wwPDB X-ray Structure Validation Summary Report (i)

Dec 3, 2020 – 01:10 pm GMT

PDB ID	:	6TTL
Title	:	crystal structure of [FeFe]-hydrogenase CbA5H (partial) from Clostridium beijerinckii in Hinact state
Authors	:	Duan, J.; Rutz, A.; Hofmann, E.; Happe, T.
Deposited on	:	2019-12-28
Resolution	:	2.90 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 following versions of software and data (see references (1)) were used in the production of this report:

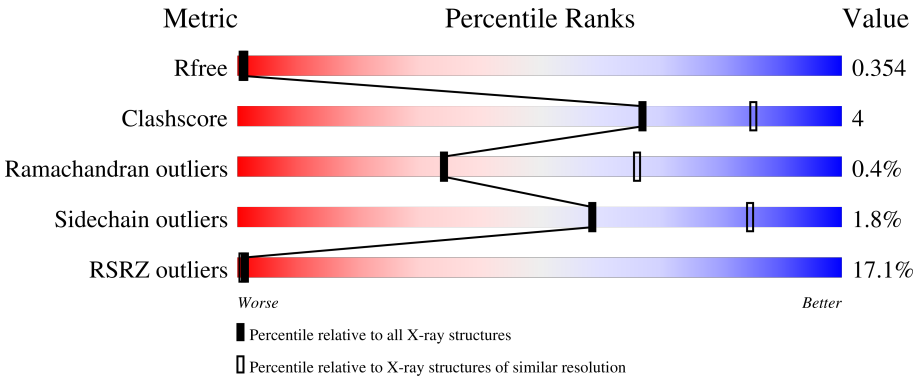
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	A	674	11% 65% 6% 29%
1	B	674	14% 62% 8% · 29%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7648 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 [FeFe]-hydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	479	Total	C	N	O	S	0	0	0
			3799	2407	643	725	24			
1	B	479	Total	C	N	O	S	0	0	0
			3799	2407	643	725	24			

There are 60 discrepancies between the modelled and reference sequences:

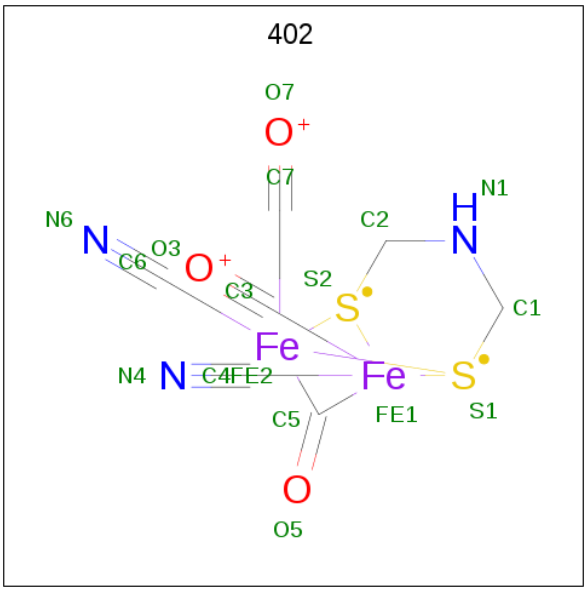
Chain	Residue	Modelled	Actual	Comment	Reference
A	645	ASP	-	expression tag	UNP A0A1I9RYV3
A	646	ILE	-	expression tag	UNP A0A1I9RYV3
A	647	TRP	-	expression tag	UNP A0A1I9RYV3
A	648	SER	-	expression tag	UNP A0A1I9RYV3
A	649	VAL	-	expression tag	UNP A0A1I9RYV3
A	650	GLY	-	expression tag	UNP A0A1I9RYV3
A	651	VAL	-	expression tag	UNP A0A1I9RYV3
A	652	LYS	-	expression tag	UNP A0A1I9RYV3
A	653	LEU	-	expression tag	UNP A0A1I9RYV3
A	654	PHE	-	expression tag	UNP A0A1I9RYV3
A	655	GLY	-	expression tag	UNP A0A1I9RYV3
A	656	GLY	-	expression tag	UNP A0A1I9RYV3
A	657	GLY	-	expression tag	UNP A0A1I9RYV3
A	658	SER	-	expression tag	UNP A0A1I9RYV3
A	659	GLY	-	expression tag	UNP A0A1I9RYV3
A	660	GLY	-	expression tag	UNP A0A1I9RYV3
A	661	GLY	-	expression tag	UNP A0A1I9RYV3
A	662	SER	-	expression tag	UNP A0A1I9RYV3
A	663	GLY	-	expression tag	UNP A0A1I9RYV3
A	664	GLY	-	expression tag	UNP A0A1I9RYV3
A	665	GLY	-	expression tag	UNP A0A1I9RYV3
A	666	SER	-	expression tag	UNP A0A1I9RYV3
A	667	TRP	-	expression tag	UNP A0A1I9RYV3
A	668	SER	-	expression tag	UNP A0A1I9RYV3
A	669	HIS	-	expression tag	UNP A0A1I9RYV3

Continued on next page...

Continued from previous page...

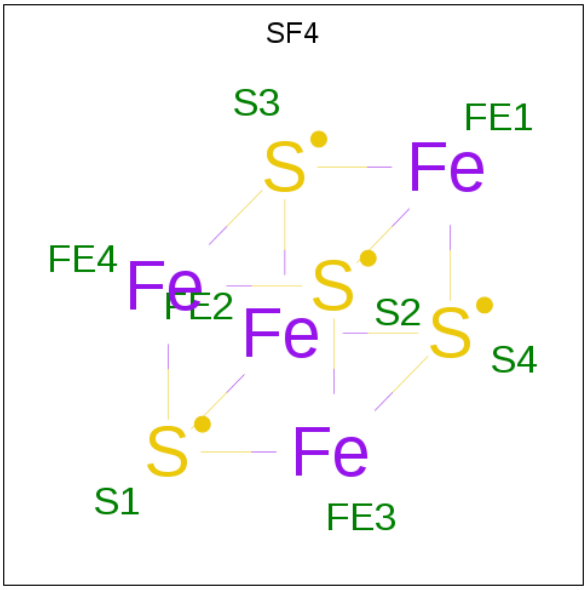
Chain	Residue	Modelled	Actual	Comment	Reference
A	670	PRO	-	expression tag	UNP A0A1I9RYV3
A	671	GLN	-	expression tag	UNP A0A1I9RYV3
A	672	PHE	-	expression tag	UNP A0A1I9RYV3
A	673	GLU	-	expression tag	UNP A0A1I9RYV3
A	674	LYS	-	expression tag	UNP A0A1I9RYV3
B	645	ASP	-	expression tag	UNP A0A1I9RYV3
B	646	ILE	-	expression tag	UNP A0A1I9RYV3
B	647	TRP	-	expression tag	UNP A0A1I9RYV3
B	648	SER	-	expression tag	UNP A0A1I9RYV3
B	649	VAL	-	expression tag	UNP A0A1I9RYV3
B	650	GLY	-	expression tag	UNP A0A1I9RYV3
B	651	VAL	-	expression tag	UNP A0A1I9RYV3
B	652	LYS	-	expression tag	UNP A0A1I9RYV3
B	653	LEU	-	expression tag	UNP A0A1I9RYV3
B	654	PHE	-	expression tag	UNP A0A1I9RYV3
B	655	GLY	-	expression tag	UNP A0A1I9RYV3
B	656	GLY	-	expression tag	UNP A0A1I9RYV3
B	657	GLY	-	expression tag	UNP A0A1I9RYV3
B	658	SER	-	expression tag	UNP A0A1I9RYV3
B	659	GLY	-	expression tag	UNP A0A1I9RYV3
B	660	GLY	-	expression tag	UNP A0A1I9RYV3
B	661	GLY	-	expression tag	UNP A0A1I9RYV3
B	662	SER	-	expression tag	UNP A0A1I9RYV3
B	663	GLY	-	expression tag	UNP A0A1I9RYV3
B	664	GLY	-	expression tag	UNP A0A1I9RYV3
B	665	GLY	-	expression tag	UNP A0A1I9RYV3
B	666	SER	-	expression tag	UNP A0A1I9RYV3
B	667	TRP	-	expression tag	UNP A0A1I9RYV3
B	668	SER	-	expression tag	UNP A0A1I9RYV3
B	669	HIS	-	expression tag	UNP A0A1I9RYV3
B	670	PRO	-	expression tag	UNP A0A1I9RYV3
B	671	GLN	-	expression tag	UNP A0A1I9RYV3
B	672	PHE	-	expression tag	UNP A0A1I9RYV3
B	673	GLU	-	expression tag	UNP A0A1I9RYV3
B	674	LYS	-	expression tag	UNP A0A1I9RYV3

- Molecule 2 is dicarbonyl[bis(cyanide-kappaC)]-mu-(iminodimethanethiolatato-1kappaS:2kappaS)-mu-(oxomethylidene)diiron(2+) (three-letter code: 402) (formula: $C_7H_5Fe_2N_3O_3S_2$) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	S	0	0
			17	7	2	3	3	2		
2	B	1	Total	C	Fe	N	O	S	0	0
			17	7	2	3	3	2		

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



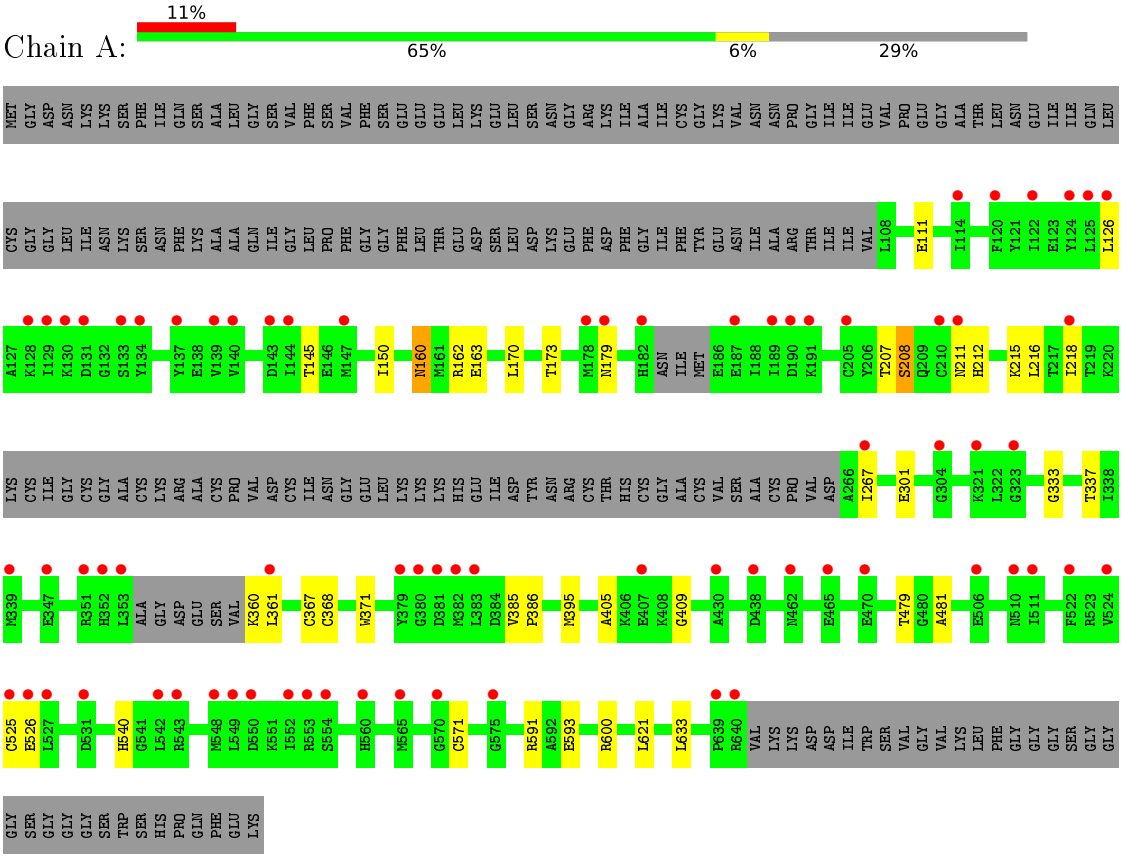
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	Fe	S		0	0
			8	4	4			
3	B	1	Total	Fe	S		0	0
			8	4	4			

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.

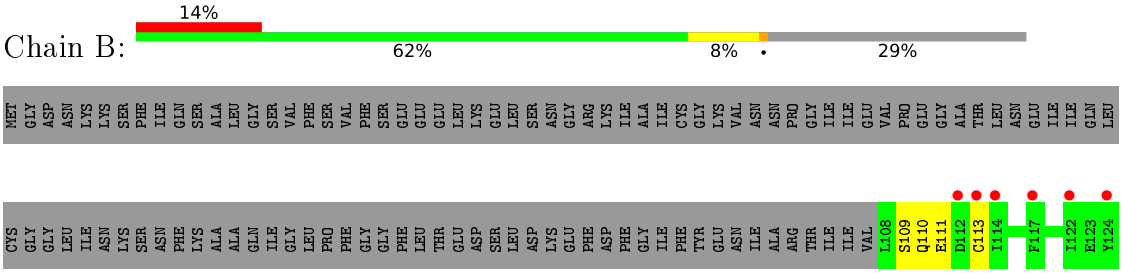
- Molecule 1: [FeFe]-hydrogenase

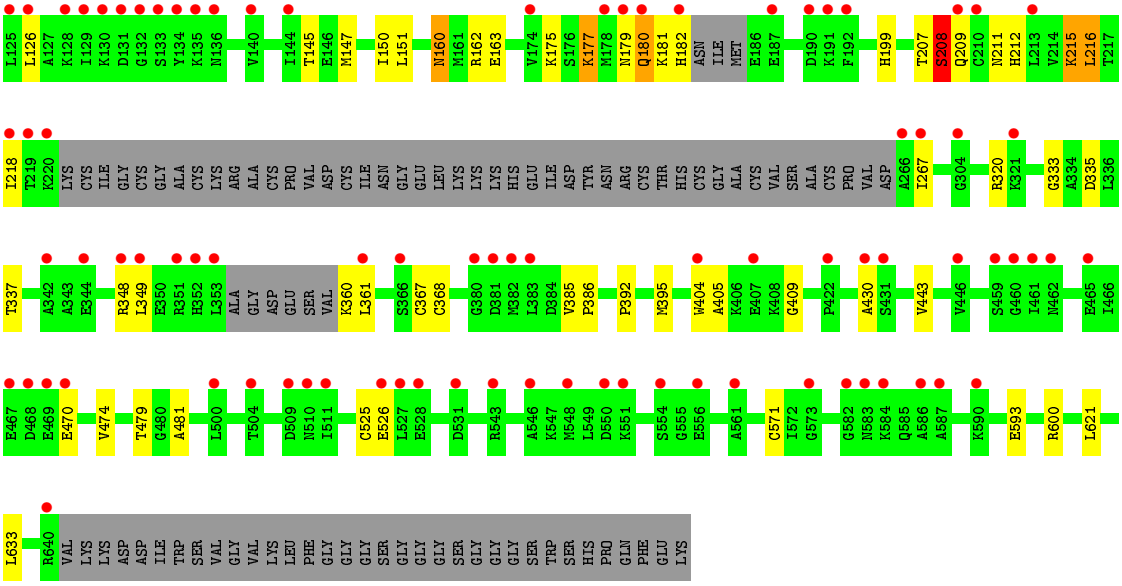
Chain A:



- Molecule 1: [FeFe]-hydrogenase

Chain B:




L125 L126 A127 K128 I129 K130 D131 G132 S133 Y134 K135 N136 V140 I144 T145 E146 M147 I150 L151 N160 M161 R162 E163 V174 K175 S176 K177 M178 N179 Q180 K181 H182 ASN ILE MET E186 E187 D190 K191 F192 H199 T207 S208 Q209 C210 N211 H212 L213 V214 K215 L216 T217
I218 T219 K220 LYS CYS ILE GLY CYS GLY ALA CYS LYS ARG ALA CYS PRO VAL ASP CYS ILE ASN GLY GLU LEU LYS LYS LYS HIS GLU ILE ASP TYR ASN ARG CYS THR HIS CYS GLY ALA CYS VAL SER ALA CYS PRO VAL ASP A266 I267 G304 R320 K321 G333 A334 D335 L336
T337 A342 A343 E344 R348 L349 E350 R351 H352 L353 ALA GLY ASP GLU SER VAL K360 L361 S366 C367 C368 G380 D381 M382 L383 D384 V385 P386 P392 M395 W404 A405 K406 E407 K408 G409 P422 A430 S431 V443 V446 S459 G460 I461 N462 E465 I466
E467 D468 E469 E470 V474 T479 G480 A481 L500 T504 D509 N510 I511 C525 E526 L527 E528 D531 R543 A546 K547 M548 L549 D550 K551 S554 G555 E556 A561 C571 I572 G573 G582 N583 K584 Q585 A586 A587 K590 E593 R600 L621
L633 R640 VAL LYS LYS ASP ASP ILE TRP SER VAL GLY VAL LYS LEU PHE GLY GLY GLY SER GLY GLY GLY SER GLY GLY GLY GLY SER TRP SER HIS PRO GLN PHE GLU LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	169.00Å 169.00Å 127.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.26 – 2.90 49.26 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.26-2.90) 100.0 (49.26-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.13	Depositor
R, R_{free}	0.339 , 0.353 0.341 , 0.354	Depositor DCC
R_{free} test set	2065 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	87.5	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	7648	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 62.04 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1968e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: SF4, 402

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	0.24	0/3859	0.41	0/5186
1	B	0.24	0/3859	0.41	0/5186
All	All	0.24	0/7718	0.41	0/10372

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

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	3799	0	3807	25	0
1	B	3799	0	3807	35	0
2	A	17	0	5	3	0
2	B	17	0	5	3	0
3	A	8	0	0	0	0
3	B	8	0	0	0	0
All	All	7648	0	7624	60	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 4.

The worst 5 of 60 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:B:207:THR:O	1:B:208:SER:HB2	1.84	0.77
1:B:405:ALA:O	1:B:409:GLY:N	2.24	0.70
1:A:405:ALA:O	1:A:409:GLY:N	2.25	0.69
1:B:160:ASN:OD1	1:B:162:ARG:NH1	2.27	0.68
1:A:160:ASN:OD1	1:A:162:ARG:NH1	2.27	0.67

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	A	471/674 (70%)	432 (92%)	37 (8%)	2 (0%)	34	66
1	B	471/674 (70%)	432 (92%)	37 (8%)	2 (0%)	34	66
All	All	942/1348 (70%)	864 (92%)	74 (8%)	4 (0%)	34	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	208	SER
1	B	208	SER
1	A	216	LEU
1	B	216	LEU

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	408/565 (72%)	403 (99%)	5 (1%)	71 91
1	B	408/565 (72%)	398 (98%)	10 (2%)	47 78
All	All	816/1130 (72%)	801 (98%)	15 (2%)	59 85

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	177	LYS
1	B	180	GLN
1	B	215	LYS
1	B	160	ASN
1	B	209	GLN

Some 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)

4 ligands are modelled in this entry.

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	SF4	B	701	1	0,12,12	0.00	-	-		
2	402	A	700	1	13,19,19	2.36	7 (53%)	2,36,36	1.23	0
3	SF4	A	701	1	0,12,12	0.00	-	-		
2	402	B	700	1	13,19,19	2.33	6 (46%)	2,36,36	1.55	1 (50%)

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
3	SF4	B	701	1	-	-	0/6/5/5
2	402	A	700	1	-	-	0/5/3/3
3	SF4	A	701	1	-	-	0/6/5/5
2	402	B	700	1	-	-	0/5/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	700	402	S1-FE2	4.13	2.32	2.26
2	A	700	402	S1-FE2	4.12	2.32	2.26
2	B	700	402	C4-N4	3.81	1.21	1.15
2	A	700	402	C4-N4	3.73	1.21	1.15
2	A	700	402	O3-C3	3.04	1.23	1.15

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	700	402	S2-C2-N1	2.18	123.17	117.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

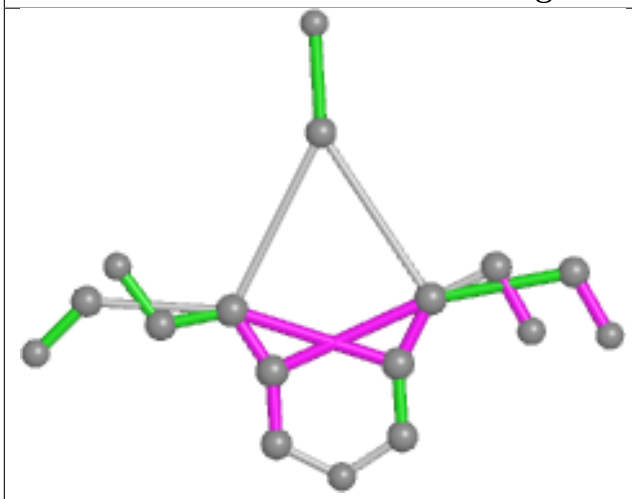
2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	700	402	3	0
2	B	700	402	3	0

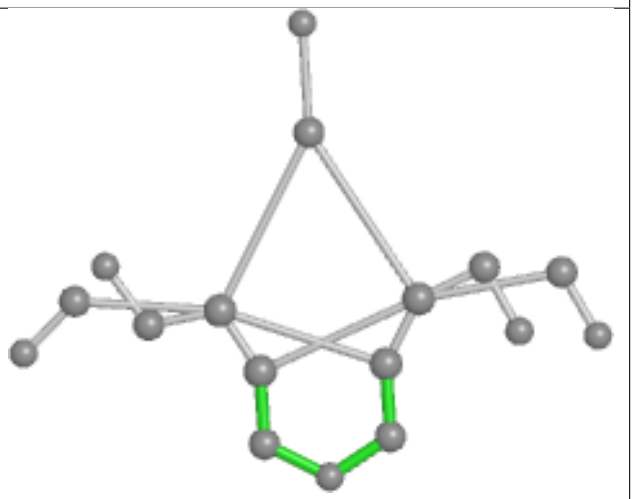
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

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.

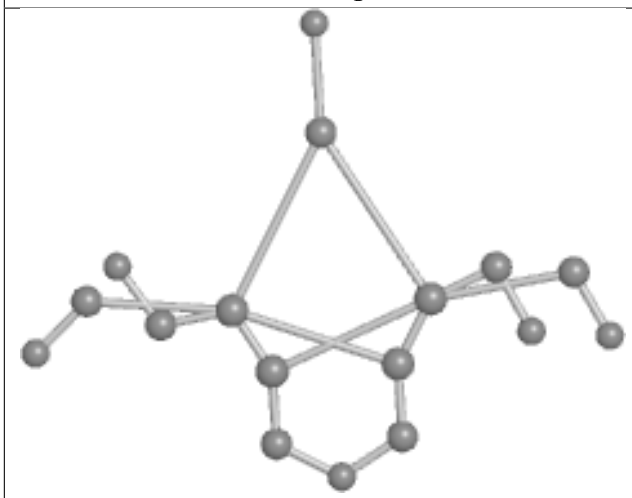
Ligand 402 A 700



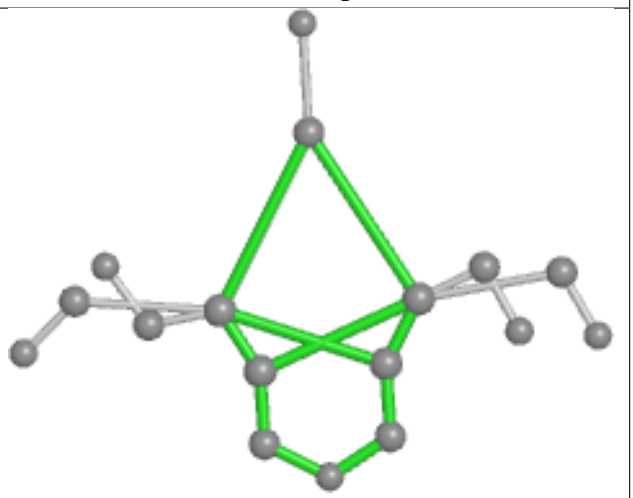
Bond lengths



Bond angles

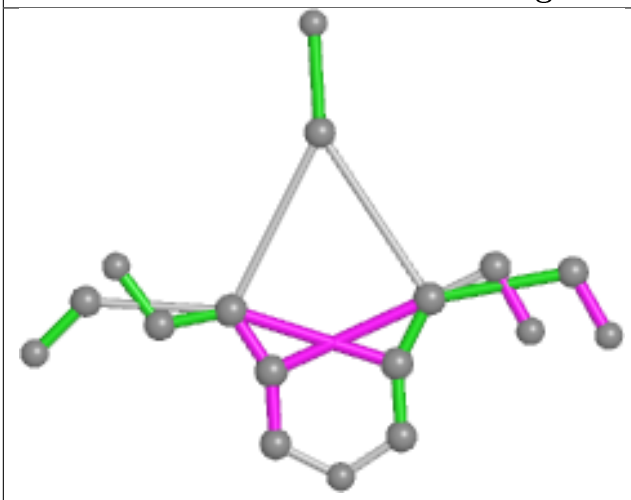


Torsions

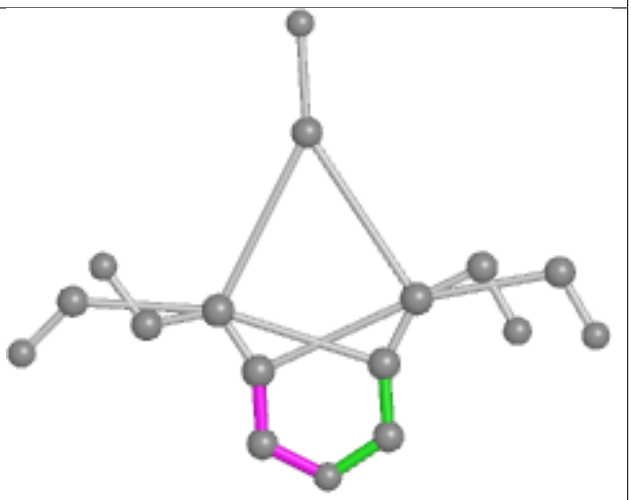


Rings

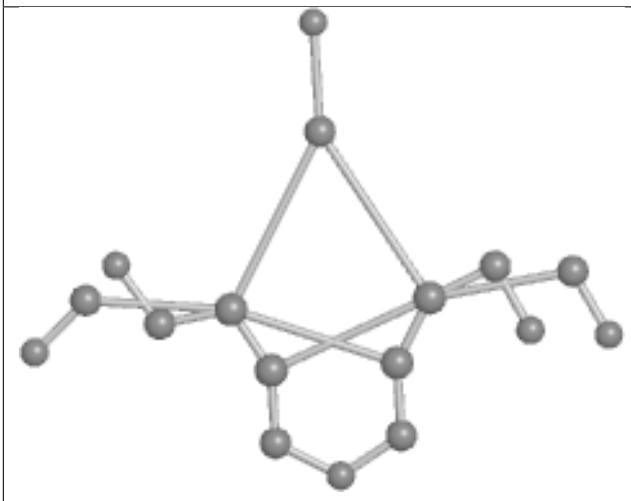
Ligand 402 B 700



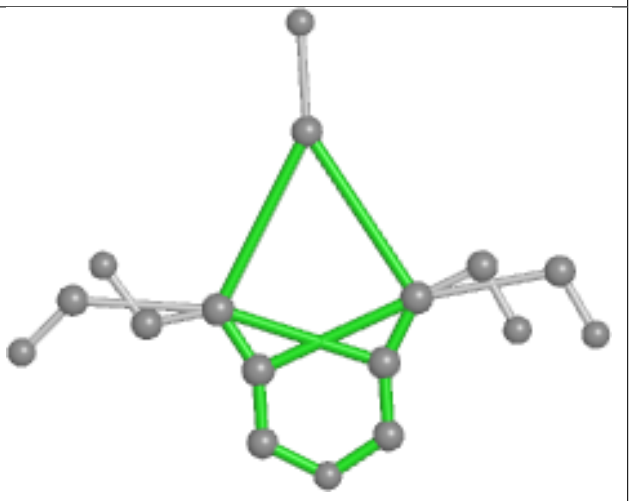
Bond lengths



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	479/674 (71%)	1.07	73 (15%)	2	1	54, 69, 90, 107	0
1	B	479/674 (71%)	1.11	91 (18%)	1	0	53, 70, 93, 109	0
All	All	958/1348 (71%)	1.09	164 (17%)	1	1	53, 69, 92, 109	0

The worst 5 of 164 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	144	ILE	7.9
1	A	134	TYR	7.2
1	A	353	LEU	6.9
1	A	140	VAL	6.7
1	A	133	SER	6.4

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

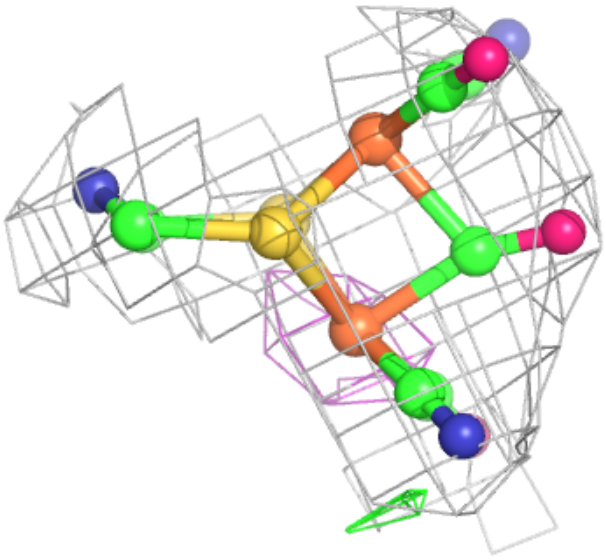
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

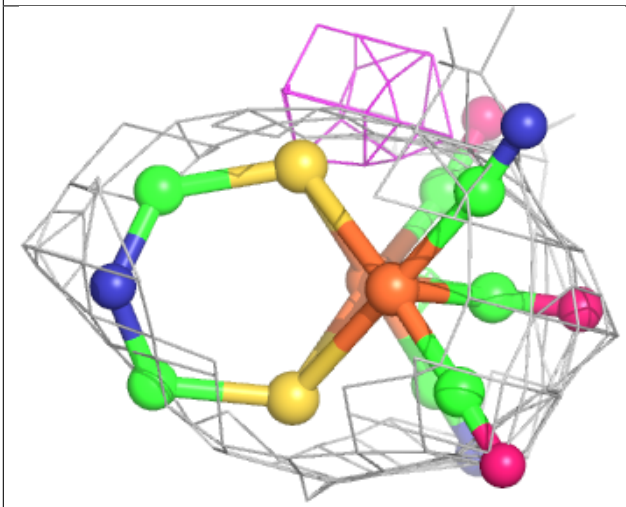
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	402	A	700	17/17	0.96	0.23	58,61,71,75	0
3	SF4	A	701	8/8	0.96	0.17	58,62,68,68	8
2	402	B	700	17/17	0.96	0.28	56,59,67,69	0
3	SF4	B	701	8/8	0.98	0.19	60,62,66,74	0

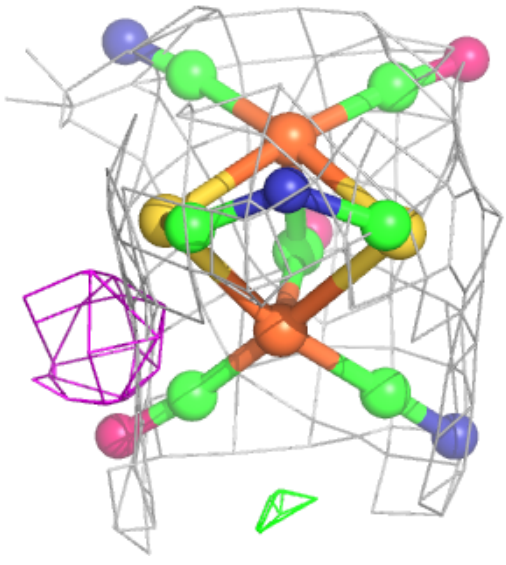
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 402 A 700:

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

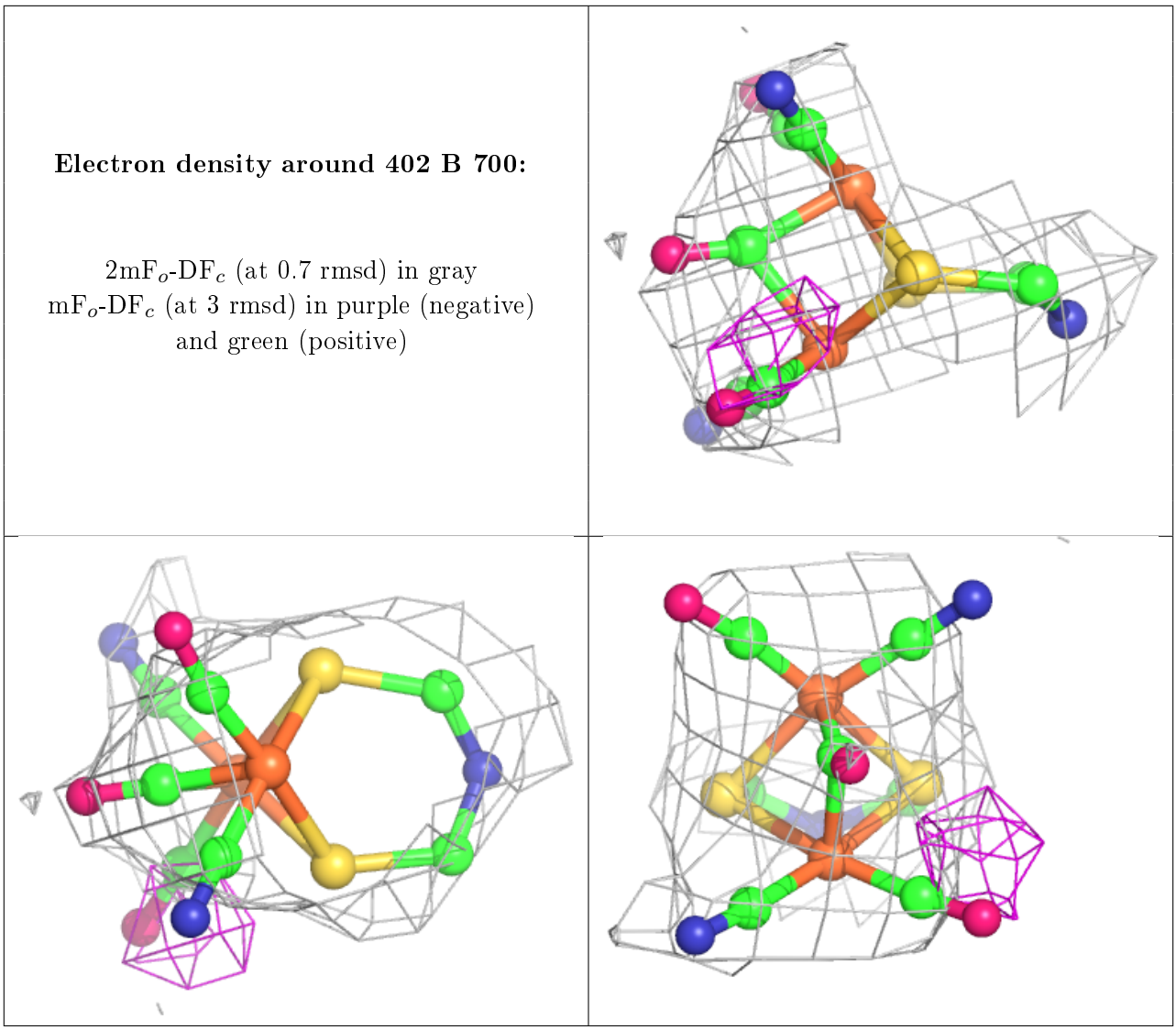






Electron density around 402 B 700:

$2mF_o$-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

There are no such residues in this entry.