Full wwPDB X-ray Structure Validation Report ⓘ

Jun 14, 2020 – 08:05 am BST

PDB ID	:	1WUJ
Title	:	Three-Dimensional Structure Of The Ni-B State Of [Nife]Hydrogenase From Desulufovibrio Vulgaris Miyazaki F
Authors	:	Ogata, H.; Hirota, S.; Nakahara, A.; Komori, H.; Shibata, N.; Kato, T.; Kano, K.; Higuchi, Y.
Deposited on	:	2004-12-07
Resolution	:	1.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ⓘ) were used in the production of this report:

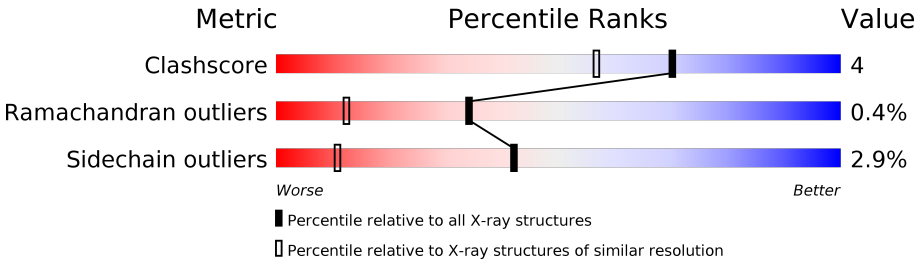
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	S	267	87% 11%
2	L	534	90% 10%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7064 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 Periplasmic [NiFe] hydrogenase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	S	267	2019	1282	342	378	17	0	0	0

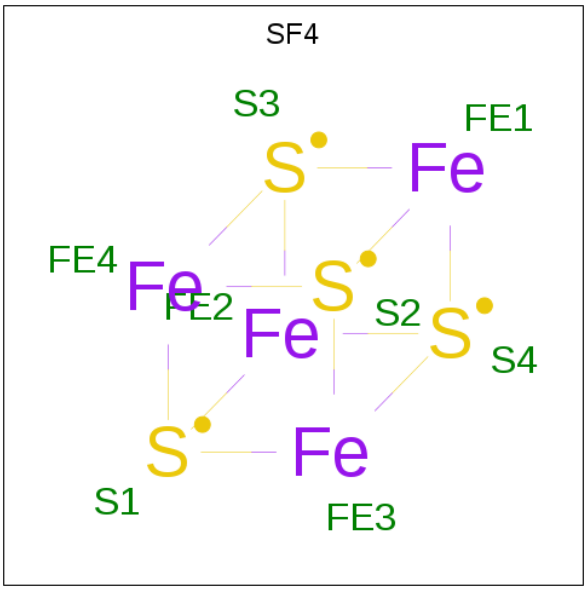
- Molecule 2 is a protein called Periplasmic [NiFe] hydrogenase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	534	4178	2674	725	764	15	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

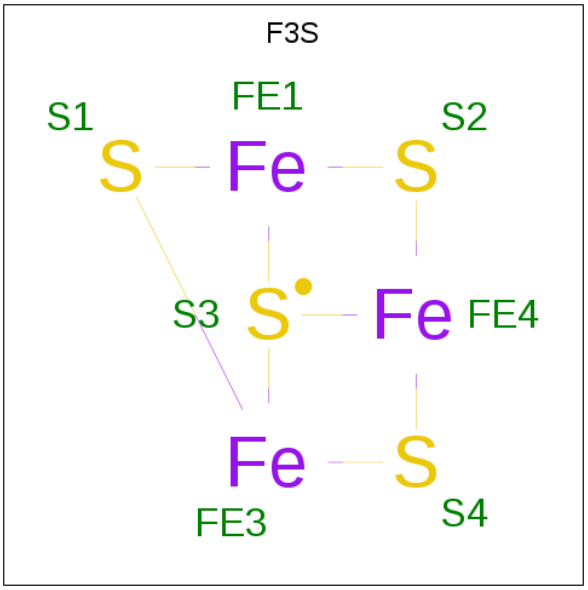
Chain	Residue	Modelled	Actual	Comment	Reference
L	514	LYS	ASN	SEE REMARK 999	UNP P21852
L	515	LEU	VAL	SEE REMARK 999	UNP P21852

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



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

- Molecule 4 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).

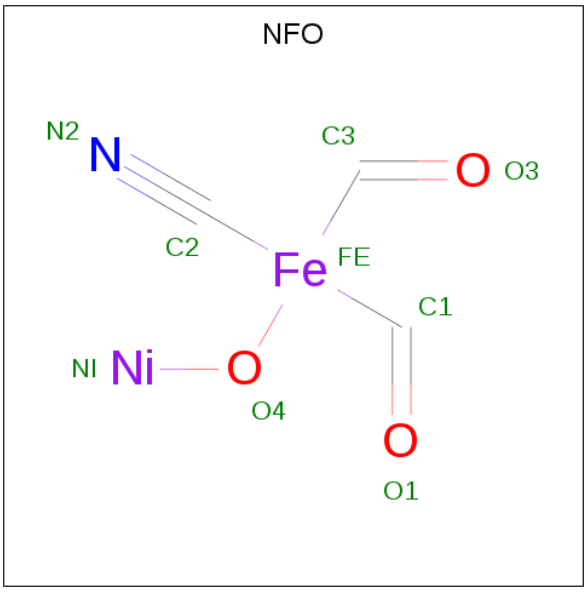


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	S	1	Total Fe S 7 3 4	0	0

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

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

- Molecule 6 is NI-FE OXIDIZED ACTIVE CENTER (three-letter code: NFO) (formula: $C_3H_2FeNNiO_3$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	L	1	Total	C	Fe	N	Ni	O	0	0
			9	3	1	1	1	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	S	287	Total	O	0	0
			287	287		
7	L	547	Total	O	0	0
			547	547		

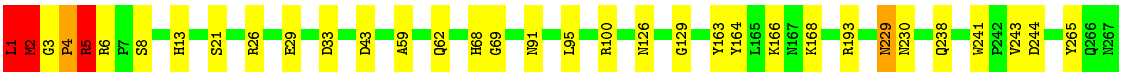
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

Note EDS was not executed.

- Molecule 1: Periplasmic [NiFe] hydrogenase small subunit

Chain S:



- Molecule 2: Periplasmic [NiFe] hydrogenase large subunit

Chain L:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	98.16Å 125.92Å 66.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.40	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.40)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.115 , 0.162	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7064	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NFO, MG, F3S, SF4, CSO

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	0.60	0/2075	1.30	23/2830 (0.8%)
2	L	0.58	1/4281 (0.0%)	1.32	32/5820 (0.5%)
All	All	0.59	1/6356 (0.0%)	1.32	55/8650 (0.6%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	552	HIS	C-O	5.82	1.34	1.23

All (55) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	224	ARG	NE-CZ-NH2	22.45	131.52	120.30
2	L	509	ARG	NE-CZ-NH1	18.75	129.67	120.30
1	S	26	ARG	NE-CZ-NH2	16.15	128.38	120.30
1	S	1	LEU	O-C-N	10.74	139.88	122.70
2	L	224	ARG	NH1-CZ-NH2	-10.72	107.61	119.40
2	L	233	ASN	CB-CG-OD1	10.43	142.46	121.60
2	L	430	ARG	NE-CZ-NH1	9.79	125.19	120.30
1	S	163	TYR	CB-CG-CD1	9.23	126.54	121.00
2	L	68	ARG	NE-CZ-NH1	9.13	124.87	120.30
2	L	424	LEU	C-N-CA	8.93	141.05	122.30
1	S	2	MET	C-N-CA	8.52	140.19	122.30
1	S	100	ARG	NE-CZ-NH2	8.33	124.47	120.30
2	L	273	TYR	CB-CG-CD1	8.29	125.97	121.00
1	S	26	ARG	NH1-CZ-NH2	-8.06	110.53	119.40
2	L	509	ARG	NE-CZ-NH2	-7.93	116.33	120.30
2	L	78	GLN	CG-CD-OE1	7.93	137.46	121.60
2	L	38	ARG	NE-CZ-NH2	-7.74	116.43	120.30
2	L	430	ARG	NE-CZ-NH2	-7.54	116.53	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	S	229	ASN	O-C-N	-7.43	110.82	122.70
2	L	195	HIS	CG-ND1-CE1	7.37	118.51	108.20
1	S	33	ASP	CB-CG-OD1	7.29	124.87	118.30
2	L	424	LEU	O-C-N	-7.25	110.88	123.20
2	L	234	PRO	C-N-CA	7.11	139.48	121.70
2	L	369	ARG	NE-CZ-NH1	6.71	123.66	120.30
1	S	229	ASN	C-N-CA	6.64	138.29	121.70
2	L	190	TYR	CB-CG-CD2	6.62	124.97	121.00
2	L	426	ARG	NE-CZ-NH1	6.60	123.60	120.30
2	L	85	THR	C-N-CA	6.53	138.02	121.70
2	L	426	ARG	NE-CZ-NH2	-6.31	117.14	120.30
1	S	129	GLY	O-C-N	-6.29	112.63	122.70
1	S	163	TYR	CB-CG-CD2	-6.18	117.29	121.00
1	S	164	TYR	CB-CG-CD2	-6.18	117.29	121.00
2	L	287	TRP	CA-CB-CG	-6.03	102.25	113.70
1	S	100	ARG	CD-NE-CZ	5.93	131.90	123.60
1	S	100	ARG	NE-CZ-NH1	-5.93	117.34	120.30
2	L	196	PRO	O-C-N	-5.85	113.35	122.70
1	S	29	GLU	OE1-CD-OE2	5.80	130.27	123.30
2	L	32	ARG	NE-CZ-NH1	5.78	123.19	120.30
1	S	163	TYR	CG-CD1-CE1	5.76	125.91	121.30
1	S	26	ARG	CG-CD-NE	-5.75	99.72	111.80
2	L	160	ARG	NE-CZ-NH1	-5.64	117.48	120.30
2	L	254	ARG	NE-CZ-NH1	-5.61	117.50	120.30
1	S	244	ASP	CB-CG-OD2	-5.57	113.29	118.30
1	S	129	GLY	C-N-CA	5.53	135.53	121.70
1	S	95	LEU	O-C-N	-5.40	114.06	122.70
2	L	120	TYR	CA-CB-CG	5.37	123.61	113.40
2	L	369	ARG	NE-CZ-NH2	-5.37	117.61	120.30
1	S	95	LEU	C-N-CA	5.27	134.88	121.70
1	S	238	GLN	CG-CD-OE1	5.25	132.10	121.60
2	L	196	PRO	C-N-CA	5.21	134.72	121.70
2	L	395	TYR	CG-CD2-CE2	5.11	125.39	121.30
1	S	193	ARG	NE-CZ-NH2	-5.11	117.75	120.30
2	L	324	PHE	CB-CG-CD1	-5.03	117.28	120.80
2	L	110	TYR	CB-CG-CD2	-5.00	118.00	121.00
2	L	273	TYR	CG-CD1-CE1	5.00	125.30	121.30

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	2019	0	1949	27	0
2	L	4178	0	4127	24	0
3	S	16	0	0	0	0
4	S	7	0	0	0	0
5	L	1	0	0	0	0
6	L	9	0	0	1	0
7	L	547	0	0	7	0
7	S	287	0	0	2	0
All	All	7064	0	6076	44	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.

All (44) 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:2:MET:HA	2:L:182:GLN:HG2	1.53	0.89
1:S:2:MET:HB2	2:L:182:GLN:HE21	1.40	0.86
1:S:2:MET:HE3	1:S:68:HIS:HB2	1.63	0.81
1:S:2:MET:HE2	7:S:5210:HOH:O	1.85	0.76
1:S:1:LEU:HD21	1:S:59:ALA:O	1.90	0.70
2:L:176:THR:O	2:L:180:THR:HG23	1.91	0.70
1:S:5:ARG:HA	1:S:5:ARG:NH1	2.06	0.70
1:S:166:LYS:HD3	7:S:5490:HOH:O	1.91	0.69
2:L:175:LYS:O	2:L:179:GLU:HG3	1.98	0.64
1:S:2:MET:HB3	1:S:43:ASP:OD1	1.98	0.63
1:S:5:ARG:HH11	1:S:5:ARG:HA	1.66	0.60
2:L:211:HIS:HE1	7:L:5242:HOH:O	1.87	0.58
1:S:5:ARG:HB3	2:L:182:GLN:OE1	2.05	0.56
1:S:2:MET:HG2	1:S:8:SER:HB2	1.88	0.55
2:L:211:HIS:HD2	2:L:276:ASP:OD2	1.90	0.55
2:L:36:HIS:HD2	7:L:5087:HOH:O	1.90	0.54
1:S:1:LEU:HA	2:L:187:THR:OG1	2.09	0.53
2:L:233:ASN:HD22	2:L:233:ASN:C	2.12	0.53
1:S:2:MET:HA	2:L:182:GLN:CG	2.33	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:S:5:ARG:HB2	7:L:5523:HOH:O	2.09	0.53
1:S:2:MET:CE	1:S:68:HIS:HB2	2.34	0.52
2:L:75:HIS:HD2	7:L:5015:HOH:O	1.92	0.51
1:S:1:LEU:HB3	2:L:187:THR:OG1	2.10	0.51
1:S:1:LEU:HD22	1:S:43:ASP:HB3	1.94	0.50
1:S:3:GLY:O	1:S:4:PRO:O	2.30	0.50
1:S:2:MET:HE1	1:S:69:GLY:O	2.13	0.49
1:S:265:TYR:OH	2:L:75:HIS:HE1	1.96	0.48
2:L:20:SER:HA	7:L:5689:HOH:O	2.14	0.47
1:S:3:GLY:HA2	1:S:68:HIS:CD2	2.50	0.46
1:S:1:LEU:HD11	1:S:62:GLN:HG2	1.98	0.46
2:L:291:GLY:HA2	2:L:521:SER:O	2.16	0.45
2:L:85:THR:OG1	2:L:235:HIS:HD2	2.00	0.45
1:S:241:TRP:CH2	1:S:243:VAL:HB	2.51	0.44
2:L:500:VAL:CG1	2:L:501:PRO:HD2	2.48	0.44
2:L:467:LYS:NZ	7:L:5491:HOH:O	2.51	0.43
2:L:498:LEU:N	2:L:498:LEU:HD12	2.33	0.43
2:L:468:GLN:NE2	7:L:5621:HOH:O	2.51	0.43
1:S:1:LEU:HD12	1:S:1:LEU:N	2.33	0.42
1:S:13:HIS:HE1	1:S:21:SER:OG	2.01	0.42
1:S:5:ARG:HD3	1:S:5:ARG:HH11	1.66	0.42
2:L:544:ASP:N	2:L:545:PRO:HD3	2.35	0.41
1:S:5:ARG:HB3	1:S:6:ARG:H	1.81	0.40
2:L:211:HIS:CD2	2:L:276:ASP:OD2	2.73	0.40
2:L:546:CSO:CB	6:L:1004:NFO:O4	2.65	0.40

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	265/267 (99%)	256 (97%)	7 (3%)	2 (1%)	19 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	L	531/534 (99%)	519 (98%)	11 (2%)	1 (0%)	47 21
All	All	796/801 (99%)	775 (97%)	18 (2%)	3 (0%)	34 12

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	S	4	PRO
2	L	235	HIS
1	S	5	ARG

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	213/213 (100%)	205 (96%)	8 (4%)	33 6
2	L	437/437 (100%)	426 (98%)	11 (2%)	47 14
All	All	650/650 (100%)	631 (97%)	19 (3%)	42 11

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	S	1	LEU
1	S	2	MET
1	S	5	ARG
1	S	91	ASN
1	S	126	ASN
1	S	168	LYS
1	S	229	ASN
1	S	230	ASN
2	L	45	ASN
2	L	49	LYS
2	L	50	ASN
2	L	132	HIS
2	L	152	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	L	161	LYS
2	L	233	ASN
2	L	279	VAL
2	L	457	ASN
2	L	468	GLN
2	L	470	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	S	13	HIS
1	S	45	HIS
1	S	68	HIS
1	S	126	ASN
1	S	139	HIS
1	S	190	GLN
1	S	230	ASN
1	S	266	GLN
1	S	267	ASN
2	L	36	HIS
2	L	50	ASN
2	L	75	HIS
2	L	78	GLN
2	L	113	ASN
2	L	132	HIS
2	L	188	ASN
2	L	205	ASN
2	L	211	HIS
2	L	233	ASN
2	L	235	HIS
2	L	334	GLN
2	L	390	GLN
2	L	446	GLN
2	L	451	ASN
2	L	476	ASN
2	L	513	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
2	CSO	L	546	2,6	3,6,7	0.51	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
2	CSO	L	546	2,6	-	1/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	L	546	CSO	N-CA-CB-SG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L	546	CSO	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 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	$\#\|Z\| > 2$
3	SF4	S	1002	1	0,12,12	0.00	-	-		
4	F3S	S	1003	1	0,9,9	0.00	-	-		
3	SF4	S	1001	1	0,12,12	0.00	-	-		
6	NFO	L	1004	2	1,8,8	5.45	1 (100%)	-		

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	S	1002	1	-	-	0/6/5/5
4	F3S	S	1003	1	-	-	0/3/3/3
3	SF4	S	1001	1	-	-	0/6/5/5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	1004	NFO	C2-N2	-5.45	1.02	1.13

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	L	1004	NFO	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.