wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2023 – 02:02 PM EDT

PDB ID	:	6O7M
Title	:	Nitrogenase MoFeP mutant F99Y from Azotobacter vinelandii in the indigo carmine oxidized state
Authors	:	Rutledge, H.L.; Williamson, L.M.; Tezcan, F.A.
Deposited on	:	2019-03-08
Resolution	:	1.40 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:

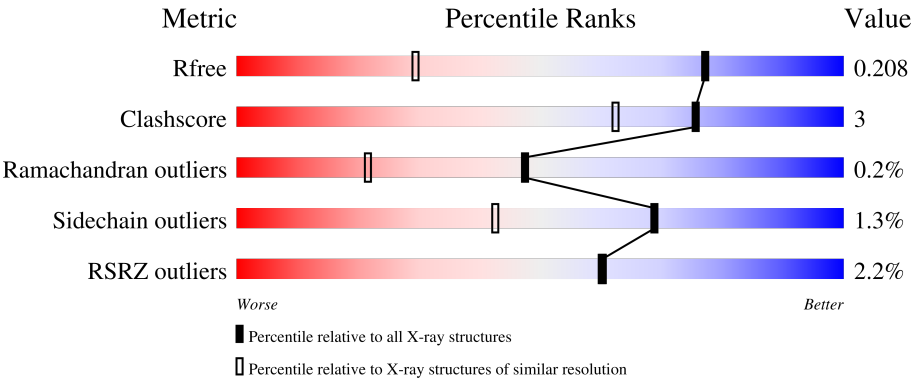
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.40 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	A	492	3% 87% 9% .
1	C	492	4% 91% 6% .
2	B	523	% 93% 7%
2	D	523	2% 94% 6%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 34288 atoms, of which 15682 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 Nitrogenase molybdenum-iron protein alpha chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	477	Total	C	H	N	O	S	0	7	0
			7525	2423	3729	646	701	26			
1	C	479	Total	C	H	N	O	S	0	6	1
			7594	2431	3772	651	712	28			

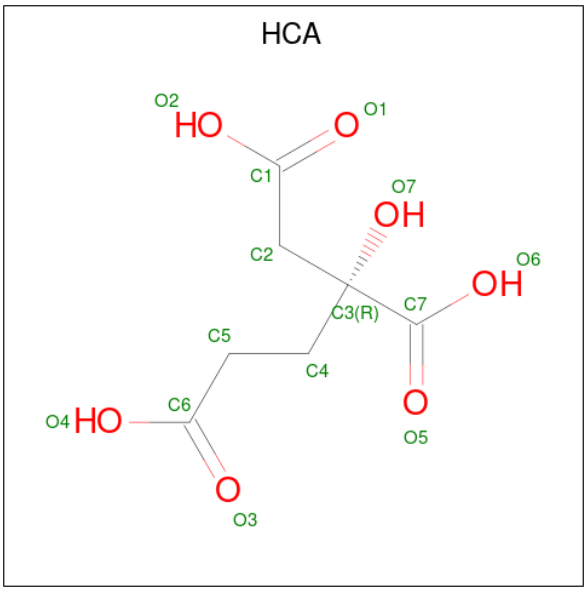
- Molecule 2 is a protein called Nitrogenase molybdenum-iron protein beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	522	Total	C	H	N	O	S	0	0	0
			8216	2657	4055	700	776	28			
2	D	522	Total	C	H	N	O	S	0	5	0
			8304	2681	4114	705	776	28			

There are 2 discrepancies between the modelled and reference sequences:

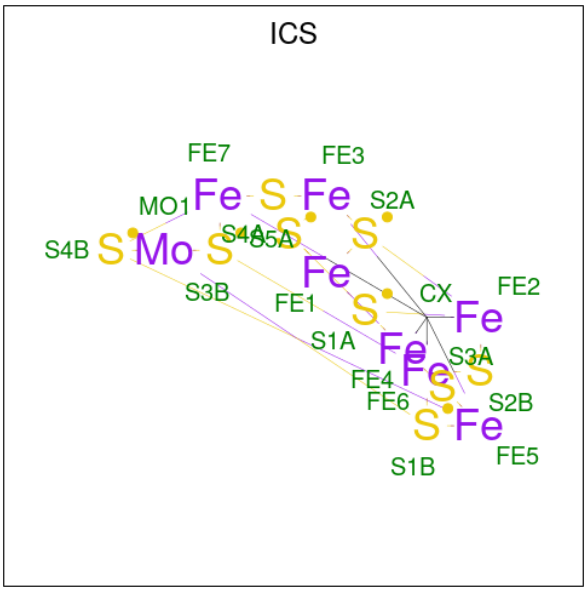
Chain	Residue	Modelled	Actual	Comment	Reference
B	99	TYR	PHE	engineered mutation	UNP P07329
D	99	TYR	PHE	engineered mutation	UNP P07329

- Molecule 3 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula: $C_7H_{10}O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			20	7	6	7		
3	C	1	Total	C	H	O	0	0
			20	7	6	7		

- Molecule 4 is iron-sulfur-molybdenum cluster with interstitial carbon (three-letter code: ICS) (formula: CFe_7MoS_9).



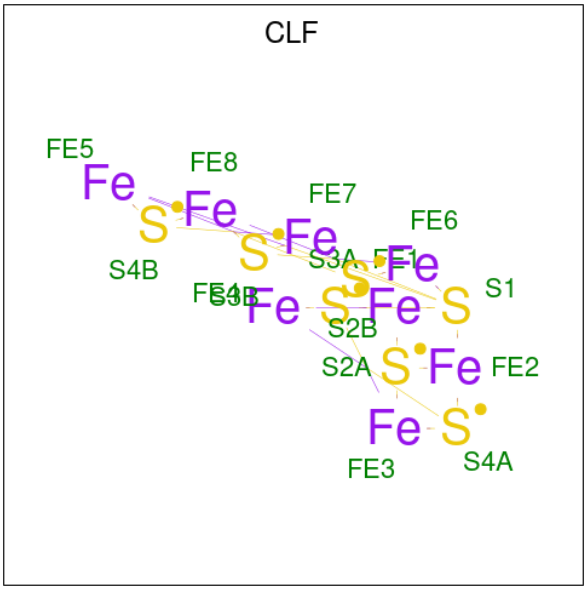
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	Fe	Mo	S	0	0
			18	1	7	1	9		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total C Fe Mo S 18 1 7 1 9	0	0

- Molecule 5 is FE(8)-S(7) CLUSTER (three-letter code: CLF) (formula: Fe_8S_7).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Fe S 15 8 7	0	0
5	C	1	Total Fe S 15 8 7	0	0

- Molecule 6 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Fe 1 1	0	0
6	D	1	Total Fe 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	579	Total O 579 579	0	0
7	B	692	Total O 692 692	0	0

Continued on next page...

Continued from previous page...

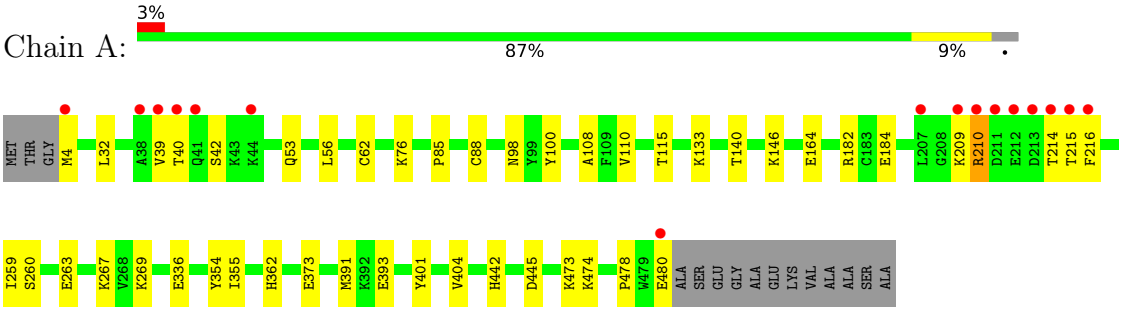
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	558	Total O 558 558	0	0
7	D	712	Total O 712 712	0	0

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.

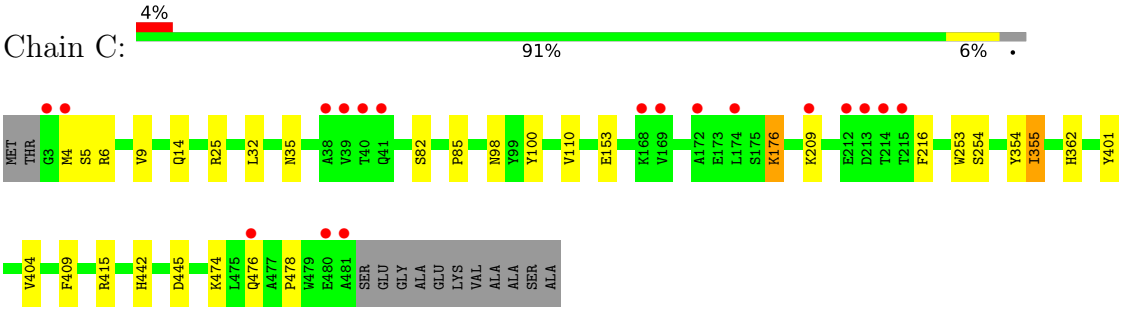
• Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

Chain A:



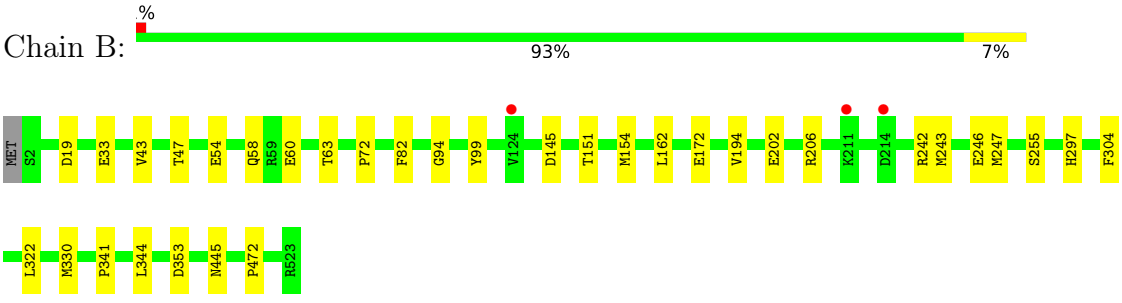
• Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

Chain C:



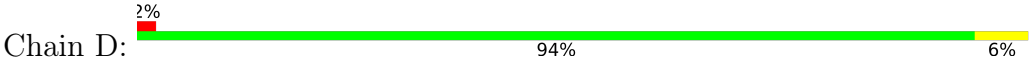
• Molecule 2: Nitrogenase molybdenum-iron protein beta chain

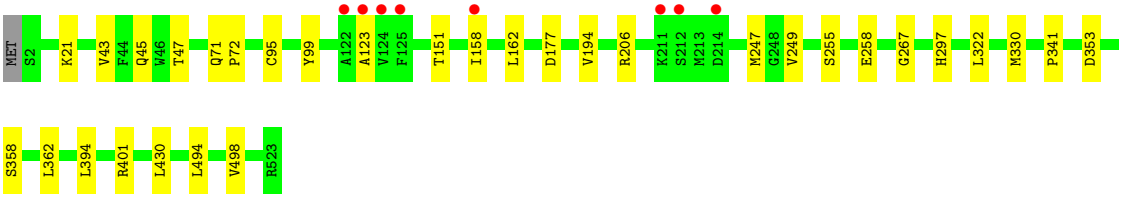
Chain B:



• Molecule 2: Nitrogenase molybdenum-iron protein beta chain

Chain D:




MET
S2
K21
V43
F44
Q45
W46
T47
Q71
P72
C95
Y99
A122
A123
V124
F125
T151
I158
L162
D177
V194
R206
K211
S212
M213
D214
M247
G248
V249
S255
E258
G267
H297
L322
M330
P341
D353
S358
L362
L394
R401
L430
L494
V498
R523

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	76.46Å 127.79Å 107.54Å 90.00° 109.00° 90.00°	Depositor
Resolution (Å)	40.11 – 1.40 40.11 – 1.40	Depositor EDS
% Data completeness (in resolution range)	93.1 (40.11-1.40) 93.5 (40.11-1.40)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 1.40Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.183 , 0.208 0.182 , 0.208	Depositor DCC
R_{free} test set	35750 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å^2)	10.5	Xtriage
Anisotropy	0.502	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 42.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	34288	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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: FE, HCA, ICS, CLF

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	0.37	0/3905	0.60	0/5266
1	C	0.37	0/3928	0.59	0/5294
2	B	0.38	0/4267	0.58	1/5773 (0.0%)
2	D	0.37	0/4311	0.57	0/5829
All	All	0.37	0/16411	0.59	1/22162 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	19	ASP	CB-CG-OD1	5.25	123.03	118.30

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	A	3796	3729	3733	39	0
1	C	3822	3772	3777	21	0
2	B	4161	4055	4054	25	0
2	D	4190	4114	4118	21	0
3	A	14	6	6	3	0
3	C	14	6	6	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	18	0	0	0	0
4	C	18	0	0	0	0
5	B	15	0	0	1	0
5	C	15	0	0	1	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
7	A	579	0	0	10	2
7	B	692	0	0	7	1
7	C	558	0	0	7	4
7	D	712	0	0	4	3
All	All	18606	15682	15694	98	5

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 98 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:A:210:ARG:HD3	1:A:263:GLU:CD	1.93	0.86
1:A:210:ARG:HD3	1:A:263:GLU:OE2	1.82	0.80
1:C:25:ARG:NH1	7:C:602:HOH:O	2.15	0.79
1:C:176:LYS:NZ	7:C:604:HOH:O	2.17	0.77
2:B:322:LEU:HD23	1:C:474:LYS:HD3	1.68	0.74

All (5) 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)
7:C:1140:HOH:O	7:D:1248:HOH:O[2_655]	1.99	0.21
7:A:909:HOH:O	7:B:1133:HOH:O[2_646]	2.04	0.16
7:A:1179:HOH:O	7:C:1136:HOH:O[1_556]	2.13	0.07
7:C:979:HOH:O	7:D:1346:HOH:O[2_655]	2.15	0.05
7:C:1104:HOH:O	7:D:1255:HOH:O[2_655]	2.19	0.01

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	482/492 (98%)	456 (95%)	25 (5%)	1 (0%)	47	21
1	C	483/492 (98%)	462 (96%)	20 (4%)	1 (0%)	47	21
2	B	520/523 (99%)	507 (98%)	12 (2%)	1 (0%)	47	21
2	D	524/523 (100%)	513 (98%)	10 (2%)	1 (0%)	47	21
All	All	2009/2030 (99%)	1938 (96%)	67 (3%)	4 (0%)	47	21

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	255	SER
2	D	255	SER
1	C	355	ILE
1	A	355	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	402/415 (97%)	392 (98%)	10 (2%)	47	14
1	C	413/415 (100%)	405 (98%)	8 (2%)	57	25
2	B	451/455 (99%)	450 (100%)	1 (0%)	93	82
2	D	457/455 (100%)	453 (99%)	4 (1%)	78	58
All	All	1723/1740 (99%)	1700 (99%)	23 (1%)	69	42

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	355	ILE
1	C	409	PHE
1	C	401	TYR
1	C	445	ASP
1	A	362	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	CLF	B	601	1,2	0,24,24	-	-	-		
4	ICS	A	502	1	18,30,30	2.81	11 (61%)	-		
4	ICS	C	502	1	18,30,30	2.95	13 (72%)	-		
3	HCA	A	501	-	13,13,13	0.91	0	14,18,18	1.65	5 (35%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	HCA	C	501	-	13,13,13	0.96	0	14,18,18	1.41	2 (14%)
5	CLF	C	503	1,2	0,24,24	-	-	-		

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	CLF	C	503	1,2	-	-	0/12/10/10
3	HCA	A	501	-	-	7/17/17/17	-
5	CLF	B	601	1,2	-	-	0/12/10/10
3	HCA	C	501	-	-	0/17/17/17	-

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	502	ICS	S3B-FE6	-4.71	2.20	2.32
4	C	502	ICS	S2A-FE2	-4.65	2.21	2.32
4	C	502	ICS	S3B-FE6	-4.59	2.21	2.32
4	C	502	ICS	S4B-FE5	-4.31	2.21	2.32
4	A	502	ICS	S4B-FE7	-4.27	2.21	2.32

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	501	HCA	O6-C7-C3	3.31	118.81	113.05
3	A	501	HCA	O6-C7-C3	2.93	118.14	113.05
3	C	501	HCA	C4-C5-C6	2.60	118.63	112.75
3	A	501	HCA	O3-C6-C5	-2.47	115.14	123.08
3	A	501	HCA	O5-C7-C3	-2.29	119.00	122.25

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	HCA	C2-C3-C4-C5
3	A	501	HCA	C4-C3-C7-O5
3	A	501	HCA	O1-C1-C2-C3
3	A	501	HCA	C4-C3-C7-O6
3	A	501	HCA	O2-C1-C2-C3

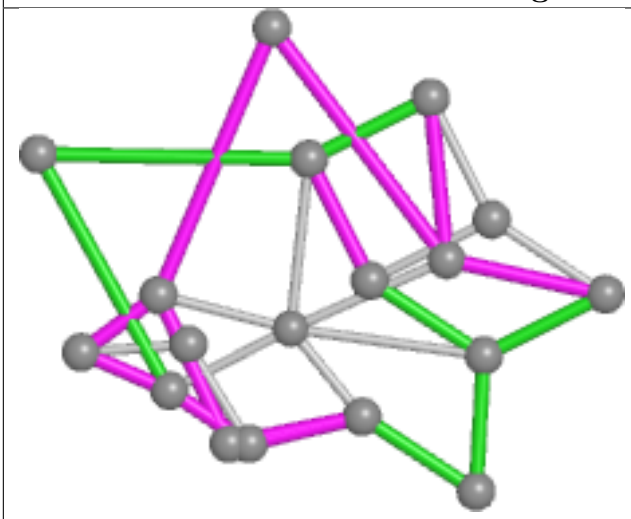
There are no ring outliers.

4 monomers are involved in 7 short contacts:

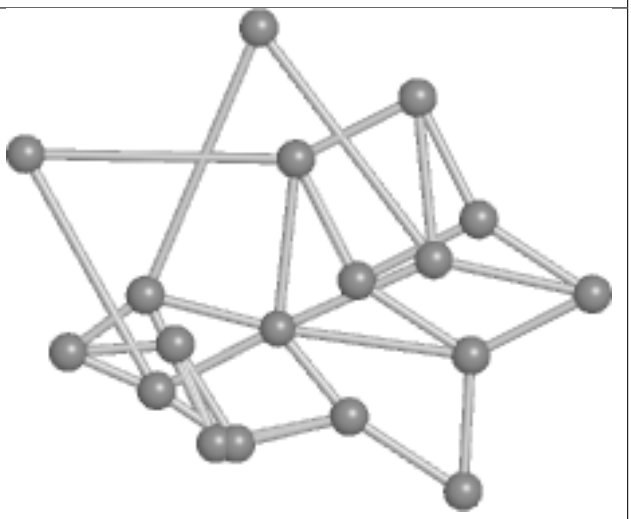
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	601	CLF	1	0
3	A	501	HCA	3	0
3	C	501	HCA	2	0
5	C	503	CLF	1	0

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 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.

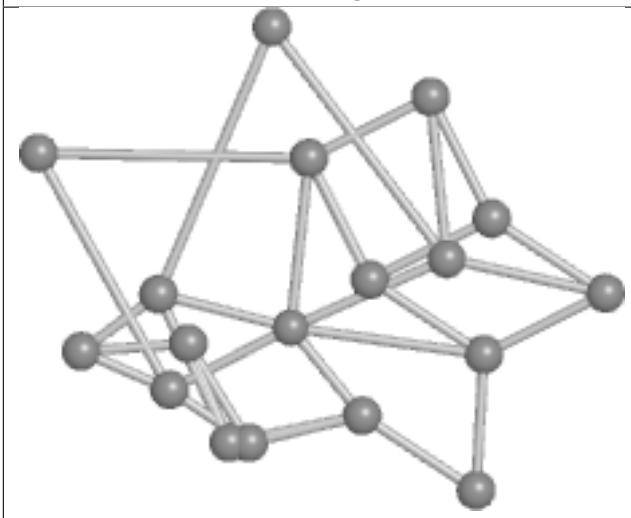
Ligand ICS A 502



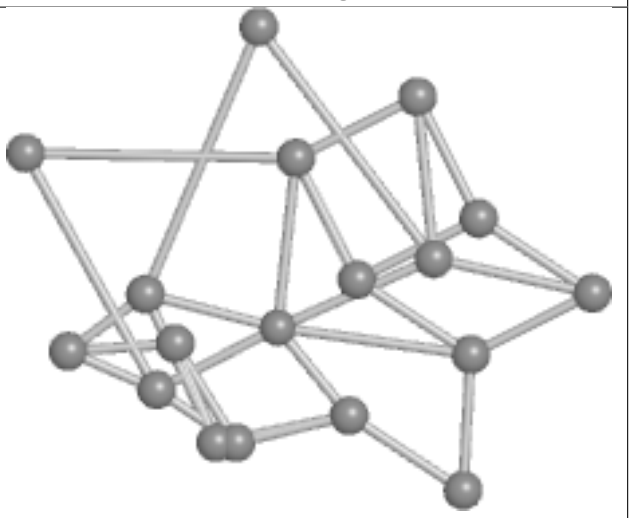
Bond lengths



Bond angles

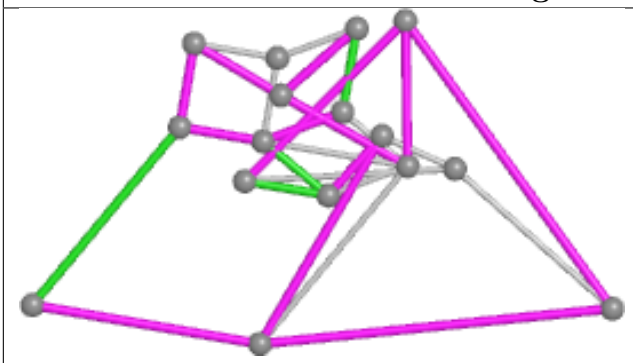


Torsions

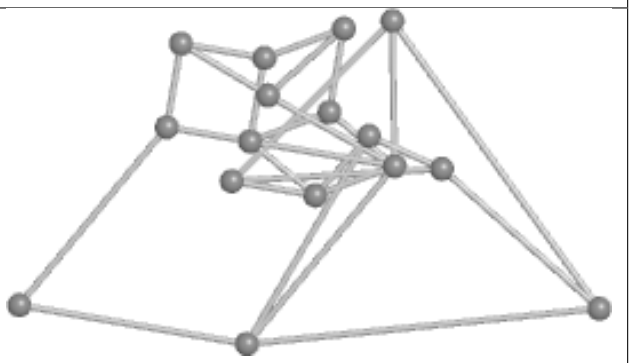


Rings

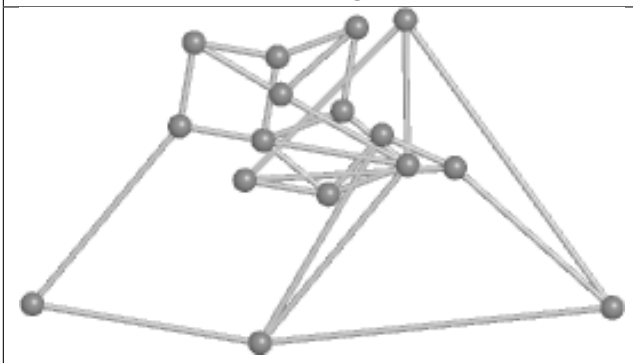
Ligand ICS C 502



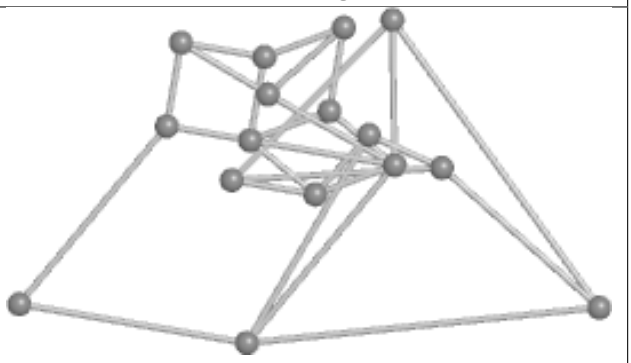
Bond lengths



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	A	477/492 (96%)	0.02	16 (3%)	45	44	8, 14, 30, 50	0
1	C	479/492 (97%)	0.02	18 (3%)	40	40	9, 14, 30, 43	0
2	B	522/523 (99%)	-0.20	3 (0%)	89	88	8, 12, 23, 35	0
2	D	522/523 (99%)	-0.16	8 (1%)	73	72	8, 13, 24, 38	0
All	All	2000/2030 (98%)	-0.08	45 (2%)	60	60	8, 13, 26, 50	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	214	THR	15.7
1	A	215	THR	10.0
1	A	210	ARG	7.0
1	A	213	ASP	5.3
1	C	481	ALA	5.2

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

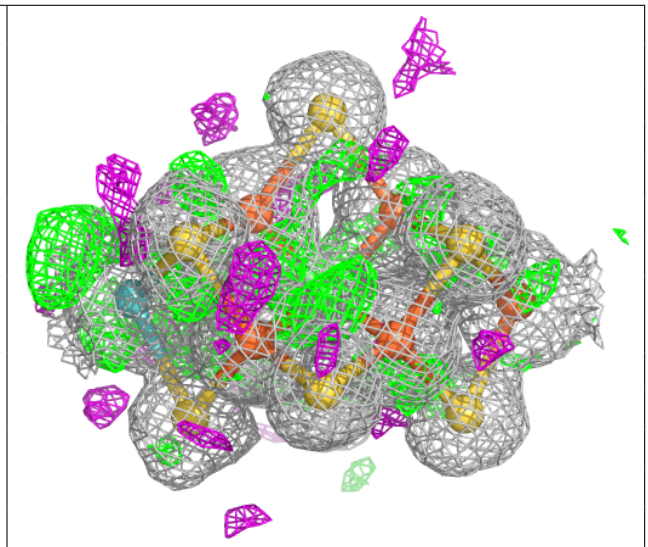
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

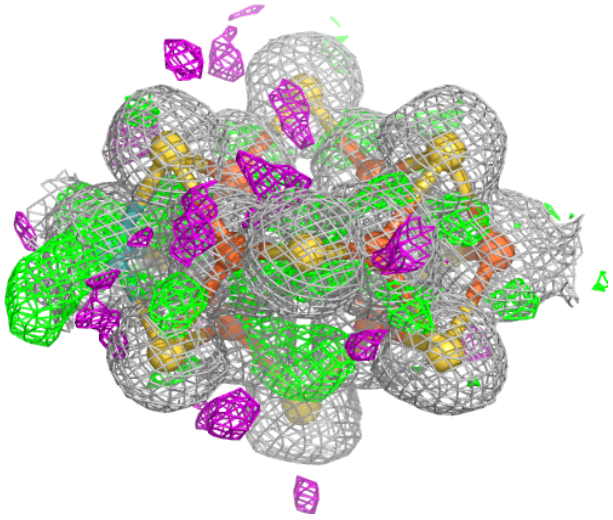
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HCA	C	501	14/14	0.96	0.11	9,11,14,15	0
6	FE	D	601	1/1	0.96	0.06	13,13,13,13	1
3	HCA	A	501	14/14	0.97	0.09	8,10,12,14	0
5	CLF	C	503	15/15	0.99	0.05	8,10,11,11	0
6	FE	B	602	1/1	0.99	0.05	13,13,13,13	1
5	CLF	B	601	15/15	0.99	0.05	7,8,9,10	0
4	ICS	A	502	18/18	1.00	0.06	5,9,10,10	0
4	ICS	C	502	18/18	1.00	0.05	6,9,10,11	0

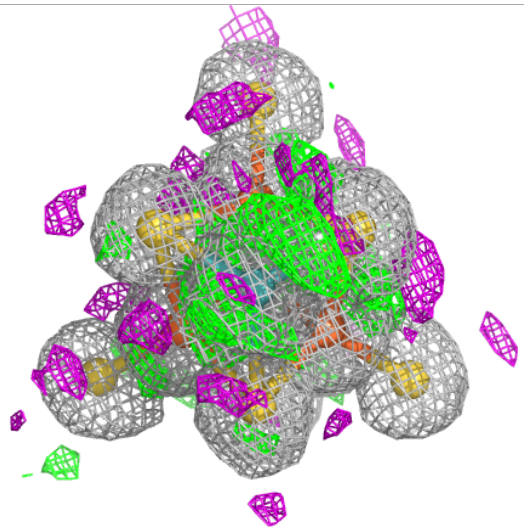
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ICS A 502:

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

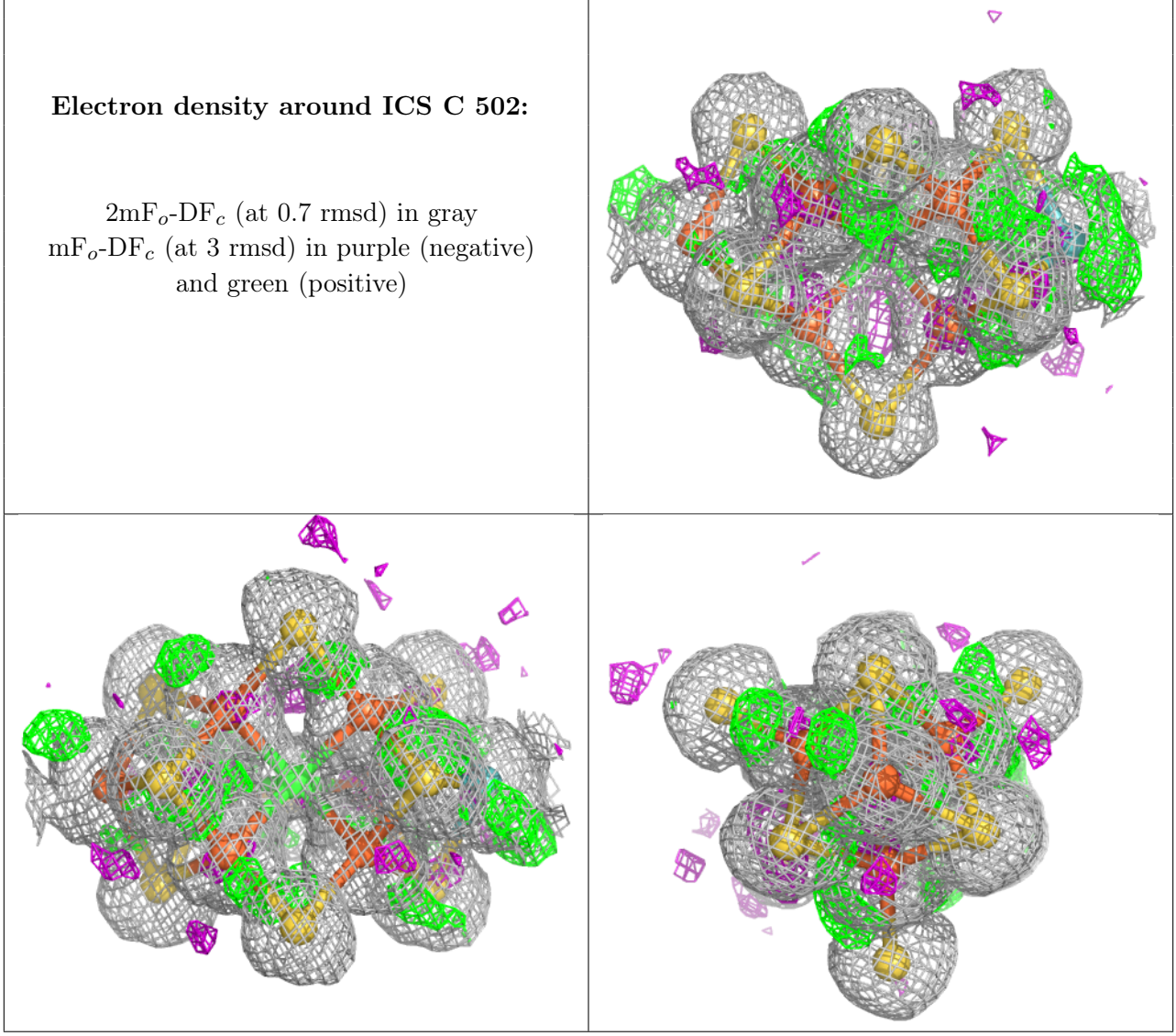






Electron density around ICS C 502:

$2mF_o$-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.