Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2020 – 10:07 PM BST

PDB ID	:	6YDI
Title	:	XFEL structure of the Soluble methane monooxygenase hydroxylase and regulatory subunit complex, from Methylosinus trichosporium OB3b, diferrous state
Authors	:	Srinivas, V.; Hogbom, M.
Deposited on	:	2020-03-20
Resolution	:	1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ⓘ) were used in the production of this report:

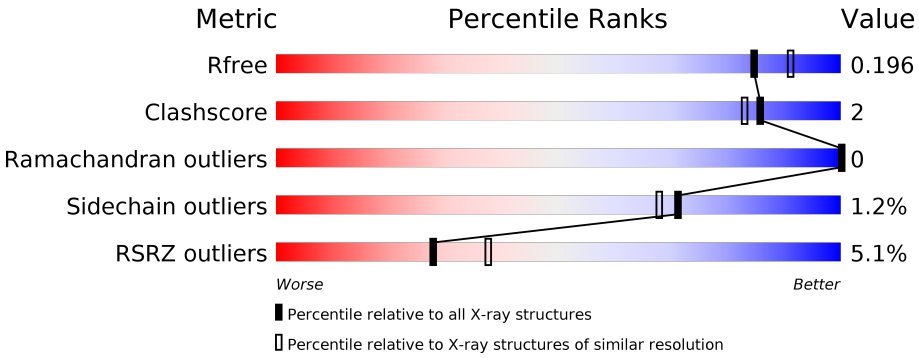
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 on 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	B	395	4% 92% 6%
2	D	526	4% 90% 6%
3	F	169	6% 96%
4	G	138	13% 95%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10205 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 Methane monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	389	Total	C	N	O	S	0	2	0
			3172	2025	554	588	5			

- Molecule 2 is a protein called Methane monooxygenase component A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	511	Total	C	N	O	S	0	1	0
			4149	2661	718	758	12			

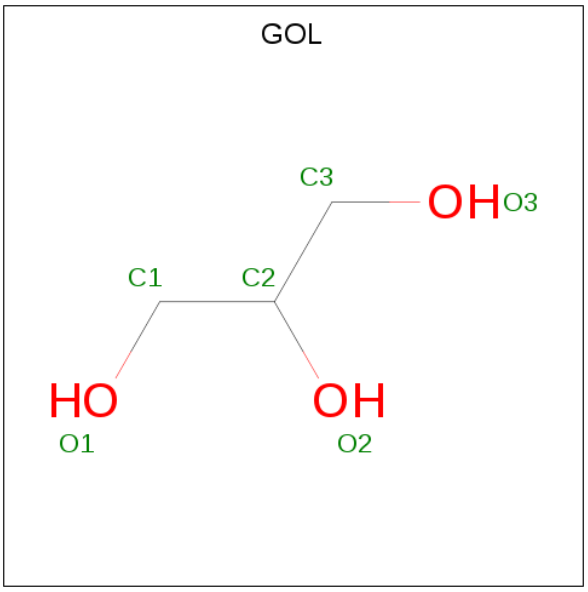
- Molecule 3 is a protein called Methane monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	168	Total	C	N	O	S	0	2	0
			1375	882	237	255	1			

- Molecule 4 is a protein called Methane monooxygenase regulatory protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	136	Total	C	N	O	S	0	0	0
			1032	659	166	204	3			

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 6 3 3	0	0

- Molecule 6 is FE (II) ION (three-letter code: FE2) (formula: Fe) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	2	Total Fe 2 2	0	0

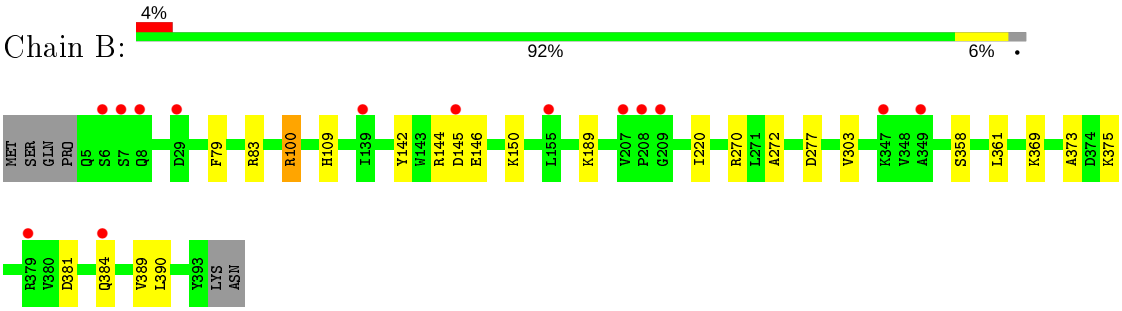
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	148	Total O 148 148	0	0
7	D	221	Total O 221 221	0	0
7	F	68	Total O 68 68	0	0
7	G	32	Total O 32 32	0	0

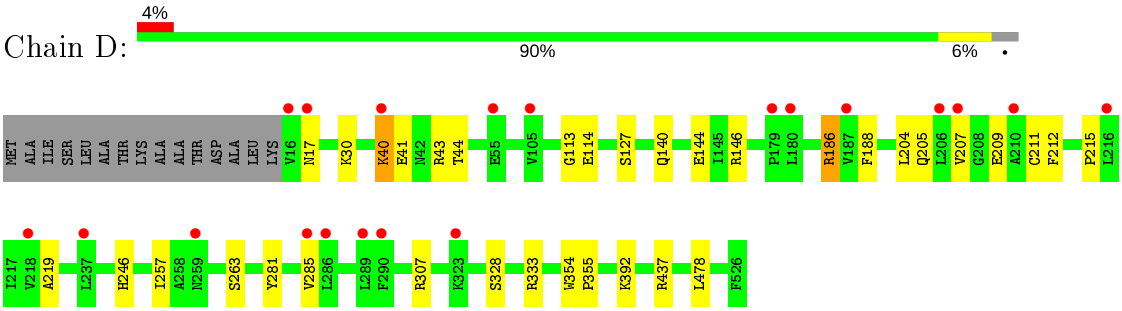
3 Residue-property plots (i)

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.

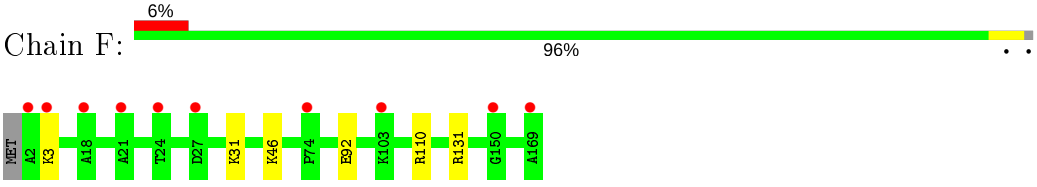
- Molecule 1: Methane monooxygenase



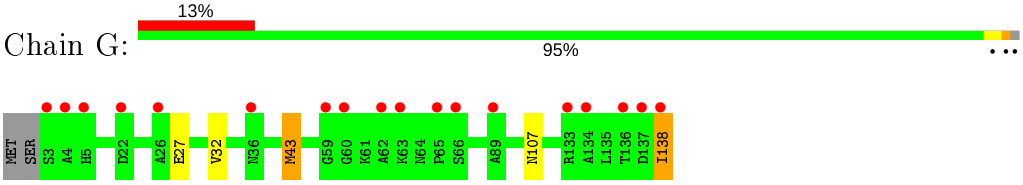
- Molecule 2: Methane monooxygenase component A alpha chain



- Molecule 3: Methane monooxygenase



- Molecule 4: Methane monooxygenase regulatory protein B



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	106.95Å 106.95Å 303.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.71 – 1.95 33.71 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (33.71-1.95) 90.5 (33.71-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.69 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.175 , 0.197 0.175 , 0.196	Depositor DCC
R_{free} test set	2000 reflections (1.55%)	wwPDB-VP
Wilson B-factor (Å^2)	36.5	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 57.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	10205	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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: GOL, 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	B	0.29	0/3270	0.46	0/4446
2	D	0.34	1/4282 (0.0%)	0.49	0/5820
3	F	0.31	0/1407	0.47	0/1902
4	G	0.29	0/1048	0.49	0/1418
All	All	0.31	1/10007 (0.0%)	0.48	0/13586

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	114	GLU	CD-OE1	-5.29	1.19	1.25

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3172	0	3016	13	0
2	D	4149	0	3948	25	0
3	F	1375	0	1417	3	0
4	G	1032	0	1028	6	0
5	B	6	0	8	0	0
6	D	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	148	0	0	0	0
7	D	221	0	0	3	0
7	F	68	0	0	1	0
7	G	32	0	0	0	0
All	All	10205	0	9417	42	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 2.

All (42) 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:D:328:SER:HB2	4:G:32:VAL:HG22	1.70	0.73
2:D:43:ARG:HG2	2:D:43:ARG:HH11	1.57	0.69
2:D:333:ARG:HH12	4:G:27:GLU:CD	2.00	0.64
1:B:358:SER:HB2	1:B:390:LEU:HD11	1.84	0.59
4:G:43:MET:HE3	4:G:107:ASN:HA	1.85	0.59
1:B:272:ALA:HB1	1:B:277:ASP:HB3	1.90	0.54
3:F:46:LYS:NZ	7:F:202:HOH:O	2.41	0.53
2:D:43:ARG:HG2	2:D:43:ARG:NH1	2.22	0.52
2:D:40:LYS:HG2	2:D:41:GLU:OE2	2.09	0.52
2:D:392:LYS:NZ	7:D:701:HOH:O	2.25	0.51
2:D:30:LYS:NZ	4:G:138:ILE:O	2.44	0.50
2:D:140:GLN:O	2:D:144:GLU:HG2	2.11	0.50
1:B:100:ARG:HG3	1:B:100:ARG:O	2.11	0.48
2:D:113:GLY:HA2	2:D:188:PHE:O	2.13	0.48
3:F:92:GLU:OE2	3:F:131:ARG:HD2	2.14	0.48
1:B:381:ASP:OD2	1:B:384:GLN:HB2	2.14	0.47
1:B:142:TYR:CE1	1:B:146:GLU:HG3	2.48	0.47
1:B:361:LEU:HD22	1:B:389:VAL:HG11	1.96	0.47
1:B:150:LYS:HB3	1:B:220:ILE:HD12	1.97	0.47
1:B:303:VAL:HG13	1:B:375:LYS:HB3	1.97	0.47
1:B:79:PHE:HB2	1:B:83:ARG:HB3	1.97	0.47
2:D:41:GLU:O	2:D:43:ARG:NH1	2.48	0.47
1:B:369:LYS:HA	1:B:373:ALA:HB3	1.98	0.46
3:F:3:LYS:HD3	3:F:3:LYS:C	2.37	0.45
2:D:333:ARG:NH1	4:G:27:GLU:OE1	2.49	0.45
2:D:281:TYR:CZ	2:D:285:VAL:HG21	2.52	0.45
2:D:186:ARG:NH1	7:D:707:HOH:O	2.41	0.45
2:D:212:PHE:O	2:D:215:PRO:HD2	2.17	0.44
2:D:215:PRO:HA	2:D:219:ALA:HB3	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:44:THR:HB	2:D:127:SER:HA	1.99	0.44
2:D:354:TRP:CG	2:D:355:PRO:HD3	2.52	0.44
1:B:144:ARG:NH2	1:B:145:ASP:OD2	2.51	0.43
2:D:209:GLU:OE1	2:D:246:HIS:HB3	2.19	0.43
2:D:41:GLU:HG3	7:D:796:HOH:O	2.19	0.42
4:G:43:MET:CE	4:G:107:ASN:HA	2.49	0.42
2:D:207:VAL:O	2:D:211:CYS:HB3	2.20	0.41
2:D:257:ILE:O	2:D:263:SER:OG	2.26	0.41
2:D:40:LYS:HE2	2:D:41:GLU:OE1	2.20	0.41
1:B:109:HIS:CE1	2:D:146:ARG:HB2	2.56	0.41
1:B:189:LYS:HD3	1:B:189:LYS:HA	1.76	0.41
2:D:140:GLN:HG3	2:D:246:HIS:CE1	2.56	0.40
2:D:204:LEU:HG	2:D:205:GLN:HG3	2.03	0.40

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	B	389/395 (98%)	378 (97%)	11 (3%)	0	100	100
2	D	510/526 (97%)	497 (98%)	13 (2%)	0	100	100
3	F	168/169 (99%)	166 (99%)	2 (1%)	0	100	100
4	G	134/138 (97%)	130 (97%)	4 (3%)	0	100	100
All	All	1201/1228 (98%)	1171 (98%)	30 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	B	323/327 (99%)	321 (99%)	2 (1%)	86	85
2	D	424/433 (98%)	418 (99%)	6 (1%)	67	62
3	F	147/146 (101%)	144 (98%)	3 (2%)	55	48
4	G	108/110 (98%)	106 (98%)	2 (2%)	57	50
All	All	1002/1016 (99%)	989 (99%)	13 (1%)	71	65

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	100	ARG
1	B	270	ARG
2	D	17	ASN
2	D	40	LYS
2	D	186	ARG
2	D	307	ARG
2	D	437	ARG
2	D	478	LEU
3	F	31	LYS
3	F	110[A]	ARG
3	F	110[B]	ARG
4	G	43	MET
4	G	138	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	17	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	GOL	B	401	-	5,5,5	0.86	0	5,5,5	1.08	1 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	B	401	-	-	4/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	401	GOL	C3-C2-C1	-2.03	103.83	111.70

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	401	GOL	C1-C2-C3-O3
5	B	401	GOL	O2-C2-C3-O3
5	B	401	GOL	O1-C1-C2-O2
5	B	401	GOL	O1-C1-C2-C3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	B	389/395 (98%)	0.16	14 (3%) 42 52	32, 43, 66, 89	0
2	D	511/526 (97%)	0.08	20 (3%) 39 49	30, 41, 59, 78	0
3	F	168/169 (99%)	0.30	10 (5%) 21 30	38, 50, 64, 80	0
4	G	136/138 (98%)	0.56	18 (13%) 3 5	35, 49, 70, 78	0
All	All	1204/1228 (98%)	0.19	62 (5%) 28 37	30, 44, 65, 89	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	169	ALA	6.1
4	G	138	ILE	5.8
4	G	4	ALA	4.6
4	G	3	SER	4.6
1	B	209	GLY	4.2
4	G	60	GLY	4.0
2	D	180	LEU	3.8
1	B	8	GLN	3.8
3	F	2	ALA	3.7
2	D	16	VAL	3.7
1	B	7	SER	3.5
3	F	27	ASP	3.5
4	G	59	GLY	3.5
4	G	136	THR	3.4
2	D	259	ASN	3.4
3	F	18	ALA	3.2
4	G	5	HIS	3.2
1	B	208	PRO	3.2
1	B	207	VAL	3.1
1	B	139	ILE	3.0
2	D	289	LEU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	F	3	LYS	3.0
2	D	105	VAL	2.9
2	D	55	GLU	2.9
4	G	133	ARG	2.9
2	D	40	LYS	2.9
4	G	134	ALA	2.8
1	B	349	ALA	2.8
1	B	384	GLN	2.7
1	B	6	SER	2.6
4	G	63	LYS	2.6
4	G	62	ALA	2.5
4	G	22	ASP	2.5
4	G	36	ASN	2.5
3	F	21	ALA	2.5
2	D	218	VAL	2.4
3	F	74	PRO	2.4
2	D	17	ASN	2.4
2	D	286	LEU	2.4
1	B	379	ARG	2.4
1	B	347	LYS	2.3
3	F	103	LYS	2.3
1	B	155	LEU	2.3
4	G	26	ALA	2.3
4	G	89	ALA	2.3
2	D	207	VAL	2.2
2	D	187	VAL	2.2
4	G	137	ASP	2.2
4	G	65	PRO	2.2
2	D	206	LEU	2.2
3	F	150	GLY	2.2
2	D	285	VAL	2.2
3	F	24	THR	2.1
2	D	179	PRO	2.1
4	G	66	SER	2.1
2	D	290	PHE	2.1
1	B	29	ASP	2.1
1	B	145	ASP	2.1
2	D	210	ALA	2.1
2	D	237	LEU	2.0
2	D	216	LEU	2.0
2	D	323	LYS	2.0

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)

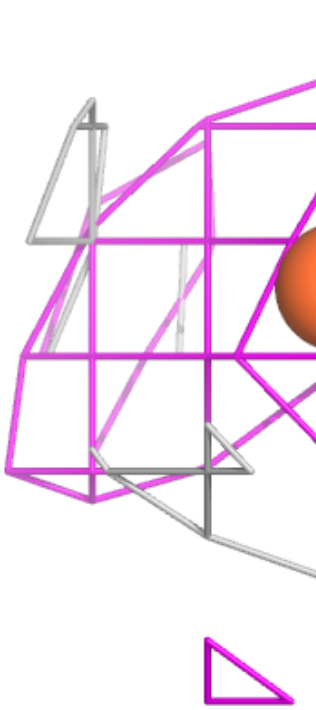
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

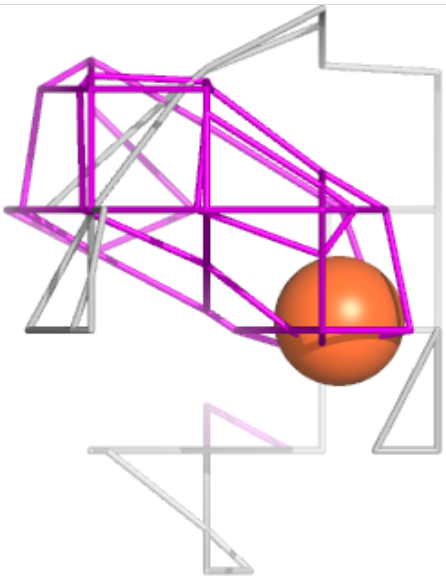
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	GOL	B	401	6/6	0.89	0.16	44,51,52,64	0
6	FE2	D	602	1/1	0.97	0.05	41,41,41,41	0
6	FE2	D	601	1/1	0.99	0.07	38,38,38,38	0

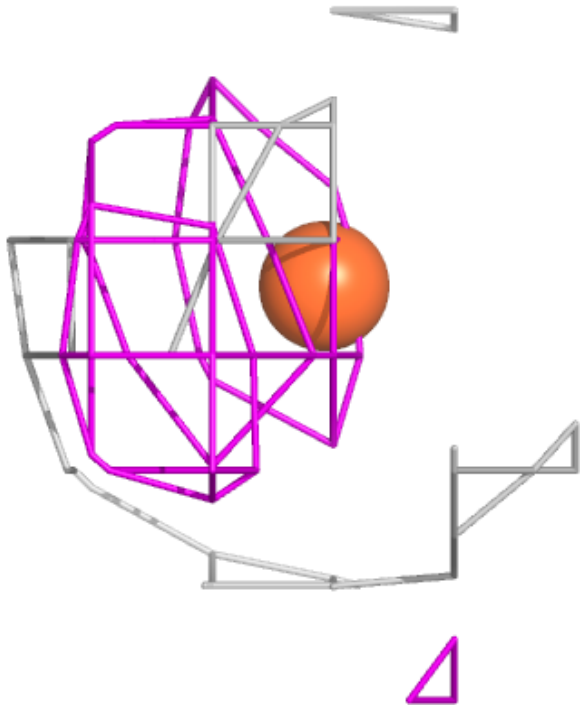
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around FE2 D 602:

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

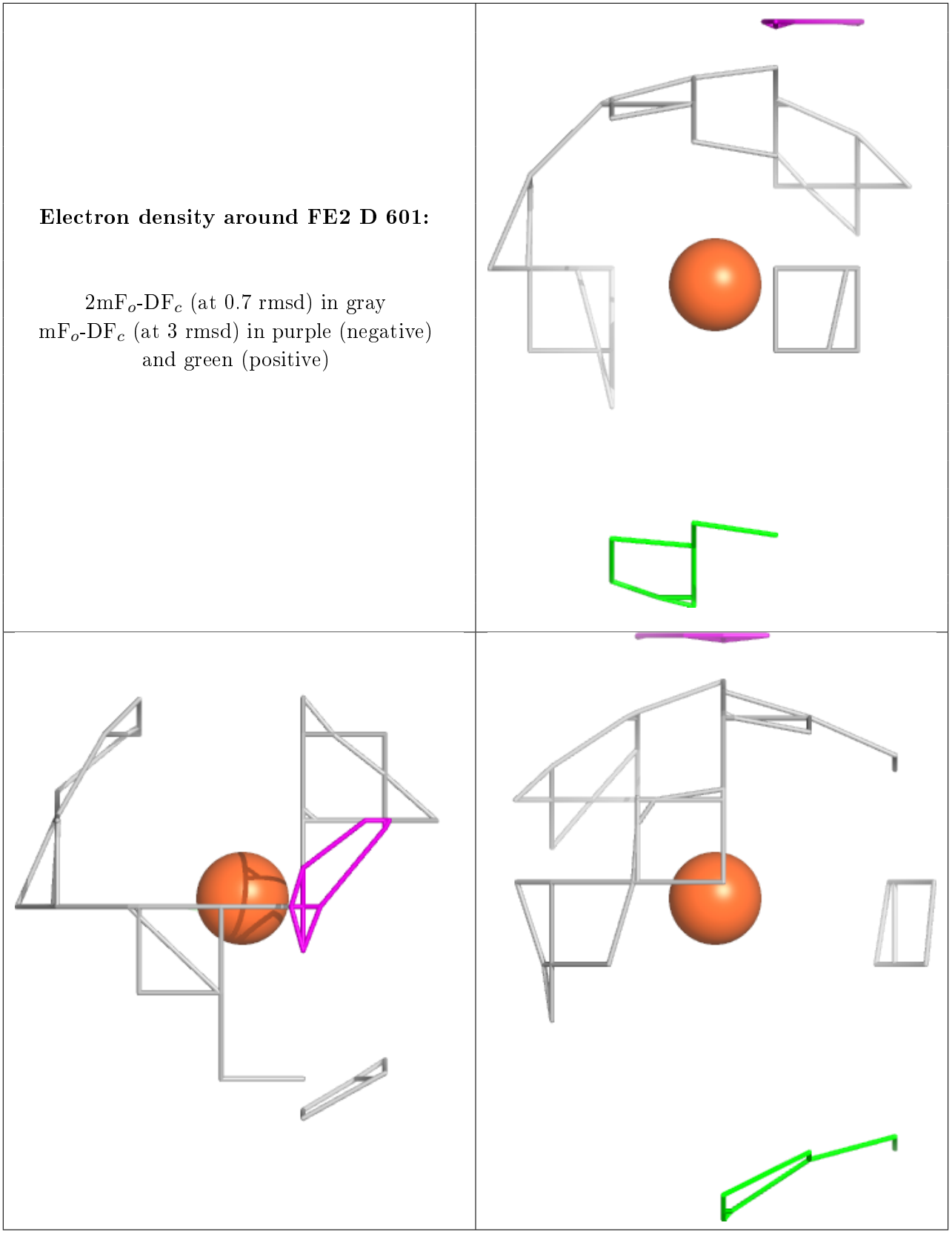






Electron density around FE2 D 601:

$2mF_o$-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.