wwPDB X-ray Structure Validation Summary Report (i)

Nov 7, 2023 – 06:51 PM JST

PDB ID	:	8W6X
Title	:	Neutron structure of [NiFe]-hydrogenase from D. vulgaris Miyazaki F in its oxidized state
Authors	:	Hiromoto, T.; Tamada, T.
Deposited on	:	2023-08-30
Resolution	:	1.04 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

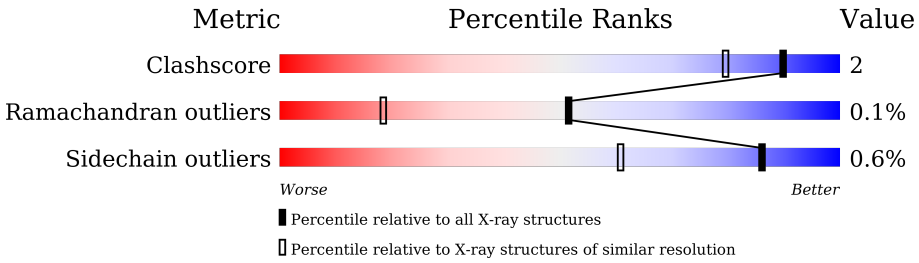
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.04 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	1677 (1.10-0.98)
Ramachandran outliers	138981	1591 (1.10-0.98)
Sidechain outliers	138945	1589 (1.10-0.98)

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 <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	L	534	98%
2	S	266	95%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 17342 atoms, of which 6520 are hydrogens and 2939 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.

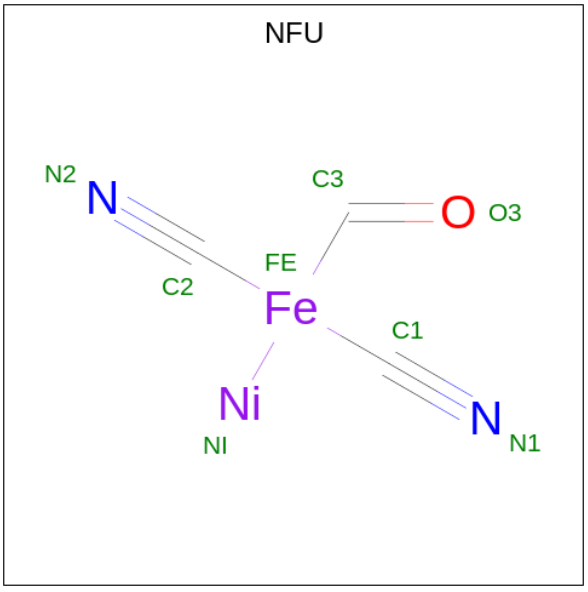
- Molecule 1 is a protein called Periplasmic [NiFe] hydrogenase large subunit.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	D	H	N	O	S			
1	L	533	9993	2942	938	4452	799	844	18	0	501	0

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

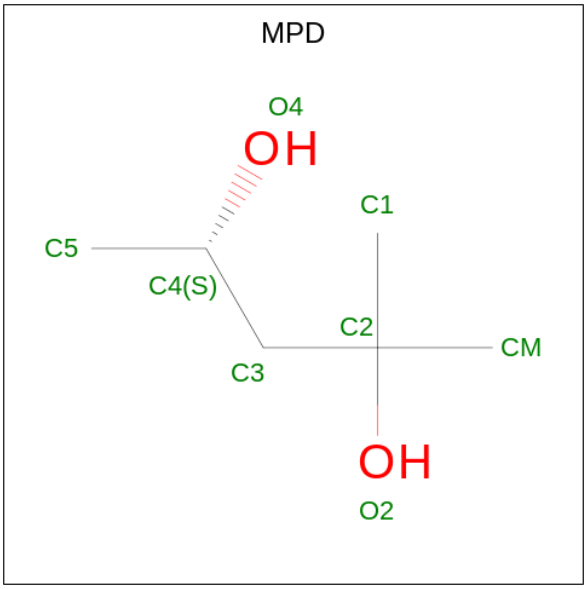
Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	D	H	N	O	S			
2	S	266	4704	1399	421	2068	375	421	20	0	241	0

- Molecule 3 is formyl[bis(hydrocyanato-1kappaC)]ironnickel(Fe-Ni) (three-letter code: NFU) (formula: C_3HFeN_2NiO) (labeled as "Ligand of Interest" by depositor).



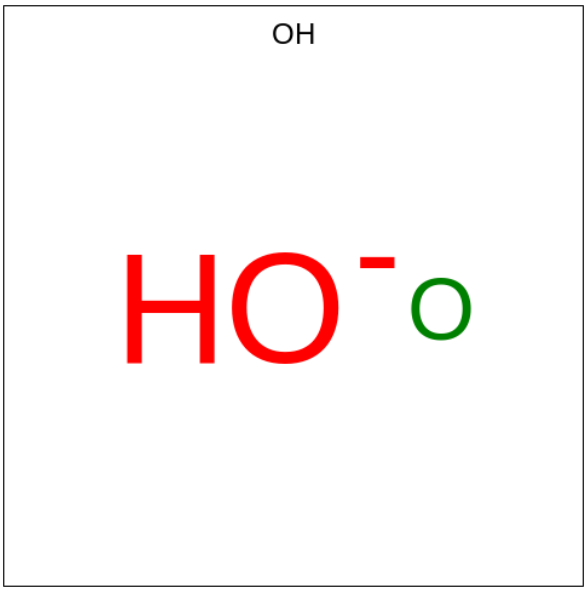
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Fe	N	Ni	O		
3	L	1	9	3	1	2	2	1	0	1

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	L	1	Total	C	D	O	0	0
			22	6	14	2		
4	L	1	Total	C	D	O	0	0
			22	6	14	2		
4	L	1	Total	C	D	O	0	0
			21	6	13	2		
4	L	1	Total	C	D	O	0	0
			22	6	14	2		
4	L	1	Total	C	D	O	0	0
			21	6	13	2		
4	L	1	Total	C	D	O	0	0
			22	6	14	2		
4	S	1	Total	C	D	O	0	0
			22	6	14	2		
4	S	1	Total	C	D	O	0	0
			22	6	14	2		
4	S	1	Total	C	D	O	0	0
			22	6	14	2		

- Molecule 5 is HYDROXIDE ION (three-letter code: OH) (formula: HO) (labeled as "Ligand of Interest" by depositor).



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

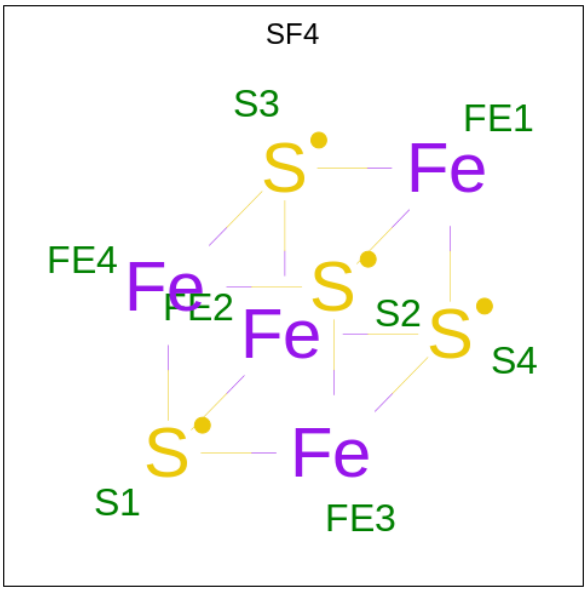
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

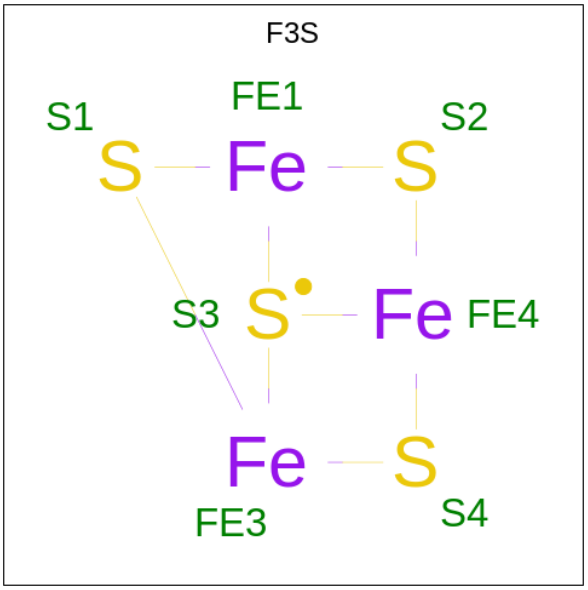
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	L	1	Total Cl 1 1	0	0

- Molecule 8 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



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

- Molecule 9 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).



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

- Molecule 10 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	L	548	Total 1575	D 958	O 617	0	69
10	S	294	Total 839	D 498	O 341	0	47

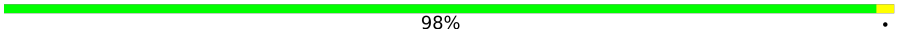
3 Residue-property plots (i)

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. 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 failed to run properly.

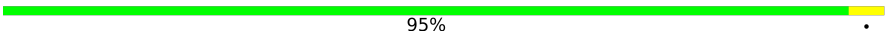
- Molecule 1: Periplasmic [NiFe] hydrogenase large subunit

Chain L:



- Molecule 2: Periplasmic [NiFe] hydrogenase small subunit

Chain S:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.70Å 98.48Å 126.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.69 – 1.04	Depositor
% Data completeness (in resolution range)	99.8 (41.69-1.04)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.19 (at 1.04Å)	Xtriage
Refinement program	PHENIX (1.15.2_3472)	Depositor
R, R_{free}	0.143 , 0.147	Depositor
Wilson B-factor (Å^2)	10.1	Xtriage
Anisotropy	0.526	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	17342	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: F3S, NFU, CL, MPD, MG, CSO, OH, SF4

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	0.55	0/8524	0.78	12/11577 (0.1%)
2	S	0.65	2/4018 (0.0%)	1.05	2/5483 (0.0%)
All	All	0.58	2/12542 (0.0%)	0.88	14/17060 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	22[A]	GLU	CD-OE2	5.21	1.31	1.25
2	S	22[B]	GLU	CD-OE2	5.21	1.31	1.25

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	L	430[A]	ARG	NE-CZ-NH1	7.75	124.18	120.30
1	L	430[B]	ARG	NE-CZ-NH1	7.75	124.18	120.30
1	L	59[A]	ARG	NE-CZ-NH2	-5.67	117.47	120.30
1	L	59[B]	ARG	NE-CZ-NH2	-5.67	117.47	120.30
1	L	59[A]	ARG	NE-CZ-NH1	5.63	123.12	120.30

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	L	5541	4452	1161	8	0
2	S	2636	2068	551	10	0
3	L	9	0	0	0	0
4	L	130	0	84	4	0
4	S	66	0	42	0	0
5	L	1	0	0	0	0
6	L	1	0	0	0	0
7	L	1	0	0	0	0
8	S	16	0	0	0	0
9	S	7	0	0	0	0
10	L	1575	0	0	5	0
10	S	839	0	0	1	0
All	All	10822	6520	1838	24	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 2.

The worst 5 of 24 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:228[B]:MET:HE2	10:S:1385:HOH:O	1.55	1.00
2:S:5[B]:ARG:NH2	2:S:5[B]:ARG:HG2	2.07	0.63
2:S:5[B]:ARG:HG2	2:S:5[B]:ARG:HH21	1.55	0.61
4:L:605:MPD:H52	4:L:605:MPD:O2	2.01	0.56
2:S:5[B]:ARG:HH21	2:S:5[B]:ARG:CG	2.16	0.50

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	1048/534 (196%)	1019 (97%)	27 (3%)	2 (0%)	47 18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	S	514/266 (193%)	504 (98%)	10 (2%)	0	100 100
All	All	1562/800 (195%)	1523 (98%)	37 (2%)	2 (0%)	51 18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	231[A]	ALA
1	L	231[B]	ALA

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	863/437 (198%)	861 (100%)	2 (0%)	93 74
2	S	410/212 (193%)	405 (99%)	5 (1%)	71 38
All	All	1273/649 (196%)	1266 (100%)	7 (0%)	86 65

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

Mol	Chain	Res	Type
2	S	5[A]	ARG
2	S	5[B]	ARG
2	S	168[B]	LYS
2	S	168[A]	LYS
2	S	2	MET

Sometimes 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)

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	$\#\|Z\| > 2$
1	CSO	L	546[B]	1,3	3,6,7	0.68	0	0,6,8	-	-

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	L	546[B]	1,3	-	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
1	L	546[B]	CSO	N-CA-CB-SG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 1 is modelled with single atom and 2 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	$\#\|Z\| > 2$
3	NFU	L	601[A]	5,1	2,7,7	0.64	0	-		
4	MPD	L	604	-	7,7,7	0.83	0	9,10,10	1.05	1 (11%)
4	MPD	L	602	-	7,7,7	0.70	0	9,10,10	1.24	1 (11%)
4	MPD	L	605	-	7,7,7	0.68	0	9,10,10	0.81	0
4	MPD	S	1004	-	7,7,7	0.62	0	9,10,10	0.65	0
9	F3S	S	1003	2	0,9,9	-	-	-		
3	NFU	L	601[B]	1	2,7,7	0.64	0	-		
4	MPD	S	1005	-	7,7,7	0.54	0	9,10,10	0.62	0
4	MPD	L	606	-	7,7,7	0.67	0	9,10,10	1.19	1 (11%)
4	MPD	L	607	-	7,7,7	0.73	0	9,10,10	0.80	0
8	SF4	S	1001	2	0,12,12	-	-	-		
4	MPD	S	1006	-	7,7,7	0.91	0	9,10,10	2.28	4 (44%)
4	MPD	L	603	-	7,7,7	0.61	0	9,10,10	1.16	0
8	SF4	S	1002	2	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
4	MPD	L	602	-	-	1/5/5/5	-
4	MPD	L	604	-	-	1/5/5/5	-
4	MPD	L	605	-	-	2/5/5/5	-
4	MPD	S	1004	-	-	0/5/5/5	-
9	F3S	S	1003	2	-	-	0/3/3/3
4	MPD	S	1005	-	-	0/5/5/5	-
4	MPD	L	606	-	-	0/5/5/5	-
4	MPD	L	607	-	-	0/5/5/5	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	SF4	S	1001	2	-	-	0/6/5/5
4	MPD	S	1006	-	-	1/5/5/5	-
4	MPD	L	603	-	-	1/5/5/5	-
8	SF4	S	1002	2	-	-	0/6/5/5

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	S	1006	MPD	O2-C2-C1	-4.43	93.87	108.08
4	L	602	MPD	CM-C2-C1	-3.26	103.78	110.57
4	S	1006	MPD	CM-C2-C1	2.91	116.64	110.57
4	L	606	MPD	CM-C2-C1	-2.87	104.59	110.57
4	S	1006	MPD	O4-C4-C3	-2.86	99.80	111.36

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	L	603	MPD	C2-C3-C4-C5
4	L	605	MPD	C2-C3-C4-O4
4	L	605	MPD	C2-C3-C4-C5
4	L	604	MPD	C2-C3-C4-C5
4	S	1006	MPD	C2-C3-C4-C5

There are no ring outliers.

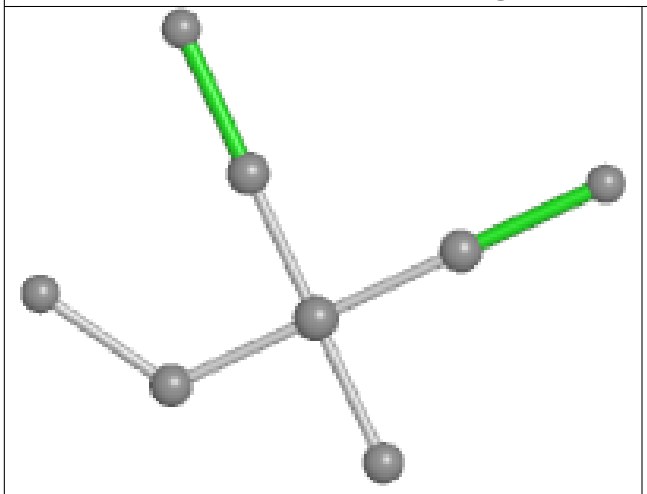
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	604	MPD	1	0
4	L	605	MPD	2	0
4	L	603	MPD	1	0

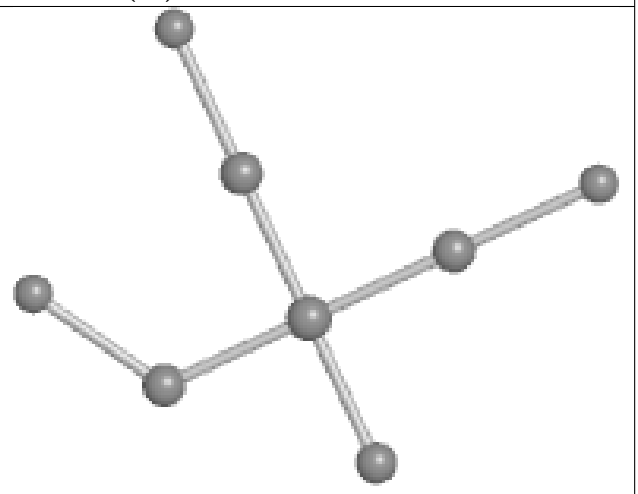
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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.

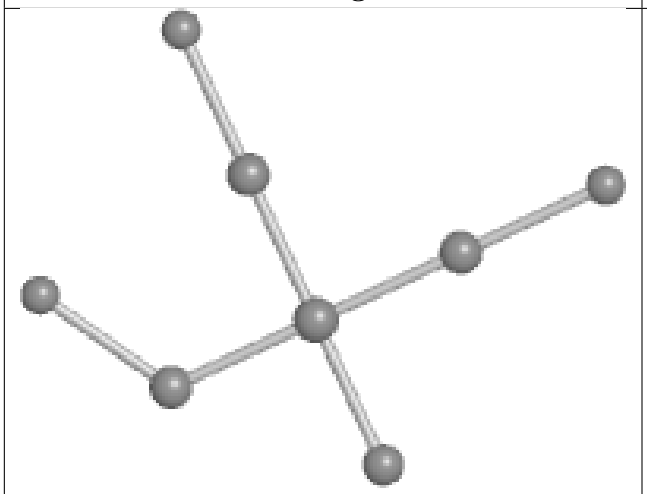
Ligand NFU L 601 (A)



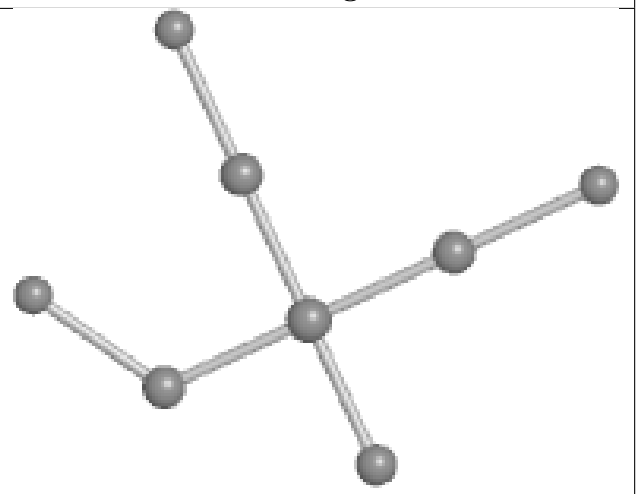
Bond lengths



Bond angles

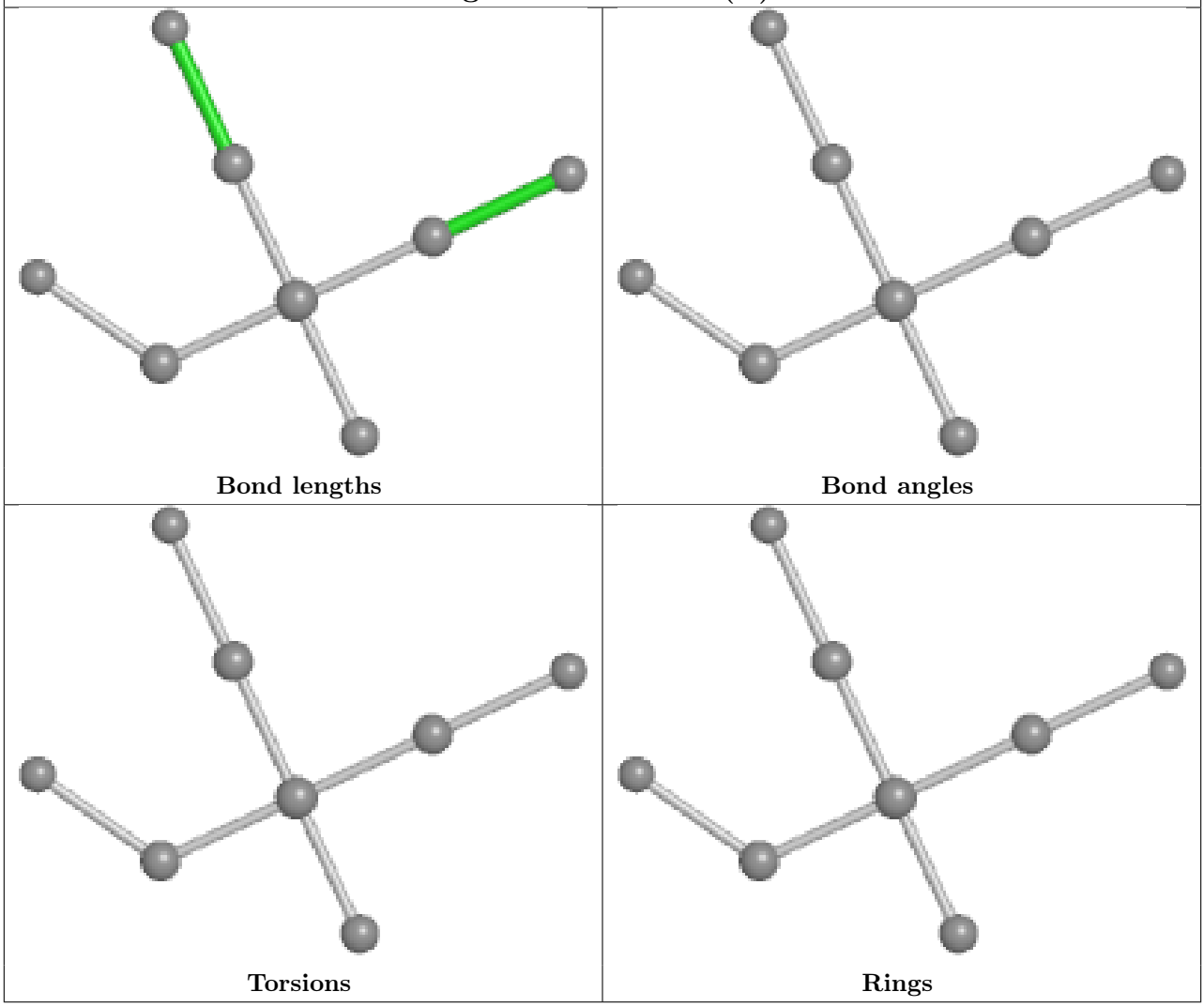


Torsions



Rings

Ligand NFU L 601 (B)



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

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.