



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 7, 2024 – 11:28 am GMT

PDB ID : 6FPW
Title : Structure of fully reduced Hydrogenase (Hyd-1)
Authors : Carr, S.B.; Armstrong, F.A.; Evans, R.M.
Deposited on : 2018-02-12
Resolution : 1.35 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

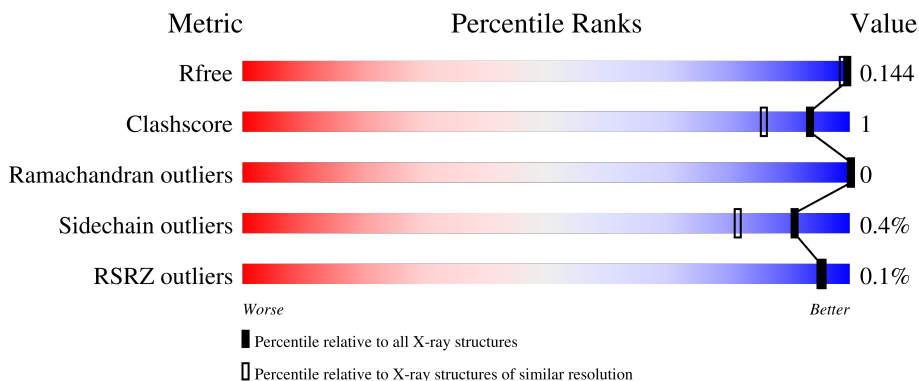
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.35 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	1509 (1.38-1.34)
Clashscore	141614	1551 (1.38-1.34)
Ramachandran outliers	138981	1530 (1.38-1.34)
Sidechain outliers	138945	1530 (1.38-1.34)
RSRZ outliers	127900	1487 (1.38-1.34)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	S	335	73% 5% 21%
1	T	335	75% 21%
2	L	582	95% 5%
2	M	582	96%

2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 28035 atoms, of which 13223 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 small chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	S	264	4107	1319	2033	356	378	21	80	6	0
1	T	264	4115	1321	2035	359	379	21	83	7	0

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
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 2 is a protein called Hydrogenase-1 large chain.

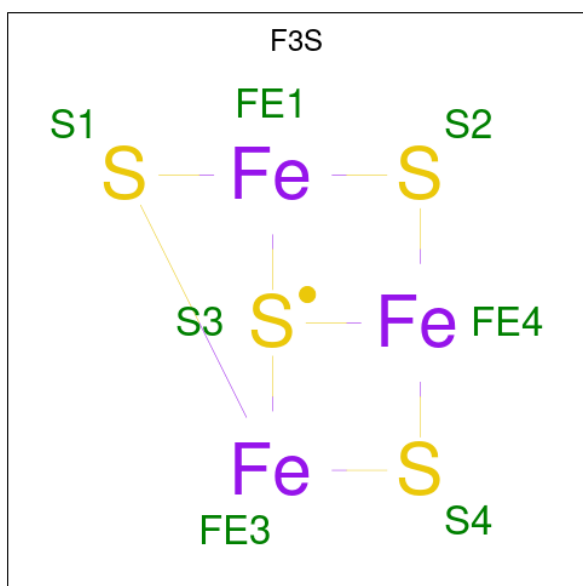
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	L	581	9209	2952	4555	817	857	28	244	18	0
2	M	581	9177	2944	4546	806	854	27	235	16	0

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



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

- Molecule 4 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe₃S₄).



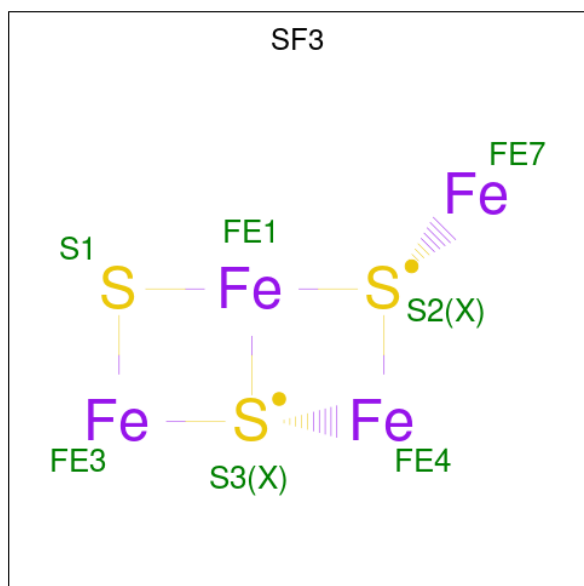
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	S	1	Total	Fe	S	0	0
			7	3	4		

Continued on next page...

Continued from previous page...

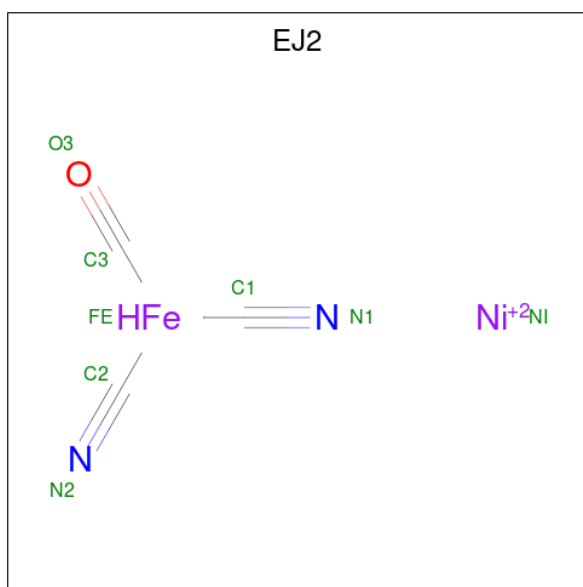
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
4	T	1	7	3	4	0	0

- Molecule 5 is FE4-S3 CLUSTER (three-letter code: SF3) (formula: Fe₄S₃).



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

- Molecule 6 is DODECYL-BETA-D-MALTOSIDE (three-letter code: LMT) (formula: C₂₄H₄₆O₁₁).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	
			Total	C	Fe	H	N	Ni			O
8	L	1	9	3	1	1	2	1	1	0	0
8	M	1	9	3	1	1	2	1	1	0	0

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	L	1	Total	O	S	0	0
			5	4	1		

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	S	191	Total	O	0	0
			191	191		
11	L	421	Total	O	0	0
			421	421		
11	T	198	Total	O	0	0
			198	198		
11	M	462	Total	O	0	0
			462	462		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.83Å 97.73Å 182.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.46 – 1.35 86.20 – 1.35	Depositor EDS
% Data completeness (in resolution range)	99.9 (91.46-1.35) 99.9 (86.20-1.35)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.48 (at 1.35Å)	Xtrriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.114 , 0.144 0.115 , 0.144	Depositor DCC
R_{free} test set	18090 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	12.9	Xtrriage
Anisotropy	0.134	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 46.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l	Xtrriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	28035	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.94% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: F3S, CL, MG, SF3, EJ2, SF4, SO4, LMT

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	S	0.83	2/2144 (0.1%)	1.06	14/2909 (0.5%)
1	T	0.80	2/2153 (0.1%)	0.92	8/2921 (0.3%)
2	L	0.74	2/4822 (0.0%)	0.86	14/6558 (0.2%)
2	M	0.72	1/4795 (0.0%)	0.86	12/6523 (0.2%)
All	All	0.76	7/13914 (0.1%)	0.90	48/18911 (0.3%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	28	GLU	C-N	-6.98	1.20	1.33
1	T	260	ARG	CZ-NH1	6.37	1.41	1.33
1	T	57	GLU	CG-CD	6.05	1.61	1.51
2	M	28	GLU	C-N	-5.92	1.22	1.33
2	L	561	GLU	CD-OE1	5.88	1.32	1.25

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	S	234	ARG	NE-CZ-NH1	-12.44	114.08	120.30
1	S	185[A]	ARG	NE-CZ-NH1	12.16	126.38	120.30
1	S	185[B]	ARG	NE-CZ-NH1	12.16	126.38	120.30
2	L	363	ARG	NE-CZ-NH2	11.31	125.96	120.30
2	L	363	ARG	NE-CZ-NH1	-10.40	115.10	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	S	2074	2033	2031	9	0
1	T	2080	2035	2032	10	0
2	L	4654	4555	4554	13	0
2	M	4631	4546	4546	6	0
3	S	8	0	0	0	0
3	T	8	0	0	0	0
4	S	7	0	0	0	0
4	T	7	0	0	0	0
5	S	7	0	0	0	0
5	T	7	0	0	0	0
6	S	14	26	25	0	0
6	T	14	26	25	1	0
7	S	3	0	0	0	0
7	T	3	0	0	0	0
8	L	8	1	0	0	0
8	M	8	1	0	0	0
9	L	1	0	0	0	0
9	M	1	0	0	0	0
10	L	5	0	0	0	0
11	L	421	0	0	5	0
11	M	462	0	0	0	0
11	S	191	0	0	1	0
11	T	198	0	0	1	0
All	All	14812	13223	13213	36	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 36 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:L:457[A]:HIS:CE1	11:L:701:HOH:O	2.20	0.92
1:T:168:ARG:CZ	11:T:503:HOH:O	2.24	0.84
2:L:280:LEU:HG	11:L:975:HOH:O	1.82	0.79
2:L:457[A]:HIS:ND1	11:L:701:HOH:O	2.13	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:L:457[A]:HIS:ND1	11:L:702:HOH:O	2.18	0.76

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	S	268/335 (80%)	257 (96%)	11 (4%)	0	100	100
1	T	269/335 (80%)	262 (97%)	7 (3%)	0	100	100
2	L	597/582 (103%)	582 (98%)	15 (2%)	0	100	100
2	M	595/582 (102%)	581 (98%)	14 (2%)	0	100	100
All	All	1729/1834 (94%)	1682 (97%)	47 (3%)	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	S	223/274 (81%)	221 (99%)	2 (1%)	78	53
1	T	223/274 (81%)	221 (99%)	2 (1%)	78	53
2	L	498/481 (104%)	497 (100%)	1 (0%)	93	84
2	M	496/481 (103%)	496 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1440/1510 (95%)	1435 (100%)	5 (0%)	91 83

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	S	191	TYR
1	S	260	ARG
2	L	524	ASP
1	T	191	TYR
1	T	242	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	L	328	ASN
1	T	244	GLN
2	M	261	GLN
2	M	387	GLN
2	M	467	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](#)

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 19 ligands modelled in this entry, 8 are monoatomic - leaving 11 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	LMT	T	404	-	13,13,36	0.40	0	12,12,47	0.37	0
6	LMT	S	404	-	13,13,36	0.45	0	12,12,47	0.54	0
5	SF3	S	403	1	0,8,8	-	-	-		
5	SF3	T	403	1	0,8,8	-	-	-		
4	F3S	S	402	1	0,9,9	-	-	-		
4	F3S	T	402	1	0,9,9	-	-	-		
10	SO4	L	603	-	4,4,4	0.19	0	6,6,6	0.45	0
3	SF4	S	401	1	0,12,12	-	-	-		
8	EJ2	M	601	2	1,6,6	1.38	0	-		
8	EJ2	L	601	2	1,6,6	0.18	0	-		
3	SF4	T	401	1	0,12,12	-	-	-		

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
6	LMT	T	404	-	-	4/11/11/61	-
6	LMT	S	404	-	-	6/11/11/61	-
5	SF3	S	403	1	-	-	0/2/2/2
5	SF3	T	403	1	-	-	0/2/2/2
4	F3S	S	402	1	-	-	0/3/3/3
4	F3S	T	402	1	-	-	0/3/3/3
3	SF4	S	401	1	-	-	0/6/5/5
3	SF4	T	401	1	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	S	404	LMT	O1'-C1-C2-C3
6	T	404	LMT	C1-C2-C3-C4
6	S	404	LMT	C2-C1-O1'-C1'

Continued on next page...

Continued from previous page...

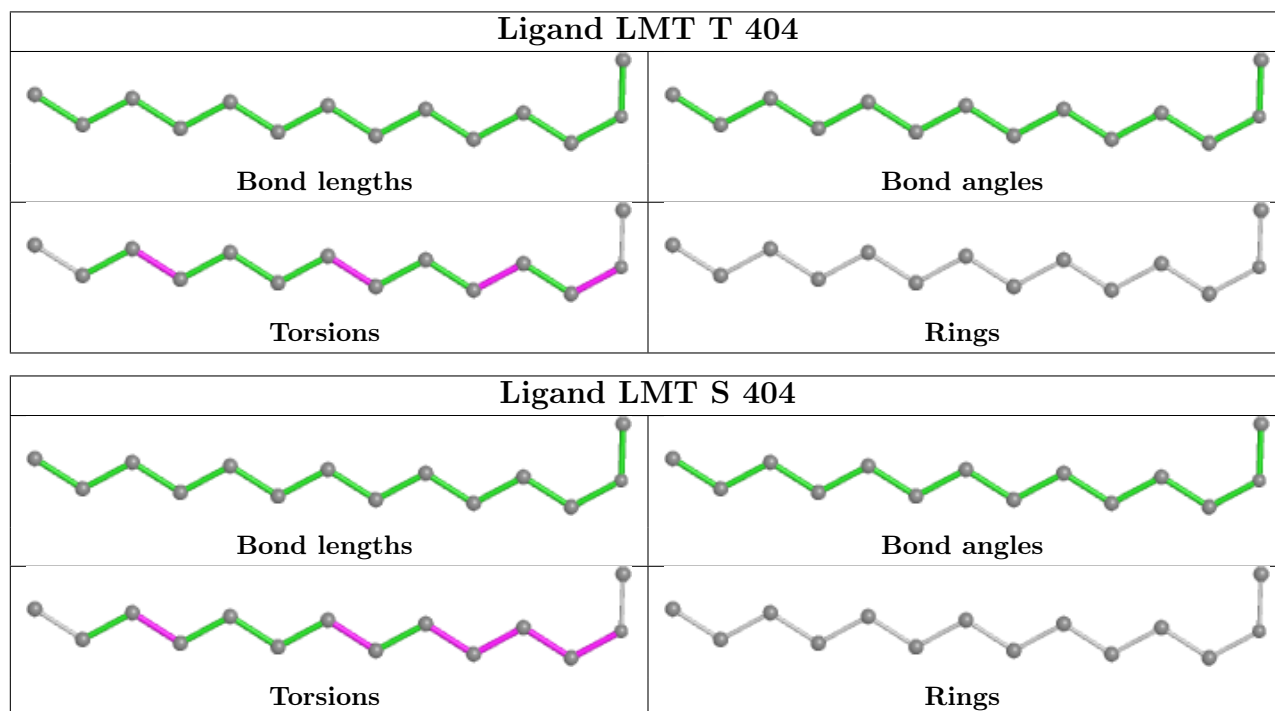
Mol	Chain	Res	Type	Atoms
6	T	404	LMT	C11-C10-C9-C8
6	S	404	LMT	C2-C3-C4-C5

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	T	404	LMT	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 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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	S	264/335 (78%)	-0.56	1 (0%) 92 93	8, 14, 24, 60	0
1	T	264/335 (78%)	-0.57	0 100 100	9, 14, 25, 49	0
2	L	581/582 (99%)	-0.67	0 100 100	9, 15, 29, 54	0
2	M	581/582 (99%)	-0.68	0 100 100	9, 15, 28, 42	0
All	All	1690/1834 (92%)	-0.64	1 (0%) 95 95	8, 15, 28, 60	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	S	4	LYS	2.8

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	LMT	S	404	14/35	0.61	0.27	0,24,34,34	40

Continued on next page...

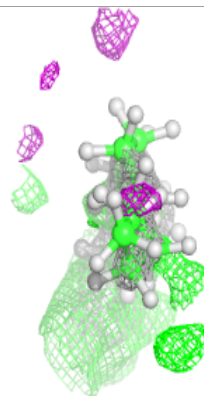
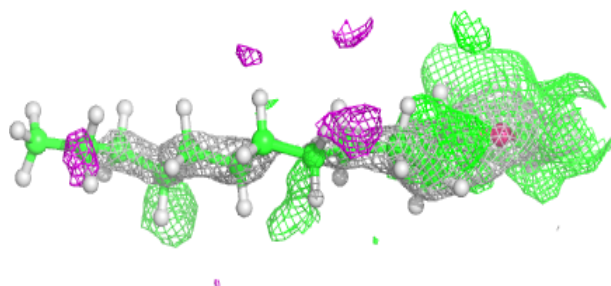
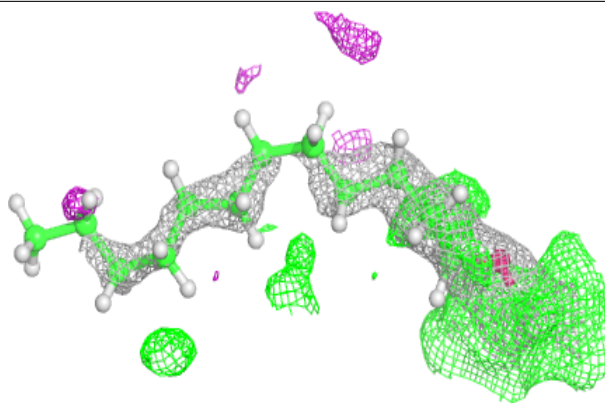
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	LMT	T	404	14/35	0.75	0.18	0,24,36,39	40
10	SO4	L	603	5/5	0.99	0.10	16,17,20,20	5
4	F3S	T	402	7/7	1.00	0.08	9,9,9,10	0
5	SF3	S	403	7/7	1.00	0.08	8,9,9,9	0
5	SF3	T	403	7/7	1.00	0.08	9,9,10,10	0
3	SF4	S	401	8/8	1.00	0.08	9,9,10,10	0
3	SF4	T	401	8/8	1.00	0.08	9,9,9,10	0
7	CL	S	405	1/1	1.00	0.05	15,15,15,15	0
7	CL	S	406	1/1	1.00	0.04	18,18,18,18	0
7	CL	S	407	1/1	1.00	0.04	16,16,16,16	0
7	CL	T	405	1/1	1.00	0.06	17,17,17,17	0
7	CL	T	406	1/1	1.00	0.03	20,20,20,20	0
7	CL	T	407	1/1	1.00	0.04	15,15,15,15	0
8	EJ2	L	601	8/8	1.00	0.07	9,9,9,10	0
8	EJ2	M	601	8/8	1.00	0.08	8,9,10,10	0
9	MG	L	602	1/1	1.00	0.08	9,9,9,9	0
9	MG	M	602	1/1	1.00	0.09	9,9,9,9	0
4	F3S	S	402	7/7	1.00	0.08	9,9,9,10	0

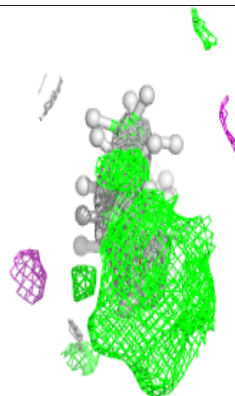
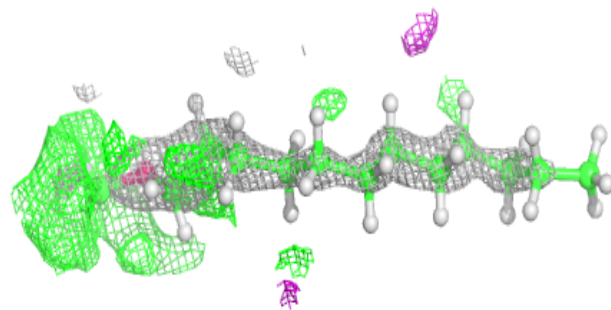
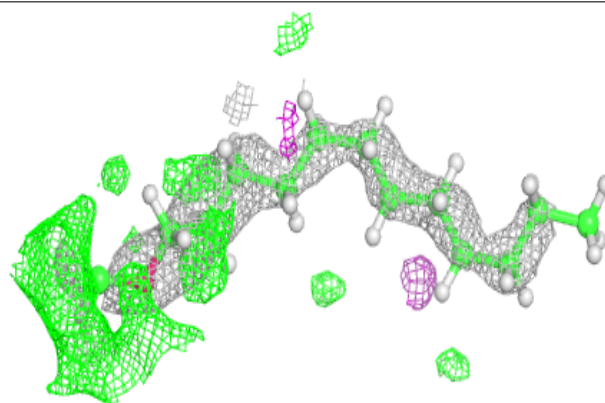
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

There are no such residues in this entry.