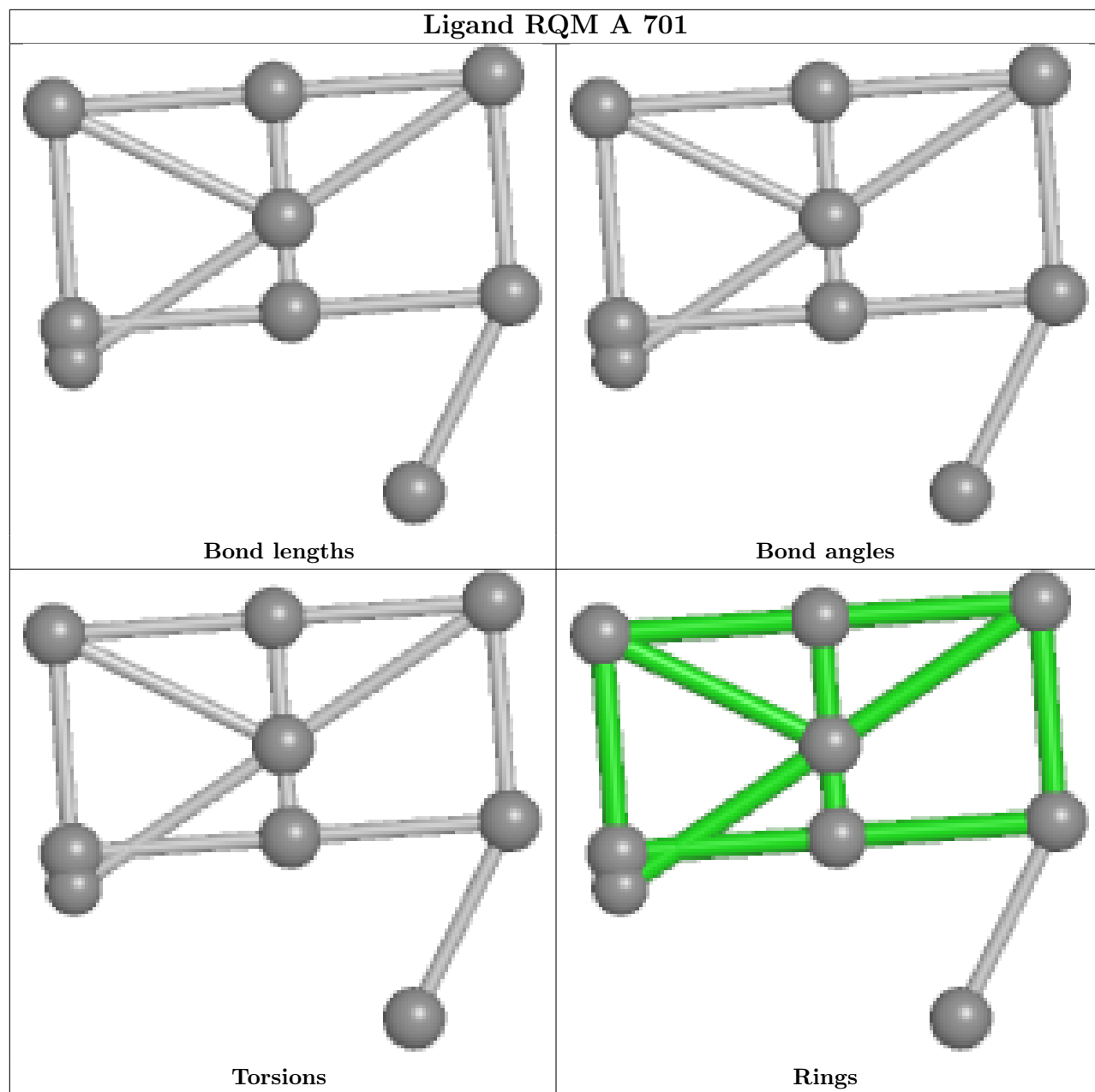
There are no torsion outliers.

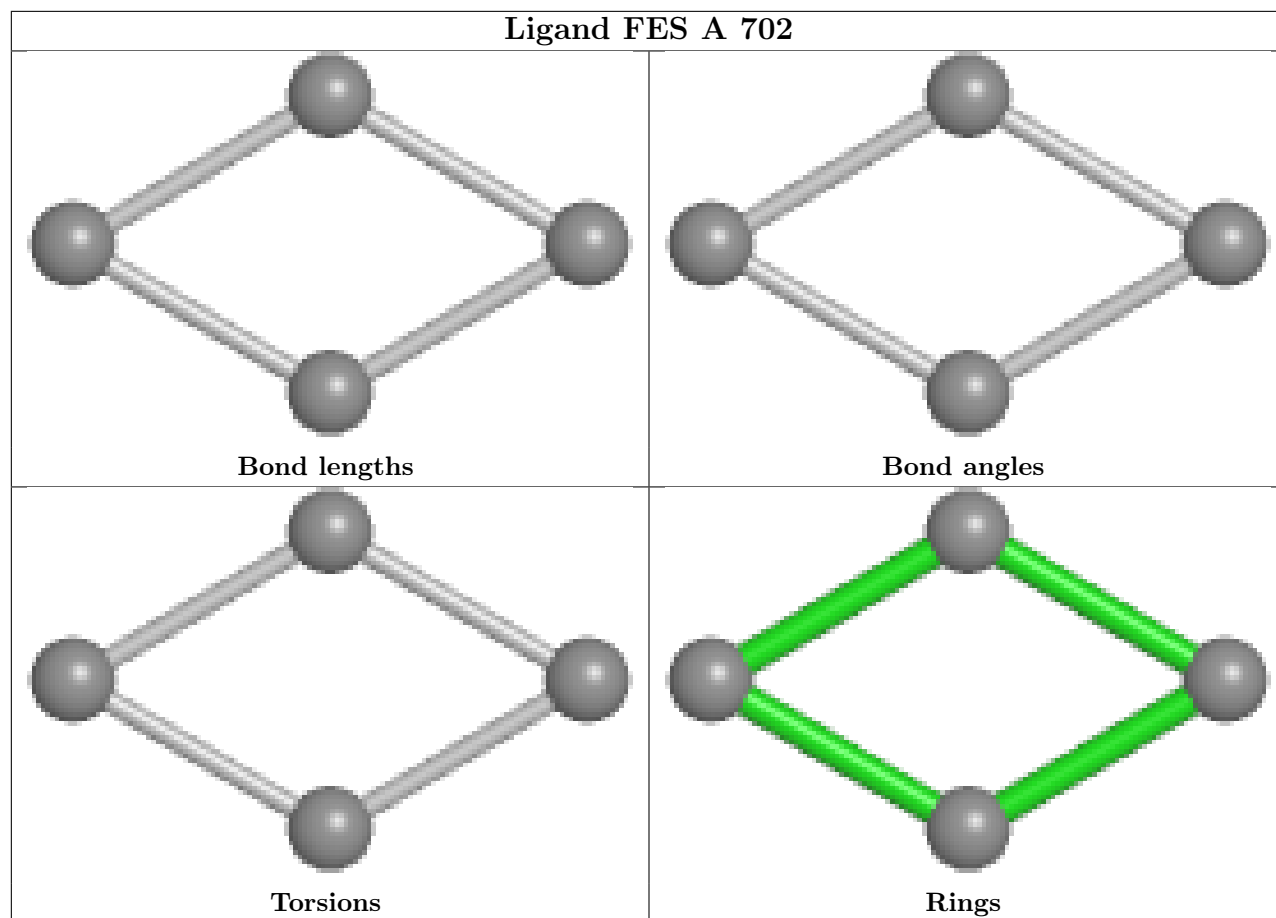
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	701	RQM	1	0
7	B	801	ACT	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	669/669 (100%)	-0.21	5 (0%) 87 89	39, 52, 71, 92	0
2	B	730/731 (99%)	-0.40	4 (0%) 91 91	41, 52, 65, 81	0
All	All	1399/1400 (99%)	-0.31	9 (0%) 89 90	39, 52, 68, 92	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	306	ALA	3.5
1	A	375[A]	ASN	2.8
2	B	580	ARG	2.6
1	A	415	VAL	2.3
1	A	407	PHE	2.2

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

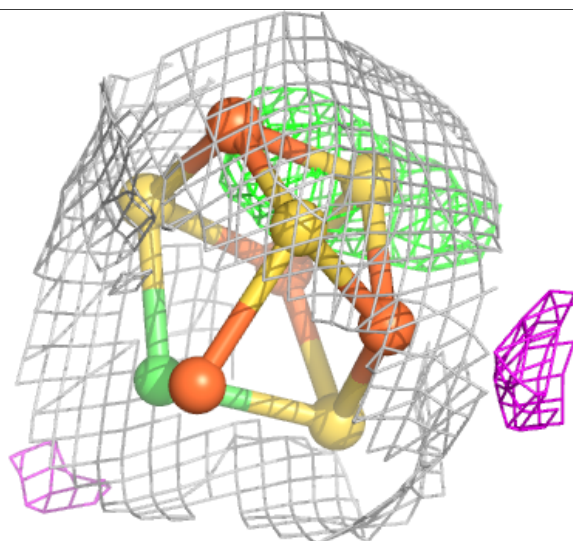
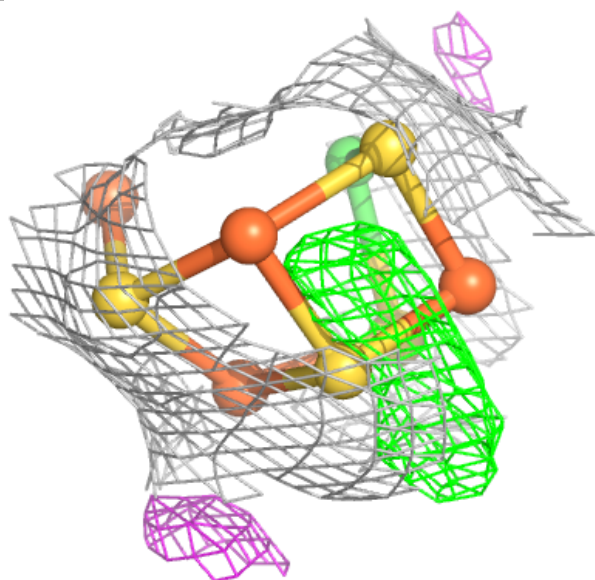
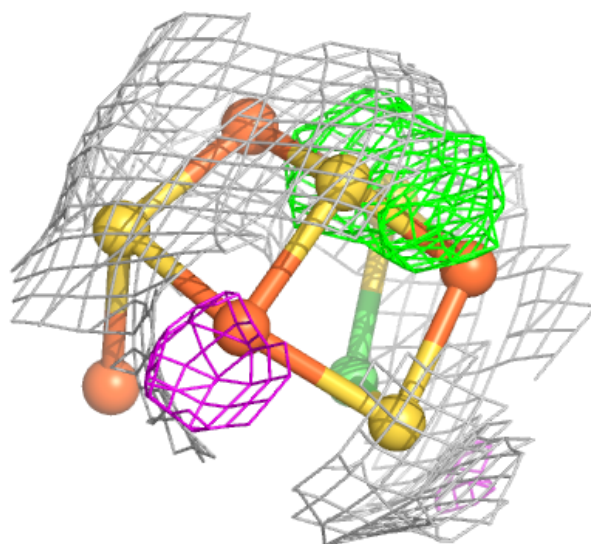
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	NA	B	802	1/1	0.92	0.05	53,53,53,53	0
7	ACT	B	801	4/4	0.95	0.17	51,51,51,51	0
3	RQM	A	701	9/9	0.95	0.16	63,75,76,79	7
6	OH	A	704	1/1	0.96	0.61	67,67,67,67	0
6	OH	A	705	1/1	0.97	0.13	56,56,56,56	0
5	SF4	B	804	8/8	0.97	0.09	48,50,51,52	0
4	FES	A	702	4/4	0.97	0.10	43,43,43,44	0
9	NI	B	805[A]	1/1	0.97	0.15	72,72,72,72	0
5	SF4	A	703	8/8	0.98	0.14	42,43,44,45	0
9	NI	B	803	1/1	0.99	0.13	57,57,57,57	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.

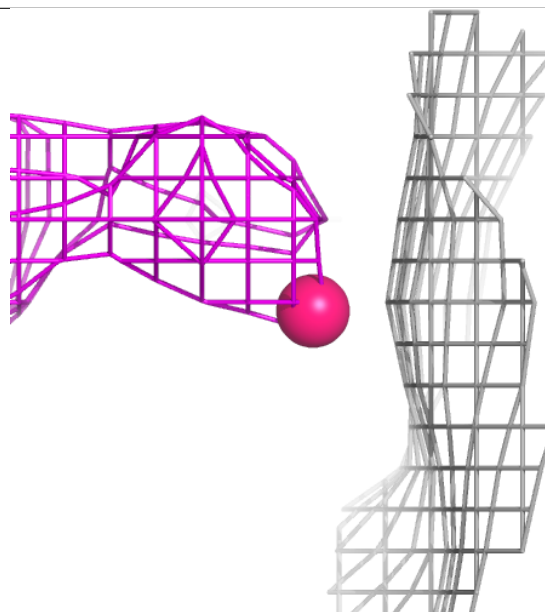
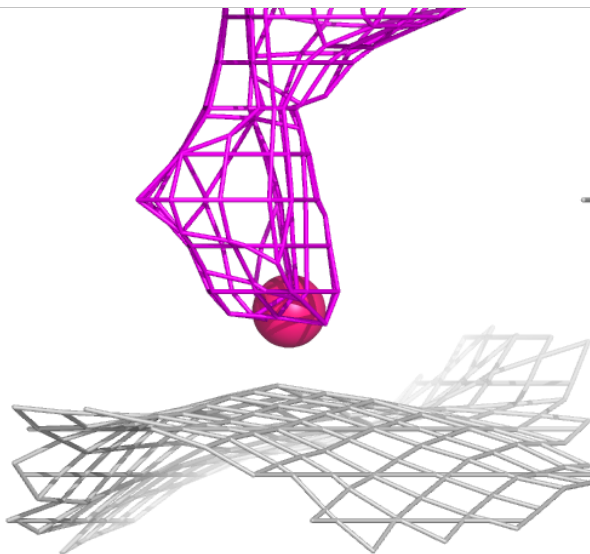
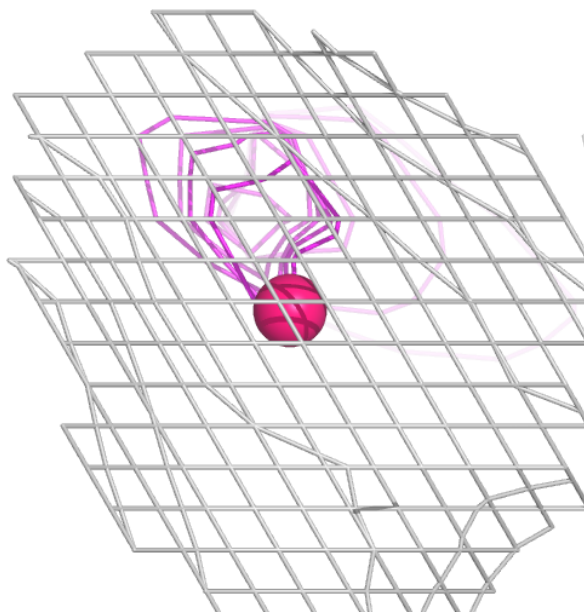
Electron density around RQM A 701:

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



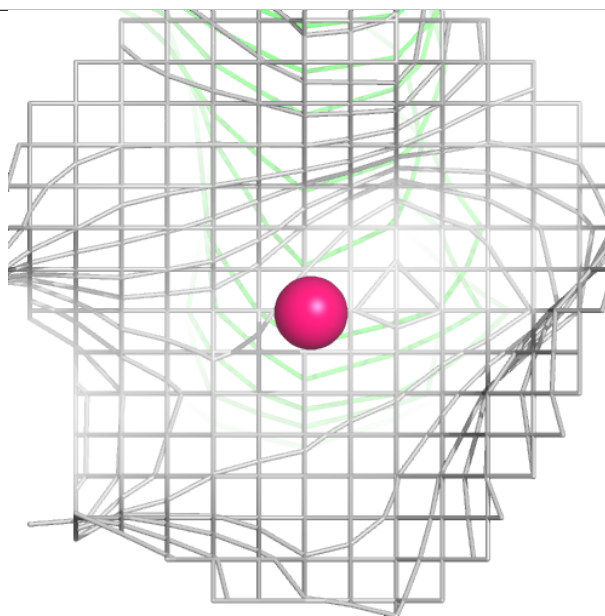
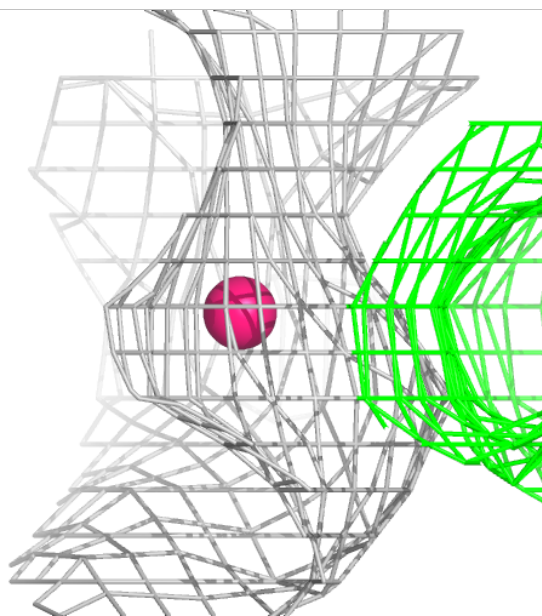
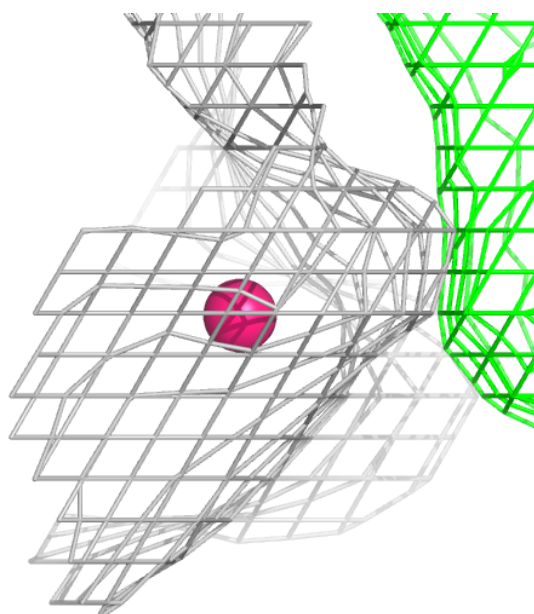
Electron density around OH A 704:

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



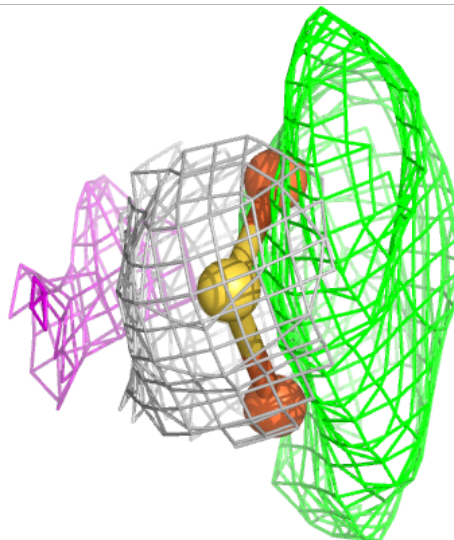
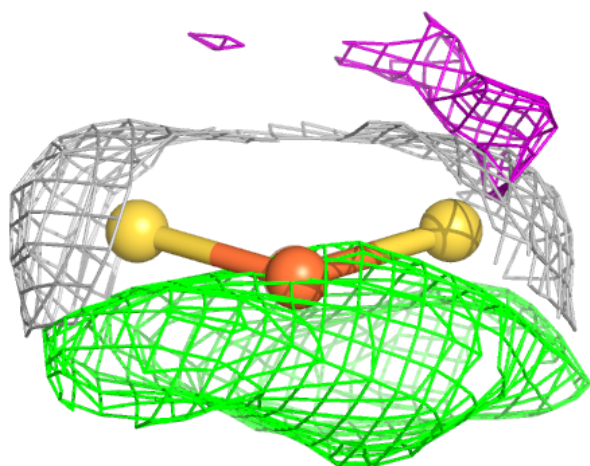
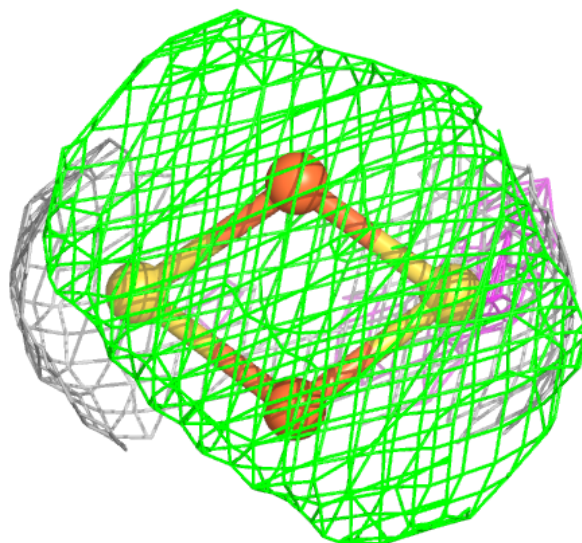
Electron density around OH A 705:

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



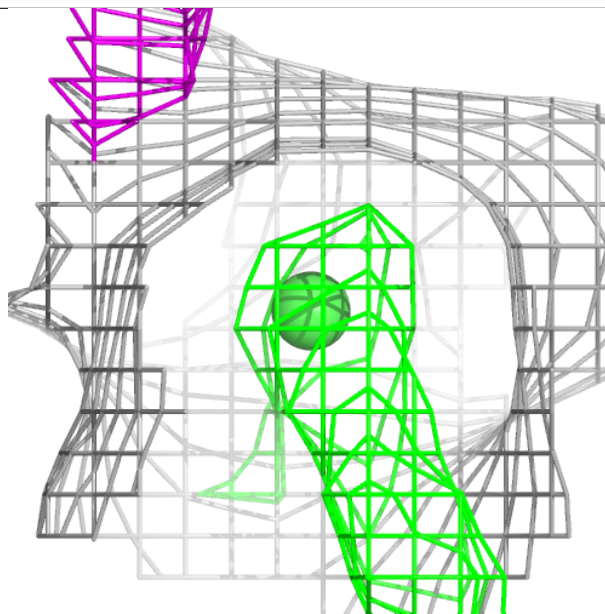
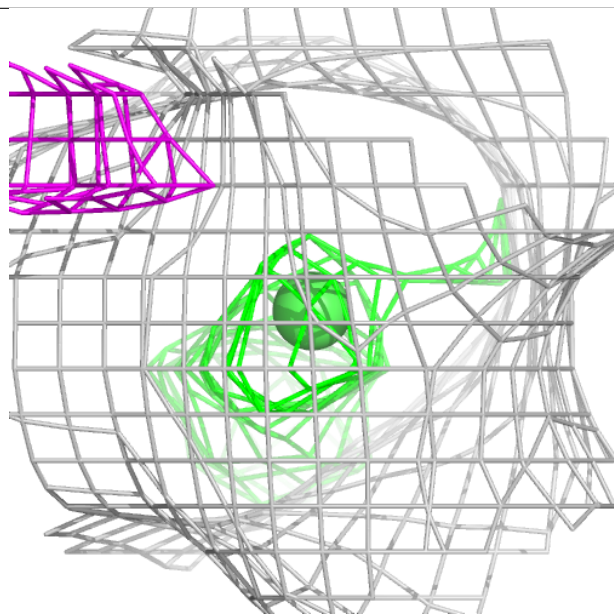
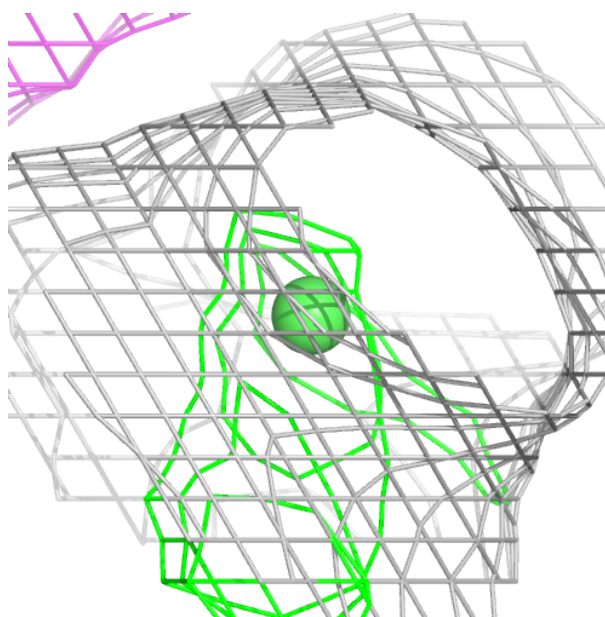
Electron density around FES A 702:

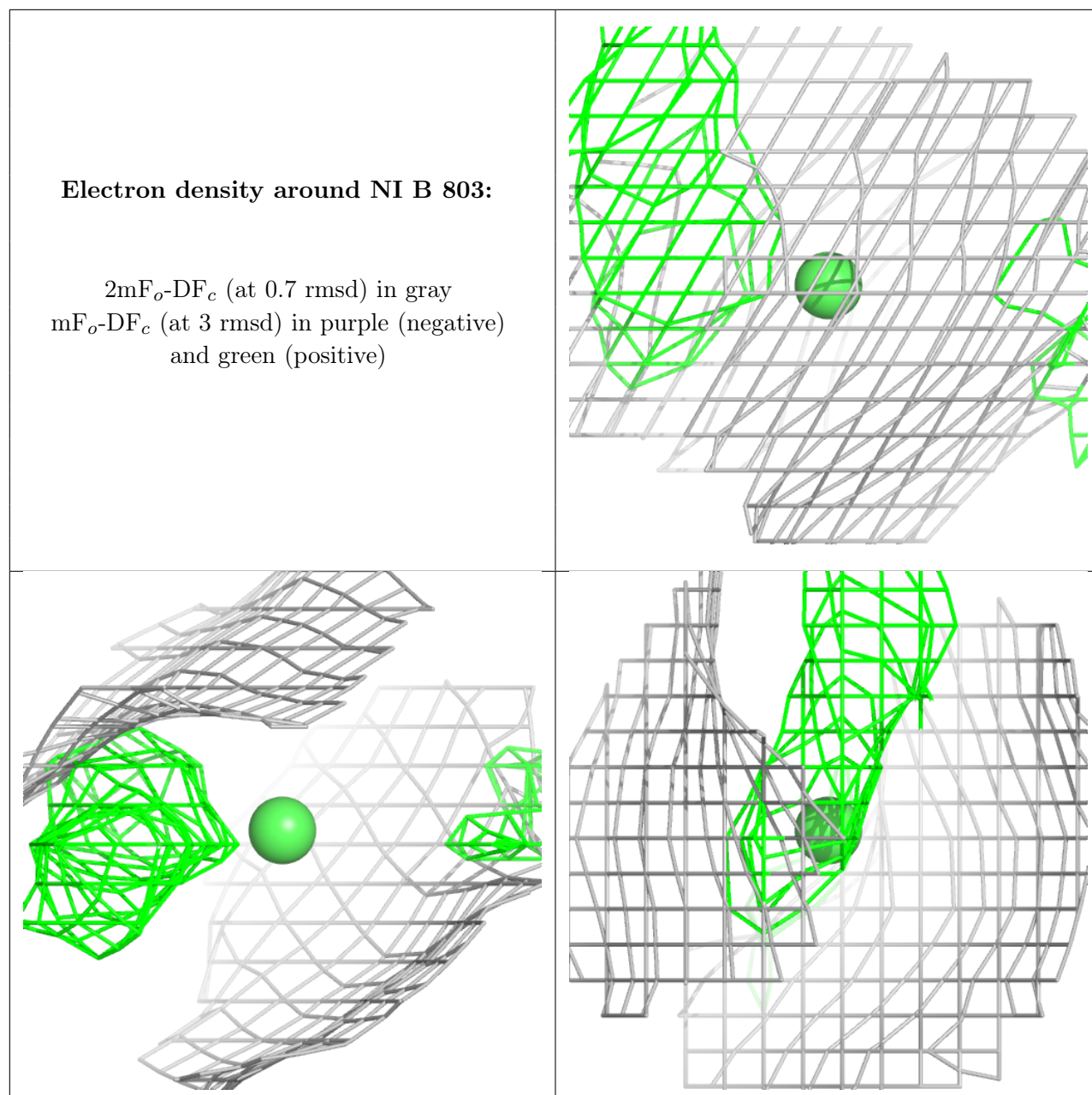
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around NI B 805 (A):

$2mF_o-DF_c$ (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.