



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 15, 2023 – 02:11 PM EDT

PDB ID : 7TXY
Title : Crystal structure of the 2-Aminophenol 1,6-dioxygenase from the ARO bacterial microcompartment of *Micromonospora rosaria*
Authors : Sutter, M.; Doron, L.; Kerfeld, C.A.
Deposited on : 2022-02-10
Resolution : 1.75 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
Xtriage (Phenix) : 1.13
EDS : 2.35
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) : 2.35

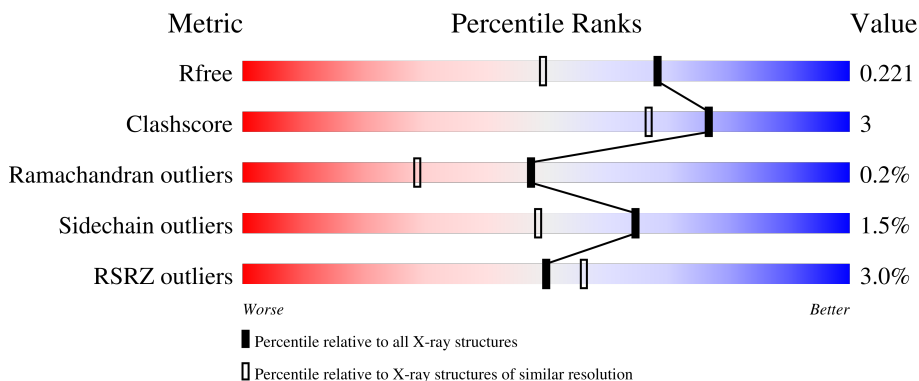
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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.



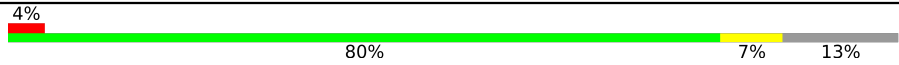

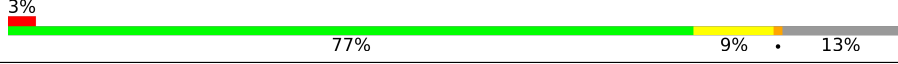
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	279	 0% 91% 7% .
1	C	279	 4% 91% 8% .
1	E	279	 3% 91% 8% .
1	G	279	 2% 91% 8% .
2	B	337	 0% 79% 8% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	337	 4% 80% 7% 13%
2	F	337	 4% 79% 8% 13%
2	H	337	 3% 77% 9% 13%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 36411 atoms, of which 17251 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 2-amino-5-chlorophenol 1,6-dioxygenase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	275	Total 4202	C 1348	H 2060	N 392	O 397	S 5	0	4	0
1	C	274	Total 4137	C 1327	H 2025	N 385	O 395	S 5	0	1	0
1	E	276	Total 4169	C 1338	H 2038	N 389	O 399	S 5	0	1	0
1	G	275	Total 4147	C 1330	H 2030	N 388	O 394	S 5	0	0	0

- Molecule 2 is a protein called 2-aminophenol 1,6-dioxygenase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	B	292	Total 4574	C 1471	H 2257	N 418	O 419	S 9	0	1	0
2	D	292	Total 4610	C 1485	H 2272	N 418	O 426	S 9	0	4	0
2	F	293	Total 4644	C 1490	H 2296	N 428	O 421	S 9	0	4	0
2	H	292	Total 4608	C 1482	H 2273	N 421	O 424	S 8	0	3	0

- Molecule 3 is FE (II) ION (three-letter code: FE2) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total 1 Fe 1	0	0
3	D	1	Total 1 Fe 1	0	0
3	F	1	Total 1 Fe 1	0	0
3	H	1	Total 1 Fe 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	155	Total 155	O 155	0	0
4	B	209	Total 209	O 209	0	0
4	C	131	Total 131	O 131	0	0
4	D	189	Total 189	O 189	0	0
4	E	140	Total 140	O 140	0	0
4	F	167	Total 167	O 167	0	0
4	G	138	Total 138	O 138	0	0
4	H	187	Total 187	O 187	0	0

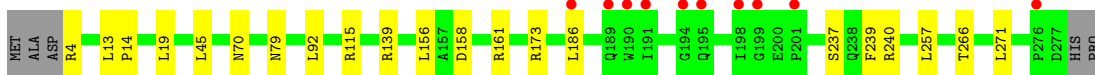
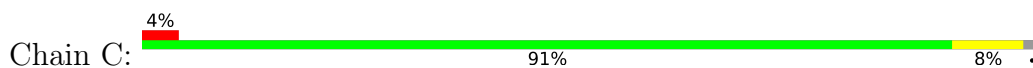
3 Residue-property plots [i](#)

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: 2-amino-5-chlorophenol 1,6-dioxygenase subunit alpha



- Molecule 1: 2-amino-5-chlorophenol 1,6-dioxygenase subunit alpha



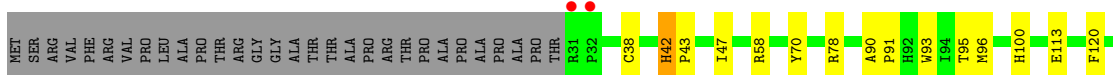
- Molecule 1: 2-amino-5-chlorophenol 1,6-dioxygenase subunit alpha



- Molecule 1: 2-amino-5-chlorophenol 1,6-dioxygenase subunit alpha

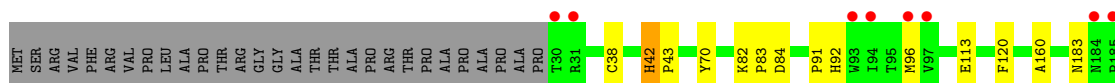
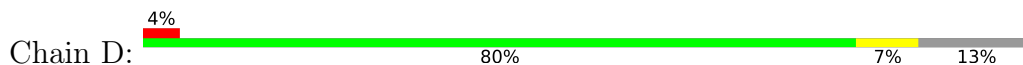


- Molecule 2: 2-aminophenol 1,6-dioxygenase subunit beta

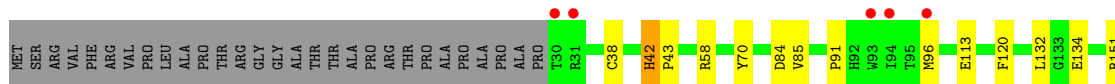
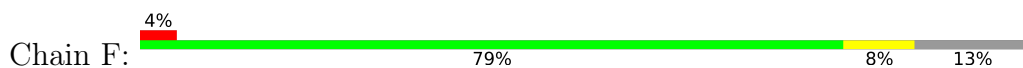




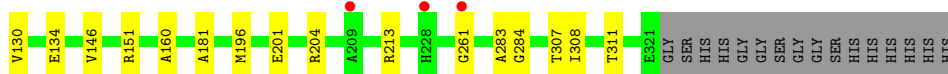
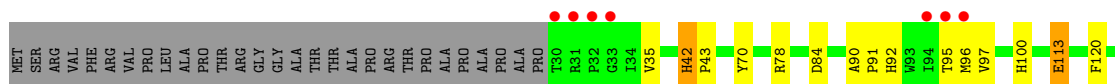
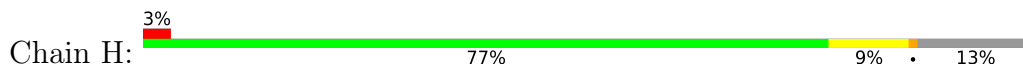
- Molecule 2: 2-aminophenol 1,6-dioxygenase subunit beta



- Molecule 2: 2-aminophenol 1,6-dioxygenase subunit beta



- Molecule 2: 2-aminophenol 1,6-dioxygenase subunit beta



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	68.80Å 83.86Å 110.27Å 85.59° 73.10° 89.26°	Depositor
Resolution (Å)	46.14 – 1.75 46.14 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.2 (46.14-1.75) 97.2 (46.14-1.75)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.28 (at 1.75Å)	Xtrriage
Refinement program	PHENIX 1.20rc1_4392	Depositor
R, R_{free}	0.194 , 0.222 0.194 , 0.221	Depositor DCC
R_{free} test set	1984 reflections (0.86%)	wwPDB-VP
Wilson B-factor (Å ²)	31.8	Xtrriage
Anisotropy	0.438	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 42.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.014 for h,-k,h-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	36411	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% 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: FE2

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	0.28	0/2208	0.55	0/3015
1	C	0.27	0/2168	0.54	0/2962
1	E	0.27	0/2188	0.54	0/2989
1	G	0.27	0/2171	0.54	0/2966
2	B	0.28	0/2391	0.55	0/3271
2	D	0.27	0/2421	0.54	0/3313
2	F	0.28	0/2431	0.55	0/3323
2	H	0.27	0/2415	0.54	0/3304
All	All	0.27	0/18393	0.54	0/25143

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2142	2060	2068	13	0
1	C	2112	2025	2027	11	2
1	E	2131	2038	2040	13	0
1	G	2117	2030	2030	12	1
2	B	2317	2257	2258	19	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	2338	2272	2280	12	1
2	F	2348	2296	2304	16	0
2	H	2335	2273	2278	23	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
4	A	155	0	0	2	0
4	B	209	0	0	0	0
4	C	131	0	0	1	0
4	D	189	0	0	1	0
4	E	140	0	0	3	0
4	F	167	0	0	1	0
4	G	138	0	0	0	0
4	H	187	0	0	1	0
All	All	19160	17251	17285	105	2

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 105 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:G:186:LEU:HD13	2:H:96:MET:HB2	1.75	0.68
1:C:79:ASN:OD1	4:C:301:HOH:O	2.12	0.67
1:A:104:ASP:OD1	1:A:115:ARG:NH2	2.31	0.64
1:G:105:ARG:NH1	1:G:169:ASP:OD2	2.32	0.63
1:E:61:LEU:HD21	2:F:230:GLU:HB2	1.84	0.59

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:139:ARG:NH2	1:G:46:GLU:OE1[1_445]	2.13	0.07
1:C:4:ARG:NE	2:D:84:ASP:OD1[1_455]	2.18	0.02

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	277/279 (99%)	268 (97%)	9 (3%)	0	100	100
1	C	273/279 (98%)	266 (97%)	7 (3%)	0	100	100
1	E	275/279 (99%)	267 (97%)	8 (3%)	0	100	100
1	G	273/279 (98%)	264 (97%)	9 (3%)	0	100	100
2	B	291/337 (86%)	283 (97%)	7 (2%)	1 (0%)	41	22
2	D	294/337 (87%)	279 (95%)	14 (5%)	1 (0%)	41	22
2	F	295/337 (88%)	286 (97%)	8 (3%)	1 (0%)	41	22
2	H	293/337 (87%)	283 (97%)	9 (3%)	1 (0%)	41	22
All	All	2271/2464 (92%)	2196 (97%)	71 (3%)	4 (0%)	47	29

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	42	HIS
2	D	42	HIS
2	F	42	HIS
2	H	42	HIS

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	212/211 (100%)	206 (97%)	6 (3%)	43	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	208/211 (99%)	205 (99%)	3 (1%)	67	52
1	E	210/211 (100%)	205 (98%)	5 (2%)	49	26
1	G	208/211 (99%)	203 (98%)	5 (2%)	49	26
2	B	244/276 (88%)	243 (100%)	1 (0%)	91	87
2	D	248/276 (90%)	246 (99%)	2 (1%)	81	72
2	F	248/276 (90%)	244 (98%)	4 (2%)	62	45
2	H	247/276 (90%)	244 (99%)	3 (1%)	71	56
All	All	1825/1948 (94%)	1796 (98%)	29 (2%)	65	45

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	197	ARG
2	H	78	ARG
2	F	58	ARG
1	G	239	PHE
1	E	239	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	188	GLN
1	A	205	GLN
2	B	228	HIS
1	E	256	HIS

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 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.

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	275/279 (98%)	-0.07	4 (1%) 73 80	27, 39, 57, 86	0
1	C	274/279 (98%)	0.07	10 (3%) 42 49	27, 39, 62, 78	0
1	E	276/279 (98%)	0.07	8 (2%) 51 57	27, 41, 62, 93	0
1	G	275/279 (98%)	0.03	5 (1%) 68 76	28, 40, 63, 79	0
2	B	292/337 (86%)	0.05	5 (1%) 70 77	26, 35, 49, 91	0
2	D	292/337 (86%)	0.18	12 (4%) 37 44	28, 39, 60, 86	0
2	F	293/337 (86%)	0.29	13 (4%) 34 40	25, 38, 68, 103	0
2	H	292/337 (86%)	0.20	10 (3%) 45 51	25, 37, 58, 93	0
All	All	2269/2464 (92%)	0.11	67 (2%) 50 56	25, 38, 60, 103	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	30	THR	9.4
2	H	30	THR	8.2
2	F	322	GLY	6.7
2	F	187	PHE	6.5
2	B	31	ARG	5.5

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands

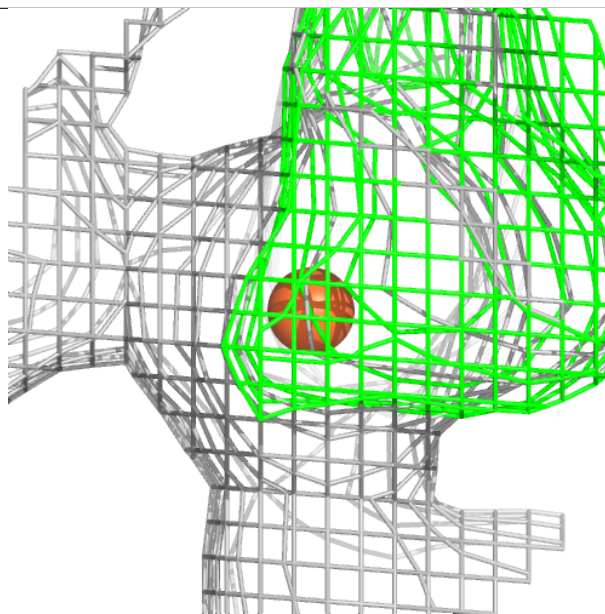
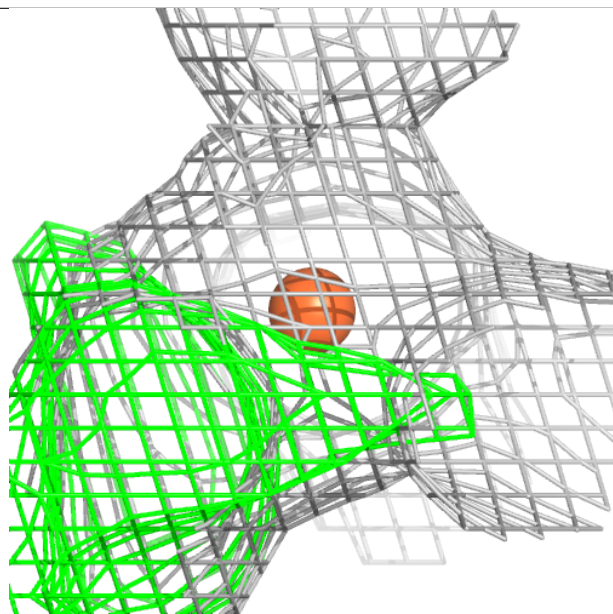
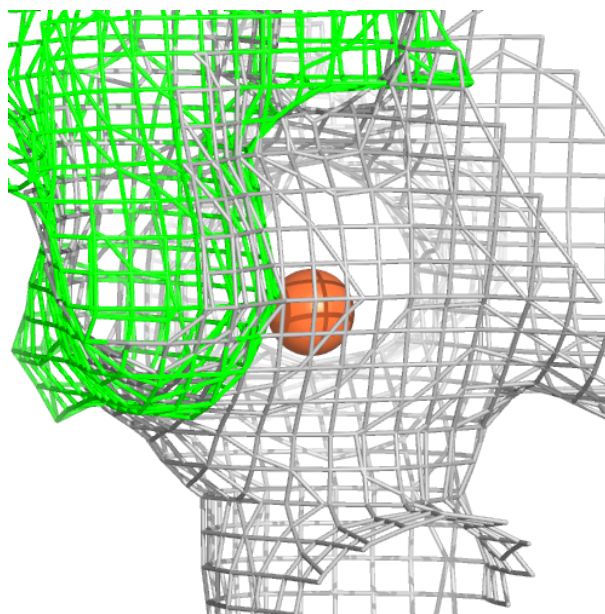
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	FE2	H	400	1/1	0.96	0.12	38,38,38,38	0
3	FE2	D	400	1/1	0.98	0.12	40,40,40,40	1
3	FE2	B	400	1/1	0.98	0.10	27,27,27,27	0
3	FE2	F	400	1/1	0.99	0.07	37,37,37,37	1

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.

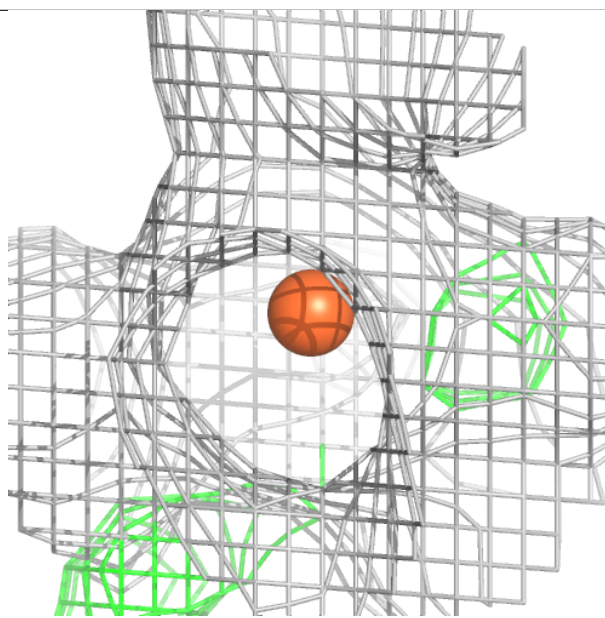
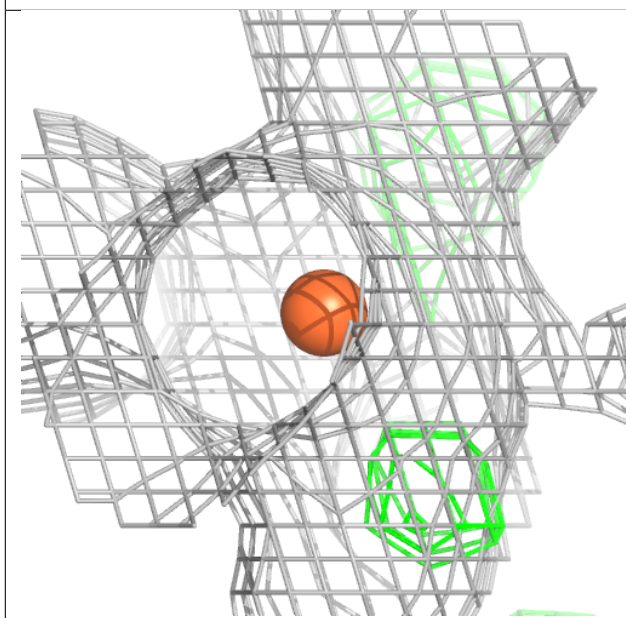
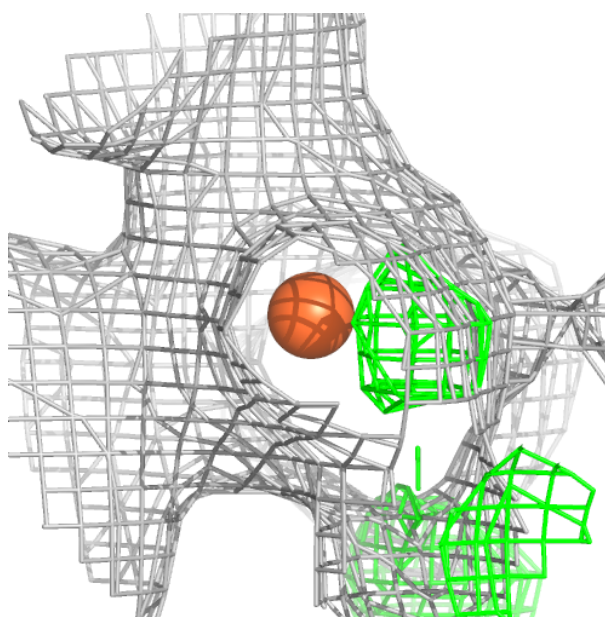
Electron density around FE2 H 400:

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



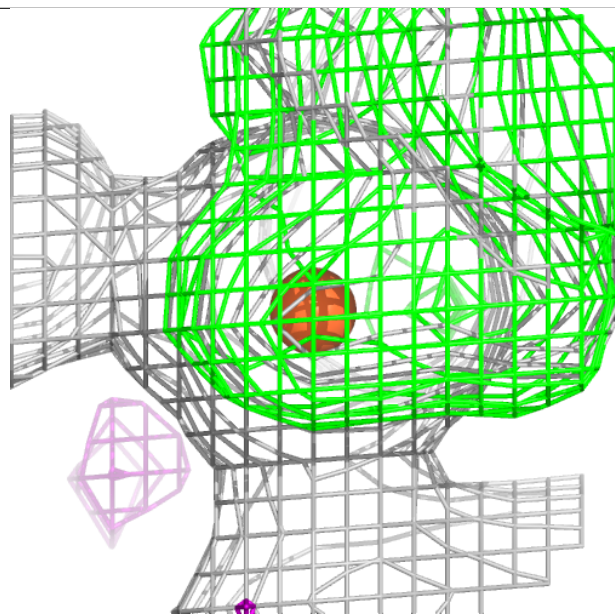
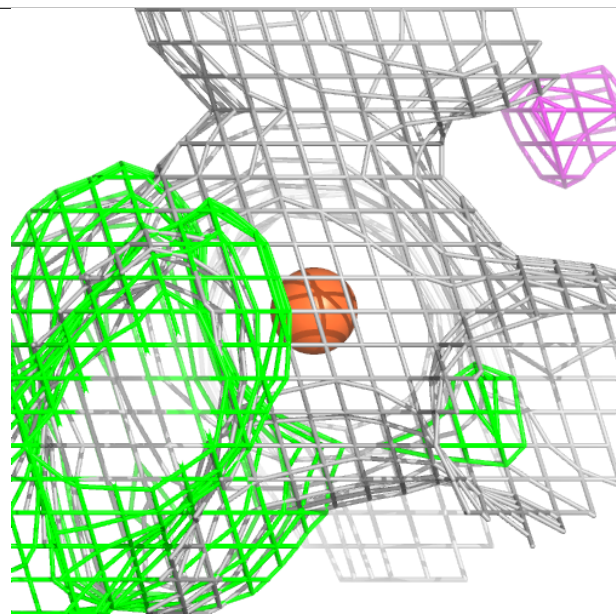
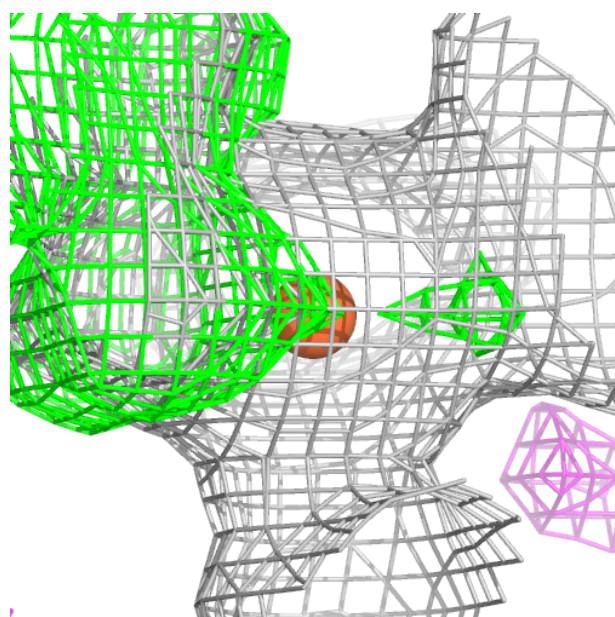
Electron density around FE2 D 400:

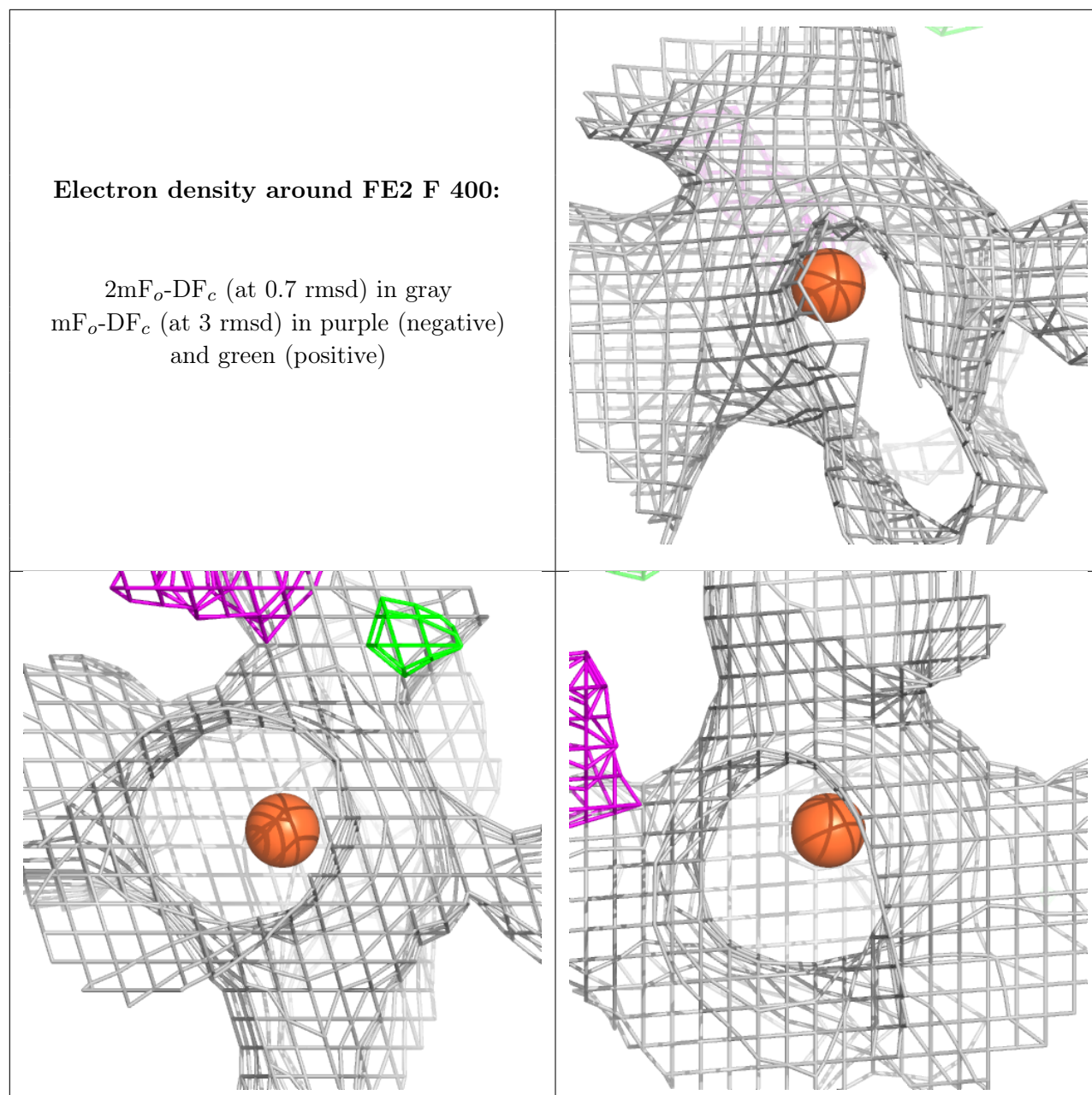
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around FE2 B 400:

$2mF_o-DF_c$ (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.