



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 18, 2026 – 11:09 pm BST

PDB ID : 6GAL / pdb_00006gal
Title : Structure of fully reduced Hydrogenase (Hyd-1) variant E28Q collected at pH 10
Authors : Carr, S.B.; Armstrong, F.A.; Evans, R.M.
Deposited on : 2018-04-11
Resolution : 1.25 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 : 4-5-2 with Phenix2.0
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.015 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

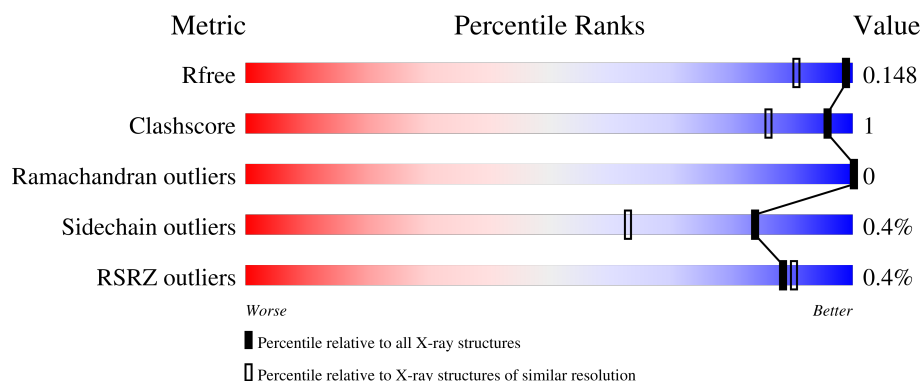
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.25 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}	180053	1693 (1.28-1.24)
Clashscore	190562	1730 (1.28-1.24)
Ramachandran outliers	187476	1695 (1.28-1.24)
Sidechain outliers	187428	1694 (1.28-1.24)
RSRZ outliers	180081	1693 (1.28-1.24)

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	% 75% 21%
1	T	335	% 76% 21%
2	L	582	% 97% .
2	M	582	% 97% .

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 28048 atoms, of which 13158 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
1	S	264	Total	C	H	N	O	S	81	7	0
			4109	1320	2034	354	379	22			
1	T	266	Total	C	H	N	O	S	84	9	0
			4164	1336	2057	363	387	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
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

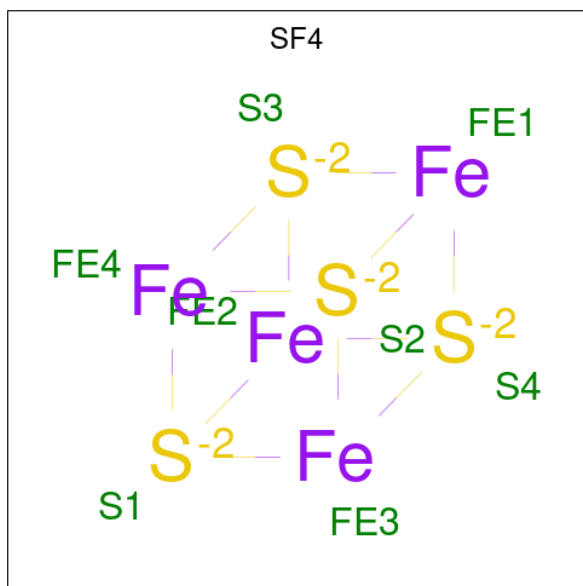
- Molecule 2 is a protein called Hydrogenase-1 large chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	L	581	Total	C	H	N	O	S	236	11	0
			9123	2926	4512	806	851	28			
2	M	581	Total	C	H	N	O	S	239	13	0
			9161	2934	4534	813	852	28			

There are 2 discrepancies between the modelled and reference sequences:

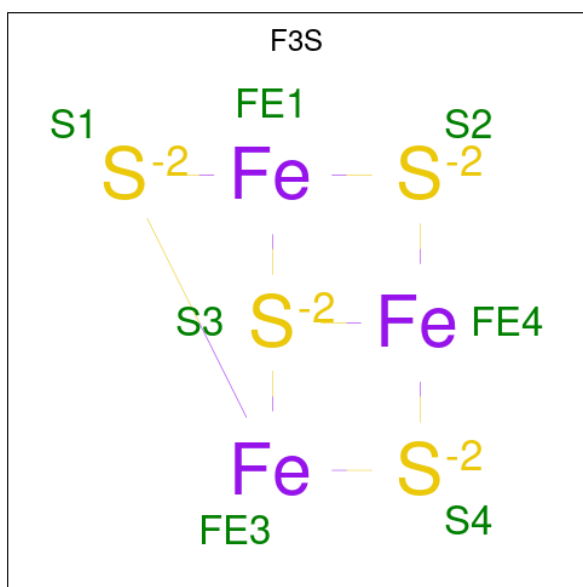
Chain	Residue	Modelled	Actual	Comment	Reference
L	28	GLN	GLU	conflict	UNP P0ACD8
M	28	GLN	GLU	conflict	UNP P0ACD8

- Molecule 3 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe_4S_4).



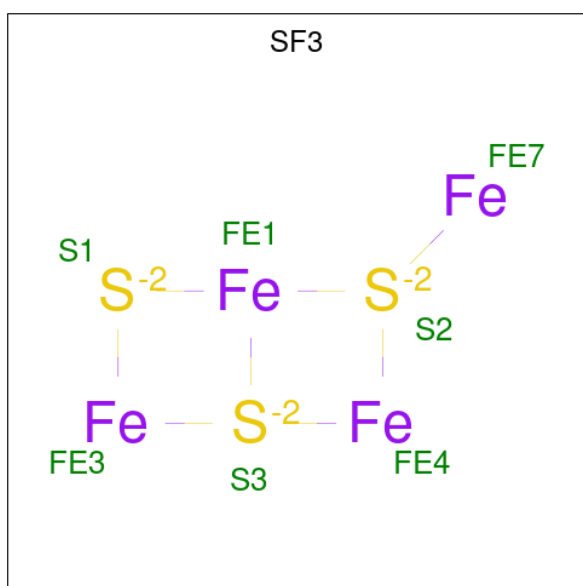
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	S	1	Total	Fe	S	0	0
			8	4	4		
3	T	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 4 is FE3-S4 CLUSTER (CCD ID: F3S) (formula: Fe_3S_4).



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

- Molecule 5 is FE4-S3 CLUSTER (CCD ID: SF3) (formula: Fe_4S_3).

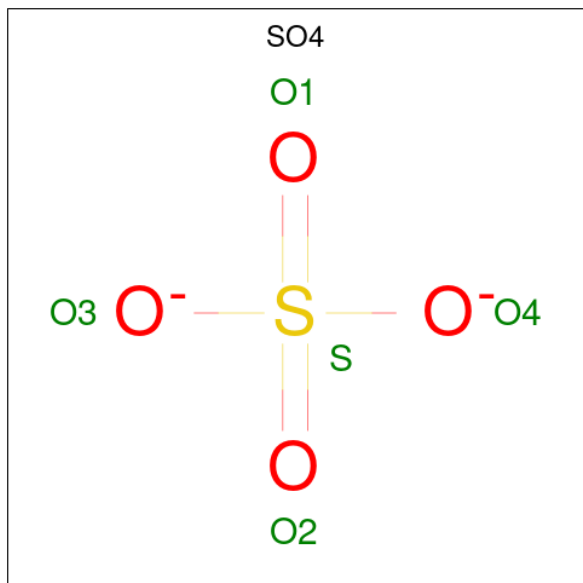


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	S	1	Total	Fe	S	0	0
			7	4	3		
5	T	1	Total	Fe	S	0	0
			7	4	3		

- Molecule 6 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

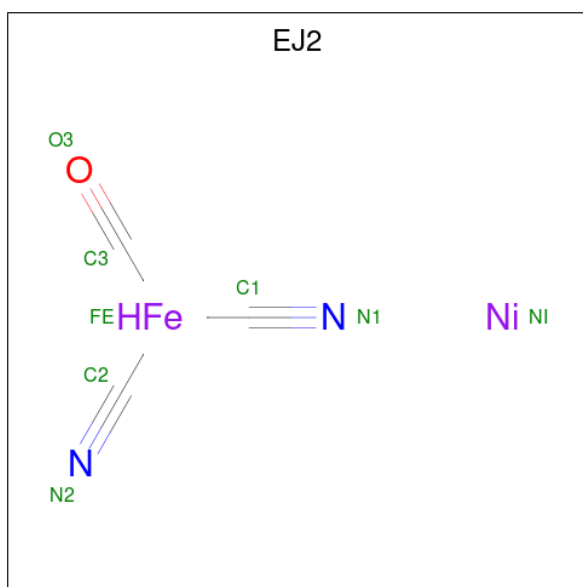
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	S	1	Total	Cl	0	0
			1	1		
6	T	1	Total	Cl	0	0
			1	1		

- Molecule 7 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

- Molecule 8 is NI-FE REDUCED ACTIVE CENTER (CCD ID: EJ2) (formula: C₃HFeN₂NiO).

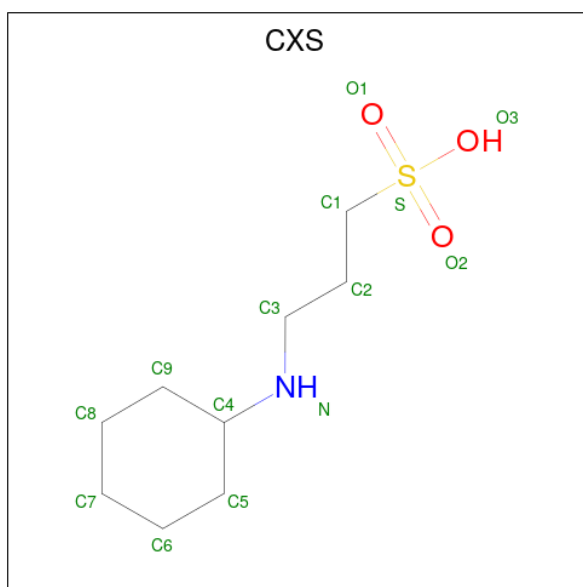


Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
8	L	1	Total	C	Fe	H	N	Ni	O	0	0
			9	3	1	1	2	1	1		
8	M	1	Total	C	Fe	H	N	Ni	O	0	0
			9	3	1	1	2	1	1		

- Molecule 9 is MAGNESIUM ION (CCD ID: 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 3-CYCLOHEXYL-1-PROPYLSULFONIC ACID (CCD ID: CXS) (formula: C₉H₁₉NO₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
10	M	1	Total	C	H	N	O	S	
			33	9	19	1	3	1	
								1	0

- Molecule 11 is LITHIUM ION (CCD ID: LI) (formula: Li).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	M	2	Total	Li		
			2	2	0	0

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	S	186	Total	O		
			186	186	0	0
12	L	511	Total	O		
			511	511	0	0
12	T	223	Total	O		
			223	223	0	0
12	M	460	Total	O		
			460	460	0	0

- Molecule 1: Hydrogenase-1 small chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	94.91Å 95.22Å 183.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	92.12 – 1.25 91.95 – 1.25	Depositor EDS
% Data completeness (in resolution range)	95.7 (92.12-1.25) 95.7 (91.95-1.25)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 1.25Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.120 , 0.147 0.121 , 0.148	Depositor DCC
R_{free} test set	22000 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å ²)	9.1	Xtriage
Anisotropy	0.211	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 40.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	28048	wwPDB-VP
Average B, all atoms (Å ²)	11.0	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 25.30 % 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 3.2521e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

5.1 Standard geometry

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

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	0.84	0/2145	0.89	0/2911
1	T	0.87	1/2180 (0.0%)	0.91	2/2960 (0.1%)
2	L	0.82	1/4754 (0.0%)	0.84	0/6469
2	M	0.79	1/4773 (0.0%)	0.85	1/6491 (0.0%)
All	All	0.82	3/13852 (0.0%)	0.86	3/18831 (0.0%)

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

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	574	ASP	CA-CB	7.95	1.58	1.52
2	M	574	ASP	C-O	-7.94	1.20	1.23
1	T	185	ARG	CZ-NH2	-5.37	1.26	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	530	VAL	N-CA-CB	-6.09	107.63	111.83
1	T	234	ARG	NE-CZ-NH1	5.54	127.04	121.50
1	T	187	HIS	CA-CB-CG	5.20	119.00	113.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	S	26	ARG	Sidechain
1	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	S	2075	2034	2032	8	0
1	T	2107	2057	2053	3	0
2	L	4611	4512	4500	7	1
2	M	4627	4534	4523	10	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	1	0	0	0	0
6	T	1	0	0	0	0
7	M	5	0	0	0	0
7	S	5	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	M	14	19	19	0	0
11	M	2	0	0	0	0
12	L	511	0	0	3	0
12	M	460	0	0	0	0
12	S	186	0	0	1	1
12	T	223	0	0	1	0
All	All	14890	13158	13127	28	1

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 28 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:S:175[B]:MET:HA	1:S:175[B]:MET:HE2	1.63	0.81
2:M:272:ASN:HD21	2:M:460:GLN:HG3	1.47	0.79
2:M:164:PHE:CE2	2:M:168[A]:ASN:ND2	2.60	0.69
2:L:269[B]:ASP:OD1	12:L:701:HOH:O	2.11	0.66
1:S:197:ASP:HB2	12:T:522:HOH:O	1.96	0.64

All (1) 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)
2:L:332:ASN:HD22	12:S:675:HOH:O[4_545]	1.55	0.05

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	269/335 (80%)	258 (96%)	11 (4%)	0	100	100
1	T	273/335 (82%)	260 (95%)	13 (5%)	0	100	100
2	L	590/582 (101%)	572 (97%)	18 (3%)	0	100	100
2	M	592/582 (102%)	576 (97%)	16 (3%)	0	100	100
All	All	1724/1834 (94%)	1666 (97%)	58 (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	224/274 (82%)	223 (100%)	1 (0%)	84	63
1	T	227/274 (83%)	226 (100%)	1 (0%)	84	63
2	L	491/481 (102%)	489 (100%)	2 (0%)	84	63
2	M	493/481 (102%)	490 (99%)	3 (1%)	78	52
All	All	1435/1510 (95%)	1428 (100%)	7 (0%)	84	56

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	T	242	PRO
2	M	254[A]	MET
2	M	524	ASP
2	M	254[B]	MET
2	L	524	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	T	68	ASN
2	M	40	GLN
2	M	562	GLN
2	M	387	GLN
2	M	544	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 17 ligands modelled in this entry, 6 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
7	SO4	S	405	-	4,4,4	0.49	0	6,6,6	0.33	0
7	SO4	M	601	-	4,4,4	0.17	0	6,6,6	0.65	0
8	EJ2	M	602	2	1,6,6	1.88	0	-		
3	SF4	T	401	1	0,12,12	-	-	-		
5	SF3	S	403	1	0,8,8	-	-	-		
8	EJ2	L	601	2	1,6,6	1.41	0	-		
10	CXS	M	604	-	14,14,14	1.04	1 (7%)	18,18,18	1.52	2 (11%)
3	SF4	S	401	1	0,12,12	-	-	-		
4	F3S	T	402	1	0,9,9	-	-	-		
5	SF3	T	403	1	0,8,8	-	-	-		
4	F3S	S	402	1	0,9,9	-	-	-		

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	T	401	1	-	-	0/6/5/5
5	SF3	S	403	1	-	-	0/2/2/2
10	CXS	M	604	-	-	0/8/16/16	0/1/1/1
3	SF4	S	401	1	-	-	0/6/5/5
4	F3S	T	402	1	-	-	0/3/3/3
5	SF3	T	403	1	-	-	0/2/2/2
4	F3S	S	402	1	-	-	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	M	604	CXS	C1-S	-3.12	1.73	1.77

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	M	604	CXS	C9-C4-C5	3.31	116.55	110.82
10	M	604	CXS	C3-N-C4	2.95	119.93	114.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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%)	-0.80	2 (0%) 82 83	4, 9, 21, 58	7 (2%)
1	T	266/335 (79%)	-0.84	3 (1%) 78 79	4, 9, 18, 44	9 (3%)
2	L	581/582 (99%)	-0.92	0 100 100	4, 9, 18, 36	11 (1%)
2	M	581/582 (99%)	-0.87	1 (0%) 91 92	4, 9, 20, 52	13 (2%)
All	All	1692/1834 (92%)	-0.87	6 (0%) 88 90	4, 9, 19, 58	40 (2%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	T	3	ASN	3.5
2	M	2	SER	2.8
1	S	4	LYS	2.7
1	S	267	VAL	2.5
1	T	267	VAL	2.4

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
7	SO4	S	405	5/5	0.93	0.09	36,37,49,51	0
11	LI	M	606	1/1	0.94	0.09	15,15,15,15	0
11	LI	M	605	1/1	0.98	0.12	12,12,12,12	0
7	SO4	M	601	5/5	0.98	0.14	12,13,16,18	5
10	CXS	M	604	14/14	0.99	0.03	7,8,10,10	1
5	SF3	T	403	7/7	1.00	0.01	7,7,7,7	0
6	CL	S	404	1/1	1.00	0.03	13,13,13,13	0
6	CL	T	404	1/1	1.00	0.03	13,13,13,13	0
3	SF4	S	401	8/8	1.00	0.01	7,7,7,7	0
3	SF4	T	401	8/8	1.00	0.01	7,8,8,8	0
8	EJ2	L	601	8/8	1.00	0.01	6,6,8,9	0
8	EJ2	M	602	8/8	1.00	0.02	6,7,8,9	0
9	MG	L	602	1/1	1.00	0.03	4,4,4,4	0
9	MG	M	603	1/1	1.00	0.04	5,5,5,5	0
4	F3S	S	402	7/7	1.00	0.01	7,7,7,8	0
4	F3S	T	402	7/7	1.00	0.01	7,7,8,8	0
5	SF3	S	403	7/7	1.00	0.01	7,7,7,8	0

6.5 Other polymers [i](#)

There are no such residues in this entry.