



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2020 – 06:40 am BST

PDB ID : 5A4I  
Title : The mechanism of Hydrogen activation by NiFe-hydrogenases  
Authors : Evans, R.M.; Brooke, E.J.; Wehlin, S.A.M.; Nomerotskaia, E.; Sargent, F.;  
Carr, S.C.; Phillips, S.E.V.; Armstrong, F.A.  
Deposited on : 2015-06-10  
Resolution : 1.23 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](mailto: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 : 2.11  
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) : 2.11

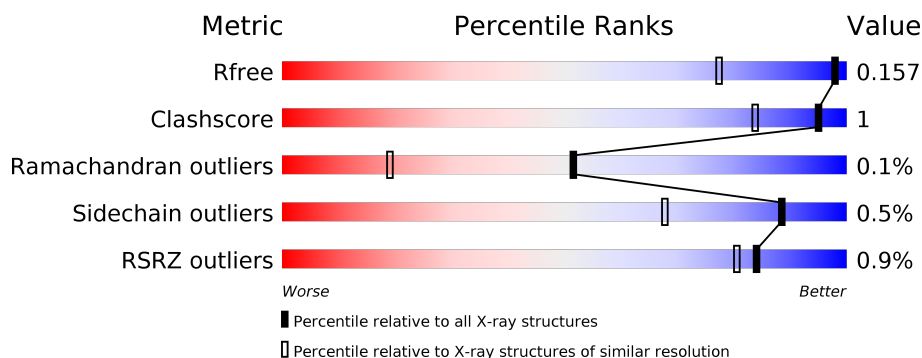
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.23 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	2024 (1.28-1.20)
Clashscore	141614	1007 (1.26-1.22)
Ramachandran outliers	138981	2053 (1.28-1.20)
Sidechain outliers	138945	2051 (1.28-1.20)
RSRZ outliers	127900	1987 (1.28-1.20)

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  $\geq 3$ , 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	L	582	<div> <div>%</div> <div> <div></div> <div>95%</div> <div>5%</div> </div> </div>
1	M	582	<div> <div>%</div> <div> <div></div> <div>95%</div> <div>5%</div> </div> </div>
2	S	335	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>21%</div> </div> </div>
2	T	335	<div> <div>%</div> <div> <div></div> <div>75%</div> <div>21%</div> </div> </div>

## 2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 14619 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 Hydrogenase-1 large chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	581	Total	C	N	O	S	0	29	0
			4717	2998	826	864	29			
1	M	581	Total	C	N	O	S	0	22	0
			4669	2972	811	858	28			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	574	ASN	ASP	conflict	UNP P0ACD8
M	574	ASN	ASP	conflict	UNP P0ACD8

- Molecule 2 is a protein called Hydrogenase-1 small chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	264	Total	C	N	O	S	0	9	0
			2092	1328	359	384	21			
2	T	264	Total	C	N	O	S	0	10	0
			2100	1334	362	383	21			

There are 16 discrepancies between the modelled and reference sequences:

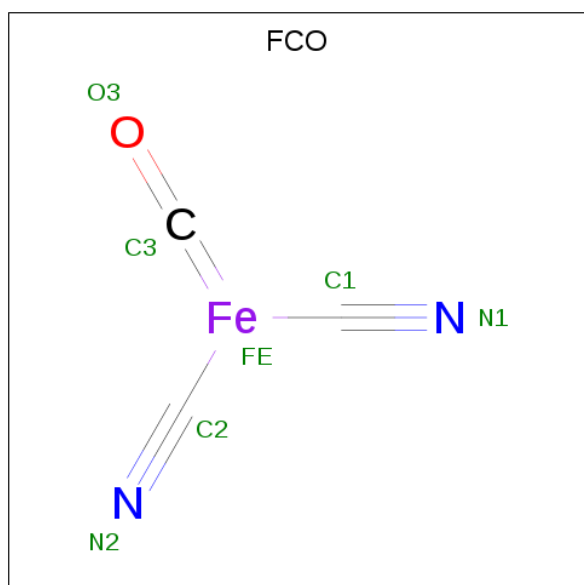
Chain	Residue	Modelled	Actual	Comment	Reference
S	328	ARG	-	expression tag	UNP P69739
S	329	SER	-	expression tag	UNP P69739
S	330	HIS	-	expression tag	UNP P69739
S	331	HIS	-	expression tag	UNP P69739
S	332	HIS	-	expression tag	UNP P69739
S	333	HIS	-	expression tag	UNP P69739
S	334	HIS	-	expression tag	UNP P69739
S	335	HIS	-	expression tag	UNP P69739
T	328	ARG	-	expression tag	UNP P69739
T	329	SER	-	expression tag	UNP P69739

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	330	HIS	-	expression tag	UNP P69739
T	331	HIS	-	expression tag	UNP P69739
T	332	HIS	-	expression tag	UNP P69739
T	333	HIS	-	expression tag	UNP P69739
T	334	HIS	-	expression tag	UNP P69739
T	335	HIS	-	expression tag	UNP P69739

- Molecule 3 is CARBONMONOXIDE-(DICYANO) IRON (three-letter code: FCO) (formula:  $C_3FeN_2O$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	L	1	Total	C	Fe	N	O	0	0
			7	3	1	2	1		
3	M	1	Total	C	Fe	N	O	0	0
			7	3	1	2	1		

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	1	Total	Ni	0	0
			1	1		
4	M	1	Total	Ni	0	0
			1	1		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	L	1	Total	Mg	0	0
			1	1		
5	M	1	Total	Mg	0	0
			1	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



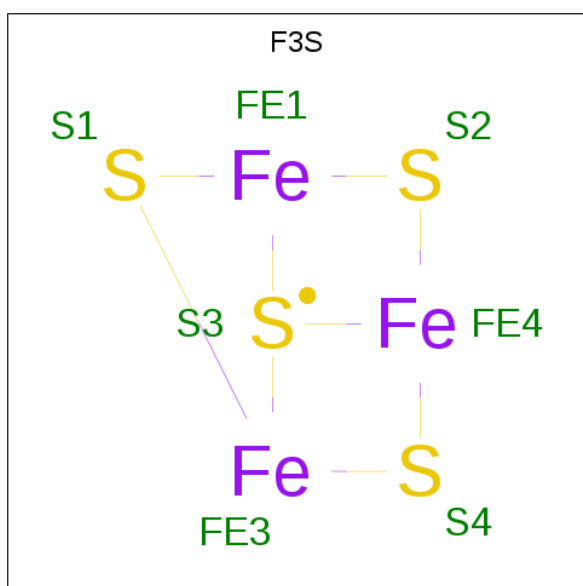
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	M	1	Total	O	S	0	0
			5	4	1		
6	S	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



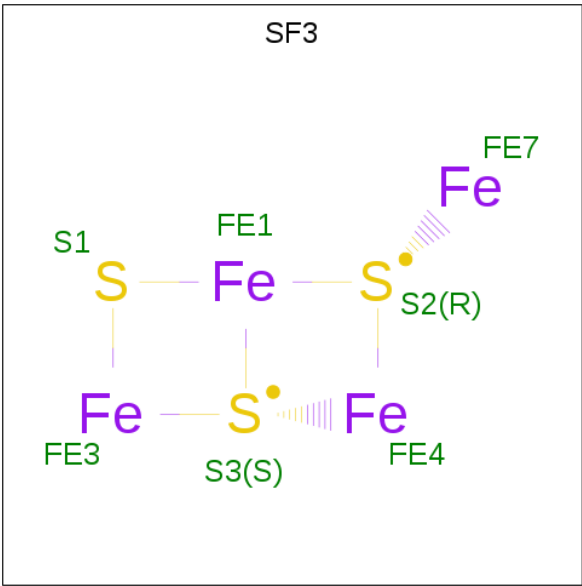
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	S	1	Total	Fe	S	0	0
			8	4	4		
7	T	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 8 is FE3-S4 CLUSTER (three-letter code: F3S) (formula:  $\text{Fe}_3\text{S}_4$ ).



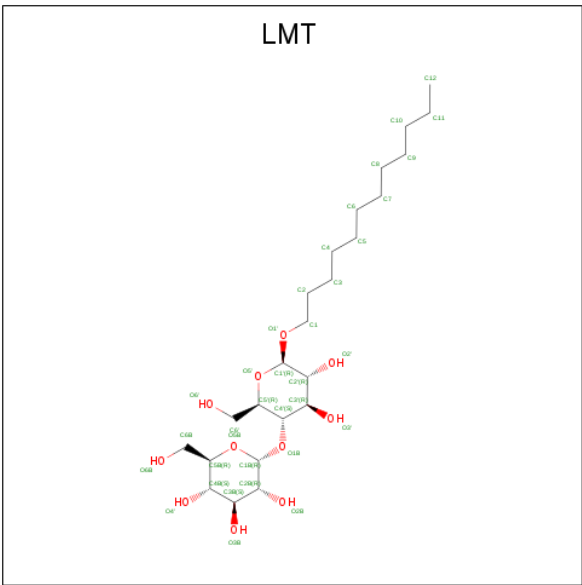
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	S	1	Total	Fe	S	0	0
			7	3	4		
8	T	1	Total	Fe	S	0	0
			7	3	4		

- Molecule 9 is FE4-S3 CLUSTER (three-letter code: SF3) (formula: Fe<sub>4</sub>S<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	S	1	Total	Fe	S	0	1
			8	5	3		
9	T	1	Total	Fe	S	0	1
			8	5	3		

- Molecule 10 is DODECYL-BETA-D-MALTOSIDE (three-letter code: LMT) (formula: C<sub>24</sub>H<sub>46</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	S	1	Total	C	O	0	0
			14	13	1		
10	T	1	Total	C	O	0	0
			14	13	1		

- Molecule 11 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	T	2	Total	Cl	0	0
			2	2		
11	S	2	Total	Cl	0	0
			2	2		

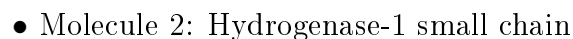
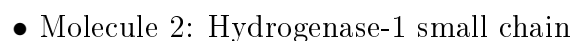
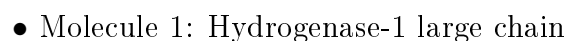
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	L	347	Total	O	0	0
			347	347		
12	M	367	Total	O	0	0
			367	367		
12	S	121	Total	O	0	0
			121	121		
12	T	100	Total	O	0	0
			100	100		





- Molecule 1: Hydrogenase-1 large chain



VAL	HIS	ALA	VAL	ALA	SER	ALA	VAL	ASP	GLN	ARG	ARG	HIS	ASN	GLN	GLN	PRO	THR	GLU	THR	GLU	HIS	GLN	PRO	GLY	ASN	GLU	ASP	LYS	GLN	ALA	ARG	SER	HIS	HIS	HIS	HIS	HIS	HIS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.78Å 97.58Å 182.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.19 – 1.23 32.17 – 1.23	Depositor EDS
% Data completeness (in resolution range)	97.2 (32.19-1.23) 97.2 (32.17-1.23)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.23Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, $R_{free}$	0.124 , 0.147 0.137 , 0.157	Depositor DCC
$R_{free}$ test set	23652 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	11.8	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.014 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	14619	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.94% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NI, CSO, CL, SF4, LMT, MG, SF3, F3S, SO4, FCO

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	L	0.70	1/4907 (0.0%)	0.87	10/6668 (0.1%)
1	M	0.71	1/4847 (0.0%)	0.86	15/6588 (0.2%)
2	S	0.78	2/2171 (0.1%)	1.00	10/2945 (0.3%)
2	T	0.75	1/2182 (0.0%)	0.96	7/2959 (0.2%)
All	All	0.73	5/14107 (0.0%)	0.90	42/19160 (0.2%)

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
2	S	0	1
2	T	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	260	ARG	CZ-NH1	7.27	1.42	1.33
2	S	260	ARG	CZ-NH1	7.07	1.42	1.33
1	L	561	GLU	CD-OE1	6.23	1.32	1.25
2	S	262	SER	CB-OG	5.91	1.50	1.42
1	M	245	GLU	CD-OE1	-5.30	1.19	1.25

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	260	ARG	NE-CZ-NH2	-12.56	114.02	120.30
2	S	197	ASP	CB-CG-OD1	11.80	128.92	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	197	ASP	CB-CG-OD1	11.45	128.61	118.30
2	S	260	ARG	NE-CZ-NH2	-11.03	114.79	120.30
2	S	260	ARG	NE-CZ-NH1	9.89	125.25	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	S	26	ARG	Sidechain
2	T	26	ARG	Sidechain

## 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	L	4717	0	4658	12	0
1	M	4669	0	4603	13	0
2	S	2092	0	2039	5	0
2	T	2100	0	2058	5	0
3	L	7	0	0	0	0
3	M	7	0	0	0	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
5	L	1	0	0	0	0
5	M	1	0	0	0	0
6	M	5	0	0	0	0
6	S	5	0	0	0	0
7	S	8	0	0	0	0
7	T	8	0	0	0	0
8	S	7	0	0	0	0
8	T	7	0	0	0	0
9	S	8	0	0	0	0
9	T	8	0	0	0	0
10	S	14	0	25	0	0
10	T	14	0	25	0	0
11	S	2	0	0	0	0
11	T	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	347	0	0	0	0
12	M	367	0	0	1	0
12	S	121	0	0	1	0
12	T	100	0	0	1	0
All	All	14619	0	13408	33	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 1.

The worst 5 of 33 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:S:61[B]:GLU:OE2	2:S:101[B]:ARG:NH1	1.60	1.34
2:S:234[A]:ARG:HD3	2:T:234[A]:ARG:NH1	1.68	1.08
1:L:275[B]:MET:HE2	1:L:456:ALA:HA	1.68	0.76
1:L:254[B]:MET:HA	1:L:254[B]:MET:HE2	1.77	0.66
1:L:378:ASP:OD1	1:L:379[A]:VAL:O	2.15	0.64

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	L	607/582 (104%)	591 (97%)	14 (2%)	2 (0%)	41	16
1	M	600/582 (103%)	587 (98%)	13 (2%)	0	100	100
2	S	271/335 (81%)	259 (96%)	12 (4%)	0	100	100
2	T	272/335 (81%)	259 (95%)	13 (5%)	0	100	100
All	All	1750/1834 (95%)	1696 (97%)	52 (3%)	2 (0%)	51	19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	381[A]	GLY
1	L	381[B]	GLY

### 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	L	507/480 (106%)	506 (100%)	1 (0%)	93	80
1	M	501/480 (104%)	499 (100%)	2 (0%)	91	77
2	S	226/274 (82%)	222 (98%)	4 (2%)	59	21
2	T	227/274 (83%)	224 (99%)	3 (1%)	69	33
All	All	1461/1508 (97%)	1451 (99%)	10 (1%)	88	59

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	S	16[B]	GLU
2	S	168	ARG
2	T	16[A]	GLU
2	S	16[A]	GLU
2	S	191	TYR

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

Mol	Chain	Res	Type
1	M	387	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](#)

2 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	CSO	M	79	1,3,4	3,6,7	0.61	0	0,6,8	0.00	-
1	CSO	L	79	1,3,4	3,6,7	0.76	0	0,6,8	0.00	-

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
1	CSO	M	79	1,3,4	-	0/1/5/7	-
1	CSO	L	79	1,3,4	-	0/1/5/7	-

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.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	M	79	CSO	3	0
1	L	79	CSO	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 8 are monoatomic - leaving 14 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
6	SO4	S	407	-	4,4,4	0.55	0	6,6,6	0.89	0
6	SO4	M	604	-	4,4,4	0.53	0	6,6,6	0.58	0
8	F3S	T	402	2	0,9,9	0.00	-	-		
10	LMT	S	404	-	13,13,36	0.48	0	12,12,47	0.96	0
9	SF3	T	403[C]	9,2	0,8,8	0.00	-	-		
9	SF3	S	403[B]	9,12,2	0,8,8	0.00	-	-		
9	SF3	T	403[B]	9,2	0,8,8	0.00	-	-		
3	FCO	M	601	1,12,4	0,6,6	0.00	-	-		
9	SF3	S	403[C]	9,12,2	0,8,8	0.00	-	-		
3	FCO	L	601	1,12,4	0,6,6	0.00	-	-		
7	SF4	S	401	2	0,12,12	0.00	-	-		
10	LMT	T	404	-	13,13,36	0.40	0	12,12,47	0.54	0
8	F3S	S	402	2	0,9,9	0.00	-	-		
7	SF4	T	401	2	0,12,12	0.00	-	-		

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
7	SF4	S	401	2	-	-	0/6/5/5
10	LMT	S	404	-	-	6/11/11/61	-
8	F3S	T	402	2	-	-	0/3/3/3
9	SF3	T	403[C]	9,2	-	-	0/2/2/2
9	SF3	S	403[B]	9,12,2	-	-	0/2/2/2
9	SF3	T	403[B]	9,2	-	-	0/2/2/2
9	SF3	S	403[C]	9,12,2	-	-	0/2/2/2
10	LMT	T	404	-	-	5/11/11/61	-
8	F3S	S	402	2	-	-	0/3/3/3
7	SF4	T	401	2	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

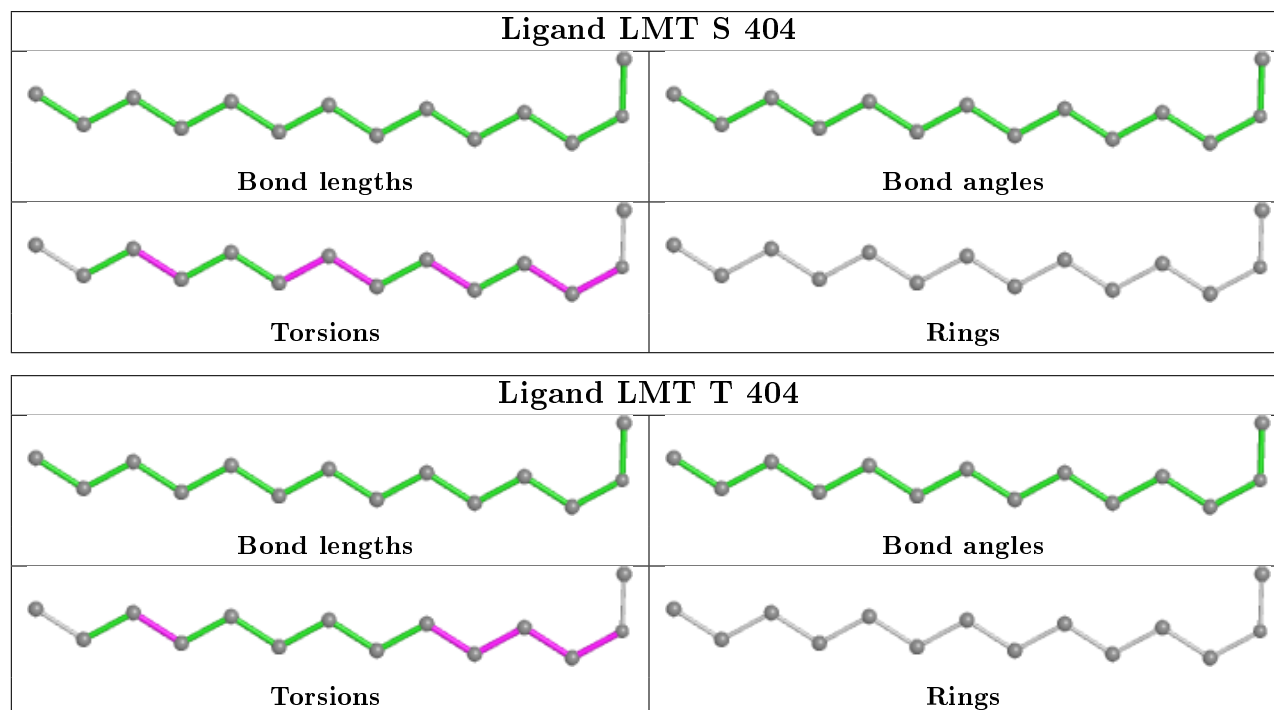
5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	S	404	LMT	C2-C3-C4-C5
10	S	404	LMT	O1'-C1-C2-C3
10	T	404	LMT	O1'-C1-C2-C3
10	S	404	LMT	C2-C1-O1'-C1'
10	S	404	LMT	C11-C10-C9-C8

There are no ring outliers.

No monomer is involved in short contacts.

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	L	580/582 (99%)	-0.43	7 (1%) 79 74	8, 13, 25, 49	0
1	M	580/582 (99%)	-0.49	3 (0%) 91 88	8, 12, 22, 34	0
2	S	264/335 (78%)	-0.50	4 (1%) 73 69	7, 11, 20, 51	0
2	T	264/335 (78%)	-0.45	2 (0%) 86 83	8, 12, 22, 55	0
All	All	1688/1834 (92%)	-0.46	16 (0%) 84 80	7, 12, 23, 55	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	4	LYS	5.2
2	T	267	VAL	4.7
1	L	2	SER	3.9
2	S	267	VAL	3.7
2	S	4	LYS	3.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	CSO	L	79	7/8	0.98	0.06	10,10,14,21	0
1	CSO	M	79	7/8	0.99	0.06	10,10,13,19	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

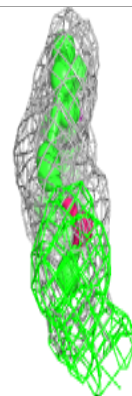
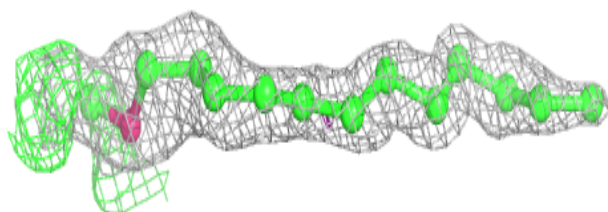
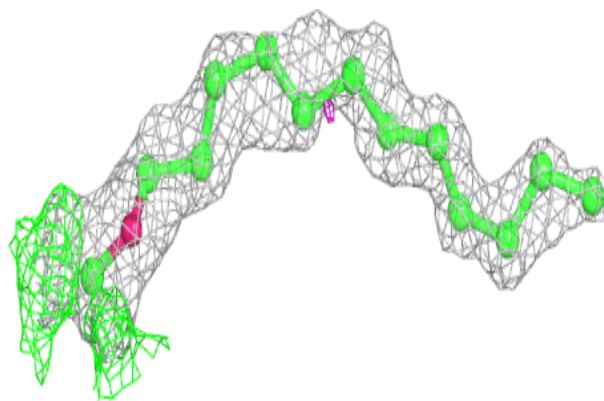
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
10	LMT	S	404	14/35	0.85	0.11	23,31,39,43	0
10	LMT	T	404	14/35	0.85	0.13	25,30,41,42	0
6	SO4	M	604	5/5	0.93	0.27	39,43,52,57	0
6	SO4	S	407	5/5	0.97	0.19	22,23,28,30	0
9	SF3	T	403[C]	7/7	0.99	0.04	10,11,14,22	1
9	SF3	T	403[B]	7/7	0.99	0.04	8,11,12,14	1
9	SF3	S	403[C]	7/7	0.99	0.04	9,11,13,20	1
9	SF3	S	403[B]	7/7	0.99	0.04	7,10,12,13	1
11	CL	S	406	1/1	1.00	0.04	22,22,22,22	0
4	NI	M	602	1/1	1.00	0.02	13,13,13,13	0
4	NI	L	602	1/1	1.00	0.02	13,13,13,13	0
11	CL	T	406	1/1	1.00	0.05	24,24,24,24	0
3	FCO	M	601	7/7	1.00	0.06	8,8,9,10	0
3	FCO	L	601	7/7	1.00	0.05	8,9,9,10	0
7	SF4	T	401	8/8	1.00	0.04	8,9,9,9	0
11	CL	T	405	1/1	1.00	0.03	14,14,14,14	0
7	SF4	S	401	8/8	1.00	0.03	8,9,9,9	0
8	F3S	T	402	7/7	1.00	0.04	9,9,10,10	0
5	MG	M	603	1/1	1.00	0.10	7,7,7,7	0
11	CL	S	405	1/1	1.00	0.03	13,13,13,13	0
8	F3S	S	402	7/7	1.00	0.04	9,9,10,10	0
5	MG	L	603	1/1	1.00	0.07	7,7,7,7	0

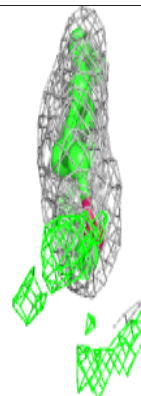
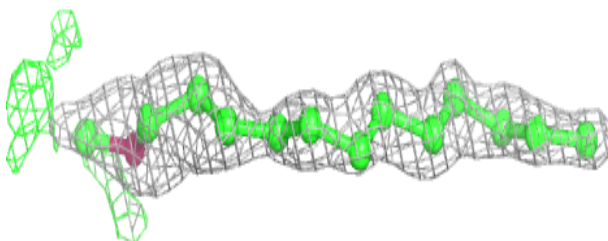
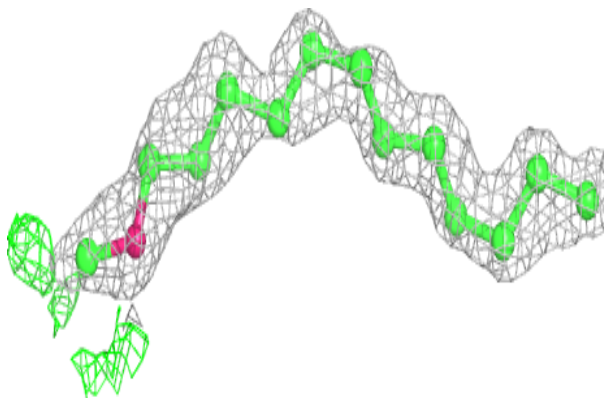
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around LMT S 404:**

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

**Electron density around LMT T 404:**

$2mF_o-DF_c$  (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.