wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2026 – 01:05 pm BST

PDB ID	:	9GNK / pdb_00009gnk
Title	:	Nitratidesulfovibrio vulgaris [FeFe]-hydrogenase variant with both subunits linked by a 13 amino acid linker peptide derived from CpI of Clostridium pasteurianum
Authors	:	Bikbaev, K.; Jaenecke, J.; Winkler, M.; Span, I.
Deposited on	:	2024-09-03
Resolution	:	1.05 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:

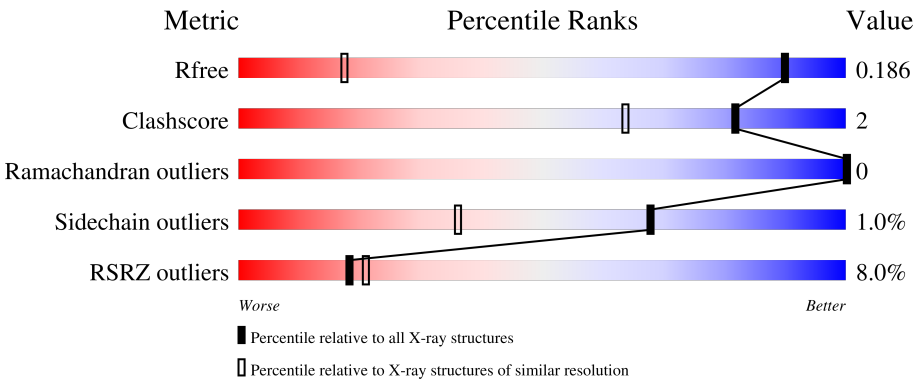
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal 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 (i)

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

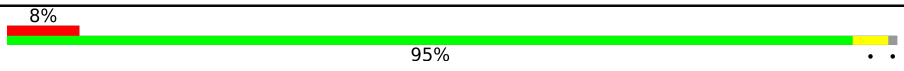
The reported resolution of this entry is 1.05 Å.

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	1182 (1.08-1.04)
Clashscore	190562	1204 (1.08-1.04)
Ramachandran outliers	187476	1175 (1.08-1.04)
Sidechain outliers	187428	1176 (1.08-1.04)
RSRZ outliers	180081	1181 (1.08-1.04)

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% 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	A	492	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 8357 atoms, of which 3886 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.

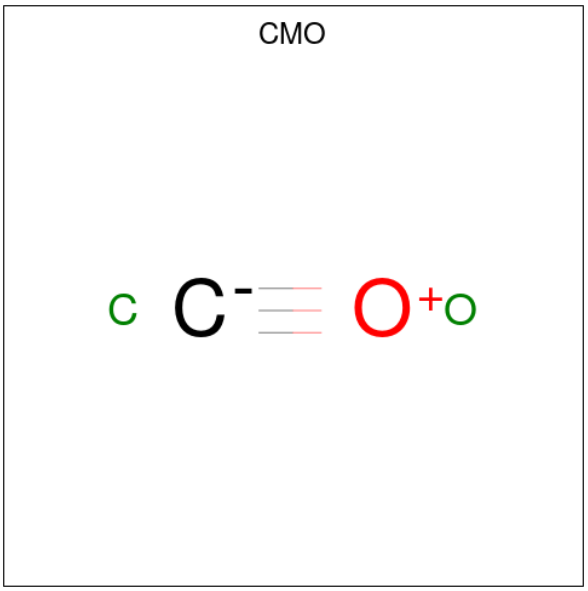
- Molecule 1 is a protein called Periplasmic [Fe] hydrogenase large subunit,Periplasmic [Fe] hydrogenase small subunit.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	485	7584	2424	3788	627	709	36	128	14	0

There are 23 discrepancies between the modelled and reference sequences:

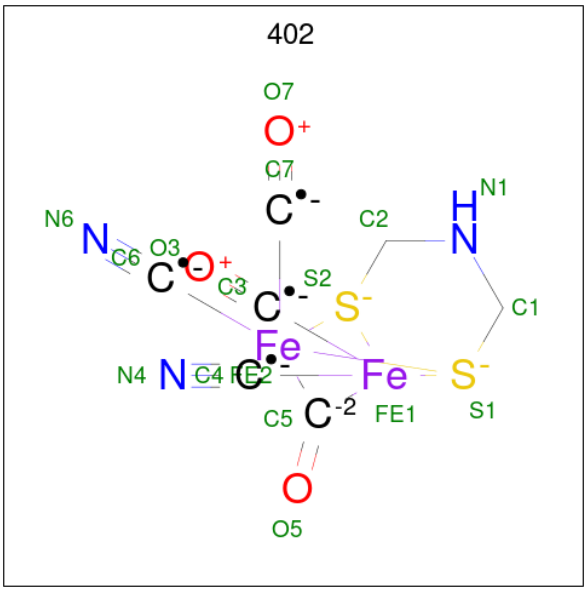
Chain	Residue	Modelled	Actual	Comment	Reference
A	391	VAL	-	linker	UNP P07598
A	392	ASN	-	linker	UNP P07598
A	393	PRO	-	linker	UNP P07598
A	394	LYS	-	linker	UNP P07598
A	395	ASP	-	linker	UNP P07598
A	396	LEU	-	linker	UNP P07598
A	397	GLU	-	linker	UNP P07598
A	398	LYS	-	linker	UNP P07598
A	399	VAL	-	linker	UNP P07598
A	400	ASP	-	linker	UNP P07598
A	401	ILE	-	linker	UNP P07598
A	402	LYS	-	linker	UNP P07598
A	403	LYS	-	linker	UNP P07598
A	483	SER	-	expression tag	UNP P07603
A	484	ALA	-	expression tag	UNP P07603
A	485	TRP	-	expression tag	UNP P07603
A	486	SER	-	expression tag	UNP P07603
A	487	HIS	-	expression tag	UNP P07603
A	488	PRO	-	expression tag	UNP P07603
A	489	GLN	-	expression tag	UNP P07603
A	490	PHE	-	expression tag	UNP P07603
A	491	GLU	-	expression tag	UNP P07603
A	492	LYS	-	expression tag	UNP P07603

- Molecule 2 is CARBON MONOXIDE (CCD ID: CMO) (formula: CO) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	2	1	1	0	0

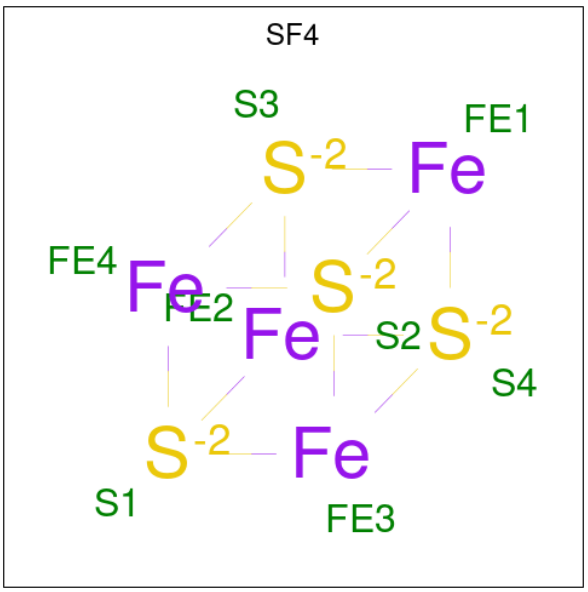
- Molecule 3 is dicarbonyl[bis(cyanide-kappaC)]-mu-(iminodimethanethiolatato-1kappaS:2kappaS)-mu-(oxomethylidene)diiron(2+) (CCD ID: 402) (formula: $C_7H_5Fe_2N_3O_3S_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
			Total	C	Fe	H	N	O	S		
3	A	1	24	7	2	7	3	3	2	1	0

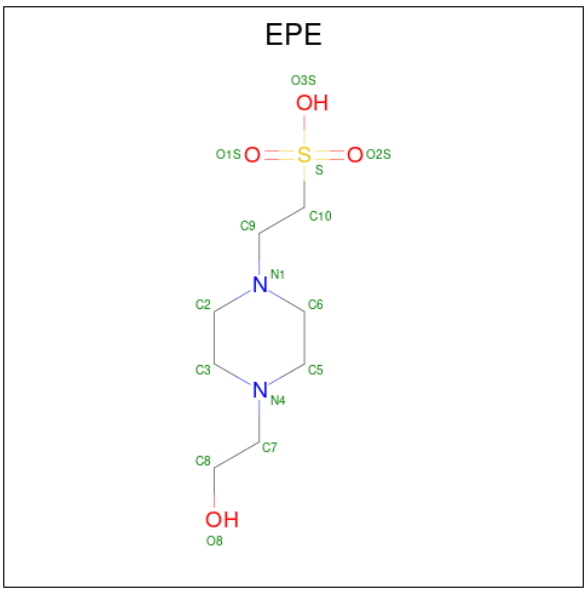
- Molecule 4 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe_4S_4) (labeled as "Lig-

and of Interest" by depositor).



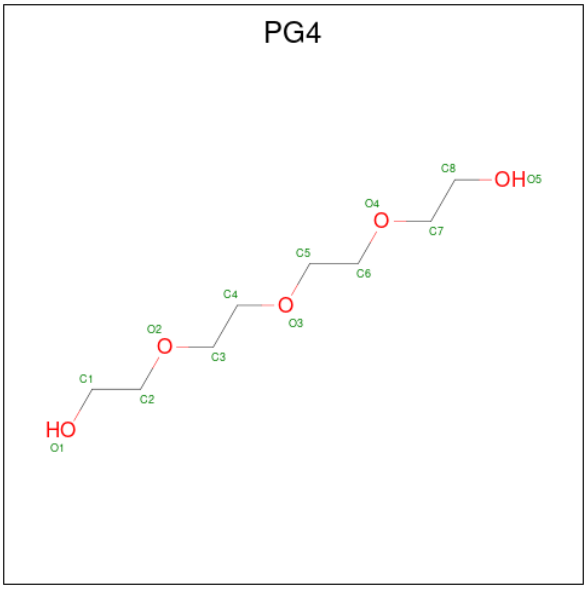
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
4	A	1	8	4	4	0	0
			Total	Fe	S		
4	A	1	8	4	4	0	0
			Total	Fe	S		
4	A	1	8	4	4	0	0

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (CCD ID: EPE) (formula: $C_8H_{18}N_2O_4S$).



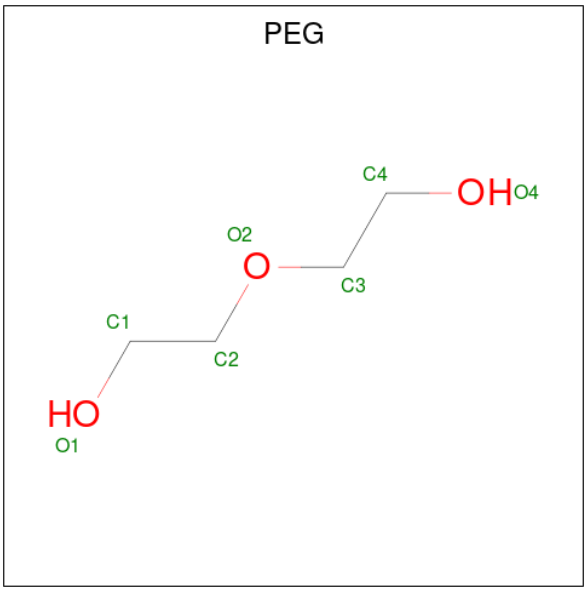
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	N	O	S		
5	A	1	33	8	18	2	4	1	2	0

- Molecule 6 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
6	A	1	31	8	18	5	2	0
			Total	C	H	O		
6	A	1	31	8	18	5	2	0
			Total	C	H	O		
6	A	1	17	5	10	2	4	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
7	A	1	17	4	10	3	2	0
			Total	C	H	O		
7	A	1	12	3	7	2	3	0
			Total	C	H	O		
7	A	1	17	4	10	3	2	0

- Molecule 8 is water.

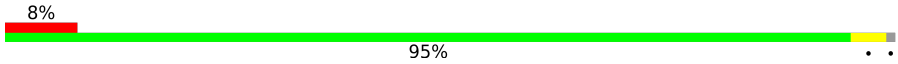
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	543	565	565	0	22

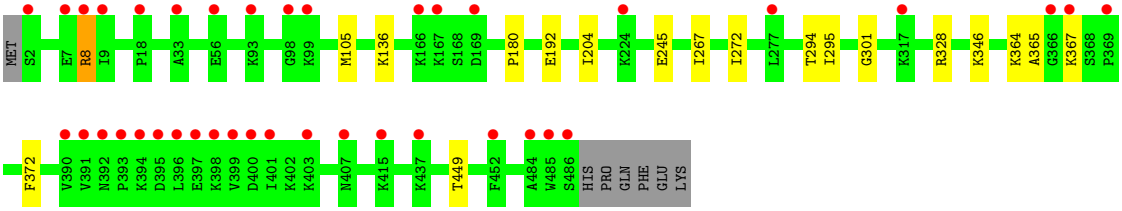
3 Residue-property plots (i)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

• Molecule 1: Periplasmic [Fe] hydrogenase large subunit,Periplasmic [Fe] hydrogenase small subunit

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.14Å 89.21Å 106.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.75 – 1.05 45.75 – 1.05	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.75-1.05) 99.8 (45.75-1.05)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 1.05Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.168 , 0.185 0.169 , 0.186	Depositor DCC
R_{free} test set	10877 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	11.3	Xtriage
Anisotropy	0.122	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 29.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8357	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CMO, PEG, 402, PG4, EPE, 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	A	0.46	0/3934	0.81	1/5319 (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	A	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	372	PHE	CA-CB-CG	5.73	119.53	113.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	8	ARG	Sidechain

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	A	3796	3788	3748	18	0
2	A	2	0	0	1	0
3	A	17	7	5	1	0
4	A	24	0	0	1	0
5	A	15	18	18	0	0
6	A	33	46	42	11	0
7	A	19	27	25	1	0
8	A	565	0	0	1	0
All	All	4471	3886	3838	19	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 19 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:A:328:ARG:HB2	7:A:510:PEG:H31	1.80	0.63
1:A:245:GLU:HB2	6:A:512:PG4:C4	2.32	0.60
1:A:449:THR:HA	6:A:512:PG4:H31	1.83	0.60
1:A:245:GLU:HG3	6:A:512:PG4:H42	1.83	0.60
1:A:367:LYS:HD3	8:A:602:HOH:O	2.03	0.58

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	A	497/492 (101%)	483 (97%)	14 (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	A	406/406 (100%)	402 (99%)	4 (1%)	68 36

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	ARG
1	A	136	LYS
1	A	294	THR
1	A	346	LYS

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

Mol	Chain	Res	Type
1	A	71	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

12 ligands 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	$\#\|Z\| > 2$
7	PEG	A	510	-	6,6,6	0.14	0	5,5,5	0.10	0
7	PEG	A	509	-	4,4,6	0.23	0	3,3,5	0.20	0
6	PG4	A	507	-	12,12,12	0.15	0	11,11,11	0.42	0
4	SF4	A	504	1	0,12,12	-	-	-		
6	PG4	A	511	-	12,12,12	0.15	0	11,11,11	0.43	0
6	PG4	A	512	-	6,6,12	0.20	0	5,5,11	0.49	0
2	CMO	A	501	-	0,1,1	-	-	-		
3	402	A	502	1	13,19,19	4.37	7 (53%)	2,36,36	0.56	0
4	SF4	A	505	1	0,12,12	-	-	-		
7	PEG	A	508	-	6,6,6	0.12	0	5,5,5	0.10	0
5	EPE	A	506	-	15,15,15	0.64	1 (6%)	18,20,20	0.76	1 (5%)
4	SF4	A	503	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
7	PEG	A	510	-	-	3/4/4/4	-
7	PEG	A	509	-	-	1/2/2/4	-
6	PG4	A	507	-	-	5/10/10/10	-
4	SF4	A	504	1	-	-	0/6/5/5
6	PG4	A	511	-	-	6/10/10/10	-
6	PG4	A	512	-	-	3/4/4/10	-
4	SF4	A	505	1	-	-	0/6/5/5
3	402	A	502	1	-	-	0/5/3/3
7	PEG	A	508	-	-	0/4/4/4	-
5	EPE	A	506	-	-	3/9/19/19	0/1/1/1
4	SF4	A	503	1	-	-	0/6/5/5

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	402	C2-S2	-10.08	1.67	1.85
3	A	502	402	C1-S1	-9.39	1.68	1.85
3	A	502	402	S1-FE2	-4.38	2.20	2.26
3	A	502	402	O5-C5	3.67	1.24	1.17
3	A	502	402	S1-FE1	-3.16	2.21	2.26

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	506	EPE	O3S-S-C10	-2.06	102.44	105.77

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	511	PG4	O2-C3-C4-O3
6	A	507	PG4	O2-C3-C4-O3
7	A	510	PEG	O2-C3-C4-O4
6	A	507	PG4	O4-C7-C8-O5
5	A	506	EPE	C8-C7-N4-C5

There are no ring outliers.

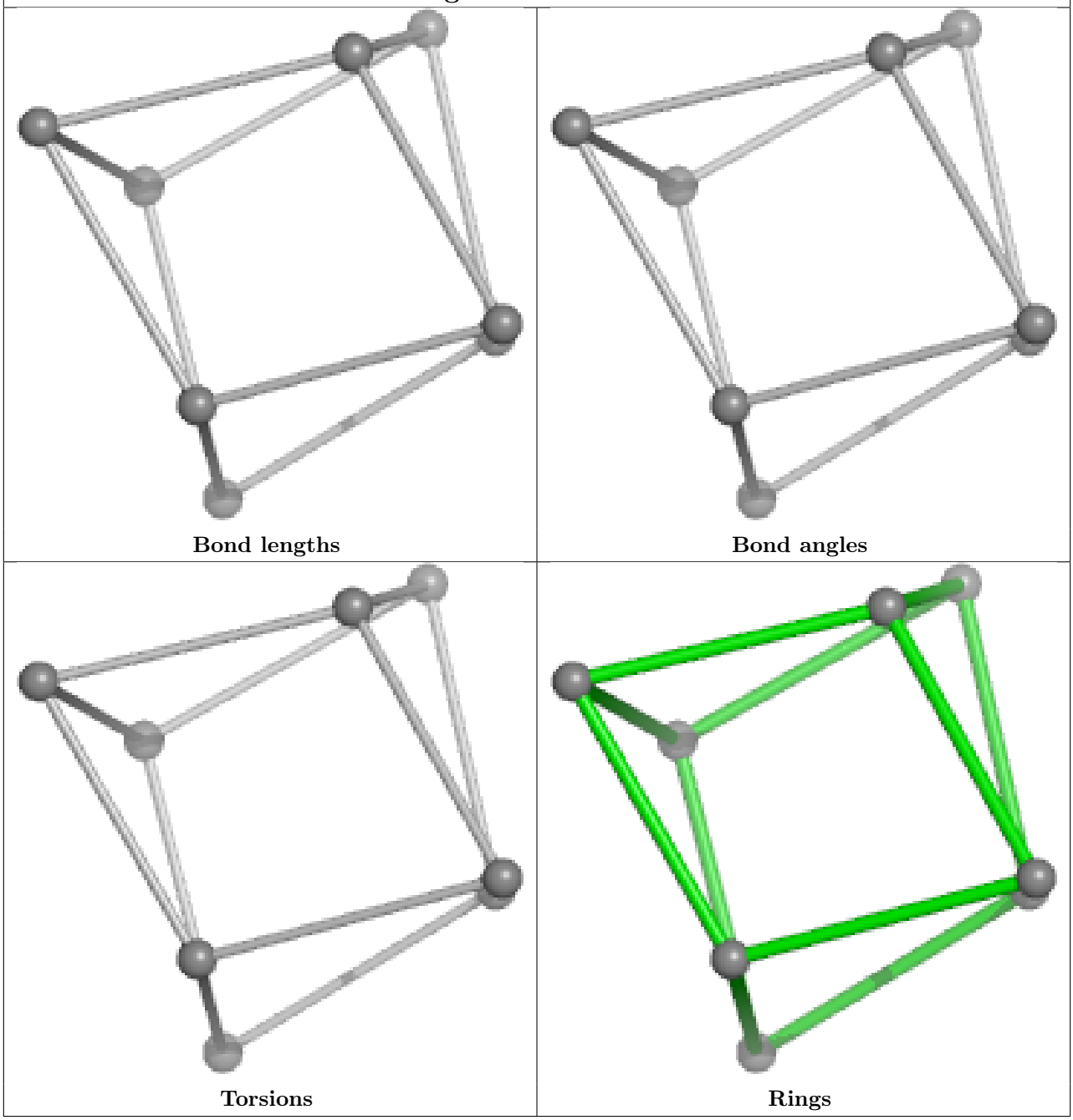
6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	510	PEG	1	0
6	A	507	PG4	5	0
6	A	512	PG4	6	0
2	A	501	CMO	1	0
3	A	502	402	1	0
4	A	503	SF4	1	0

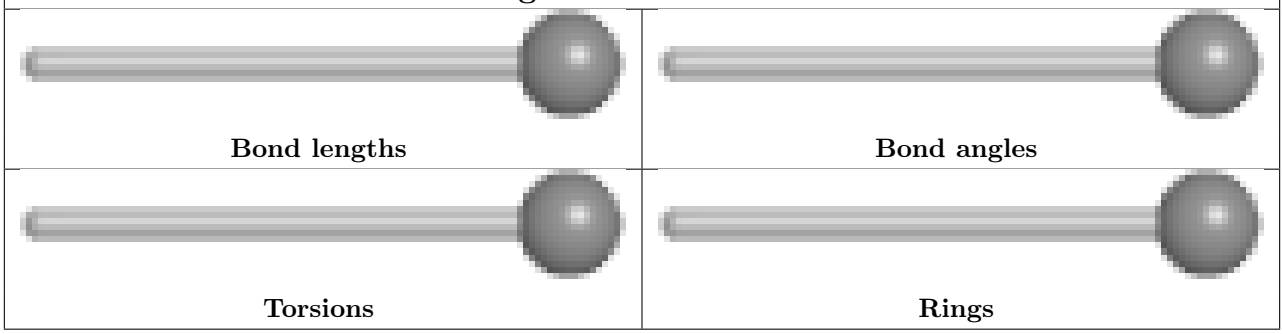
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

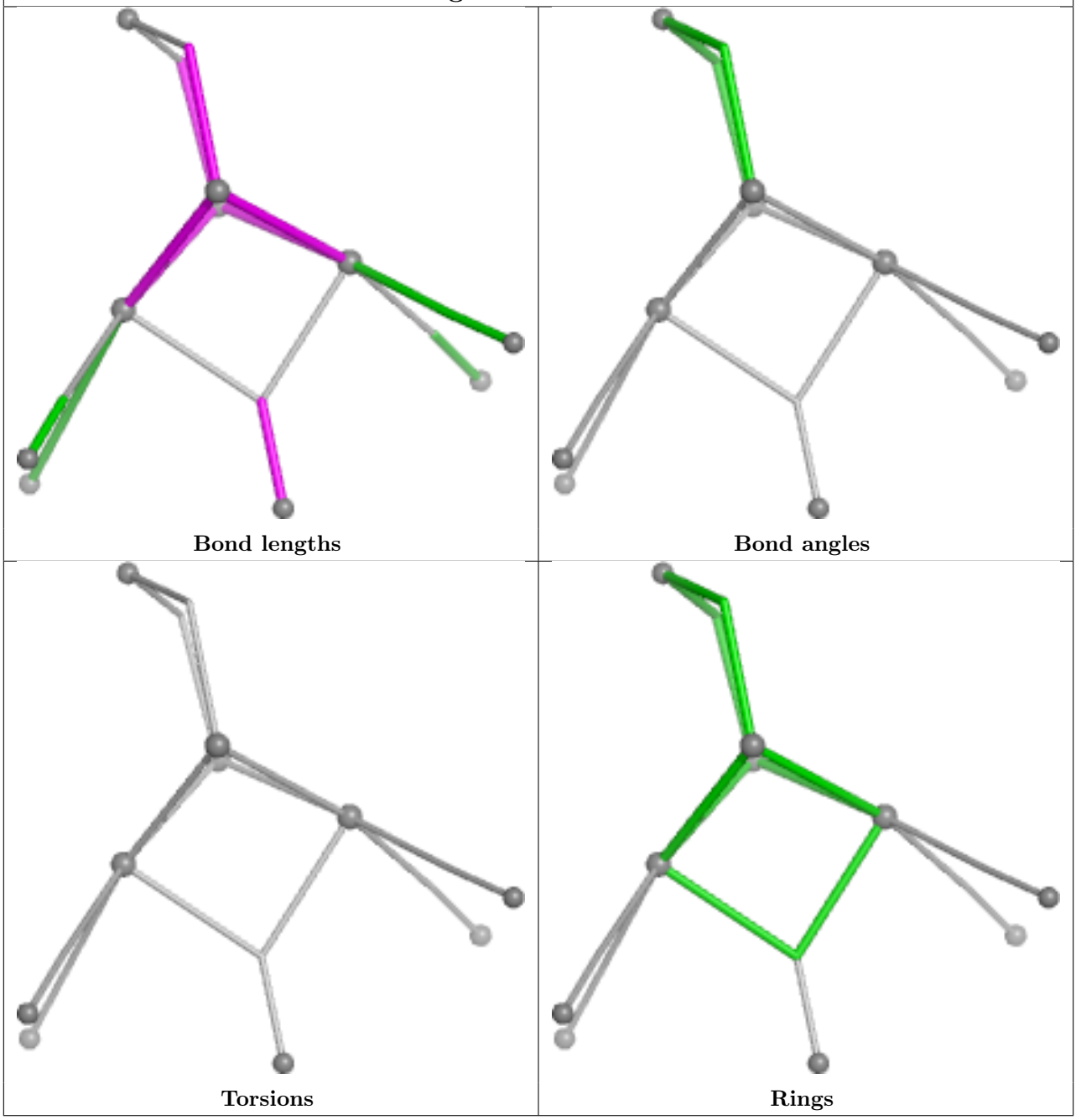
Ligand SF4 A 504



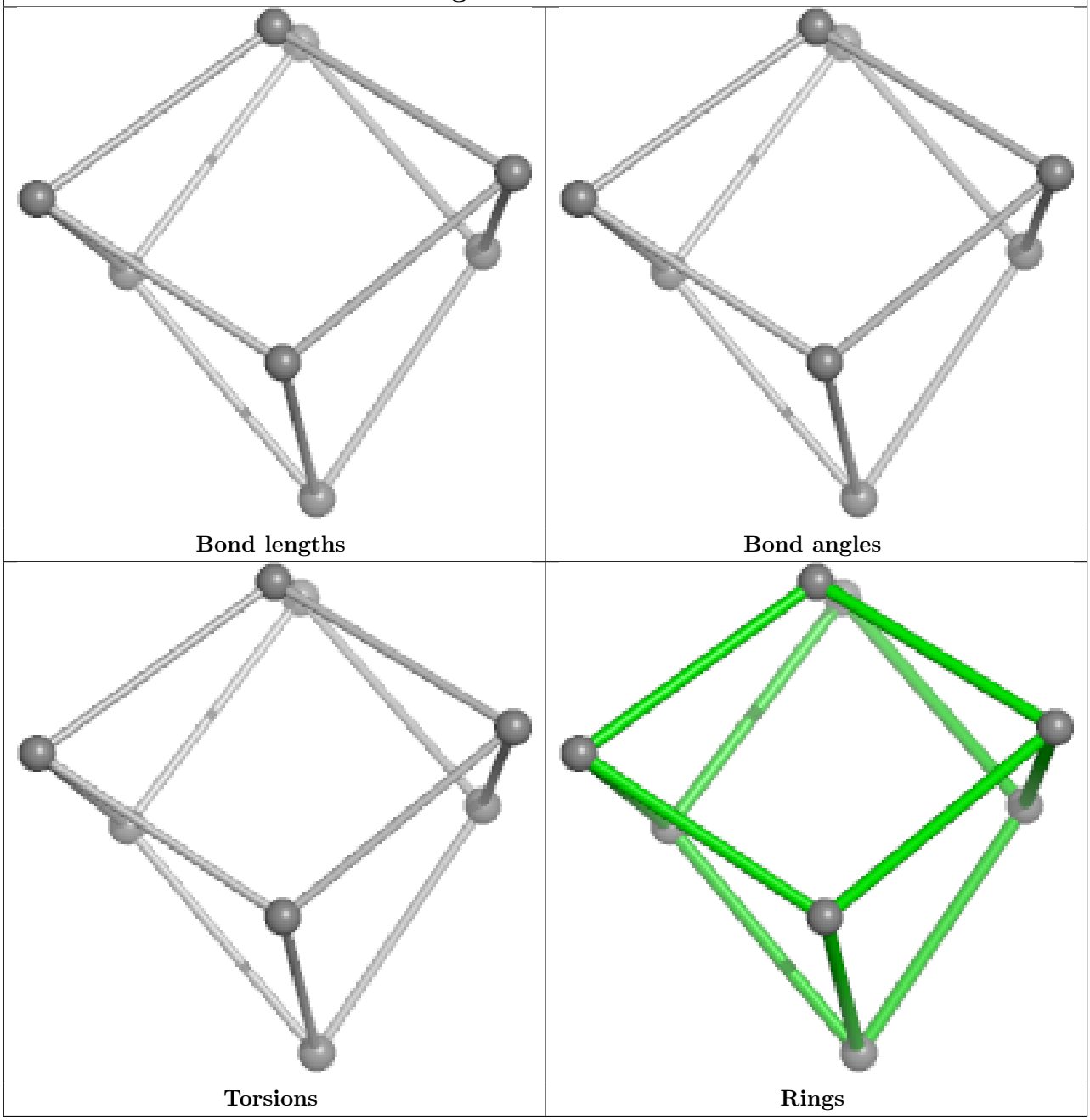
Ligand CMO A 501



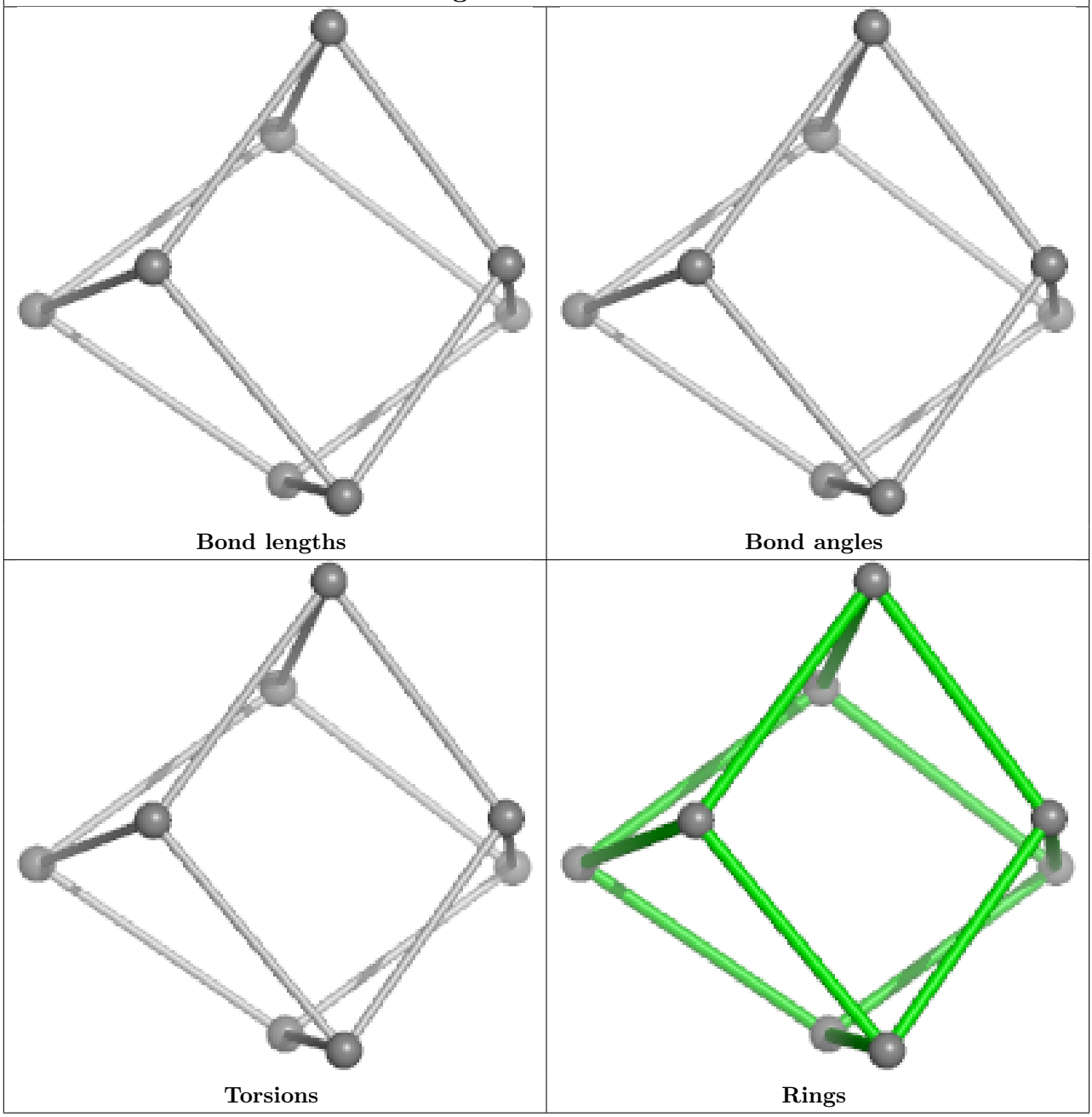
Ligand 402 A 502



Ligand SF4 A 505



Ligand SF4 A 503



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	A	485/492 (98%)	0.60	39 (8%) 18 21	5, 12, 18, 43	14 (2%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	399	VAL	8.4
1	A	2	SER	7.5
1	A	392	ASN	7.2
1	A	391	VAL	7.1
1	A	393	PRO	5.6

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(Å^2)	Q<0.9
6	PG4	A	511	13/13	0.59	0.17	24,28,32,44	2
7	PEG	A	510	7/7	0.63	0.18	28,28,29,29	2

Continued on next page...

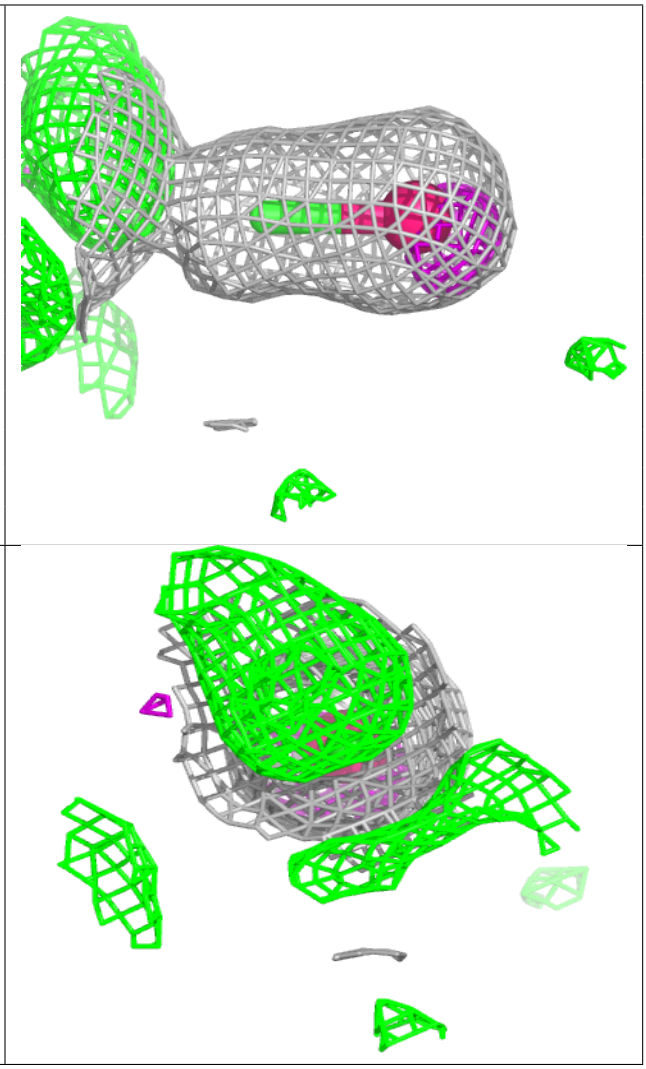
Continued from previous page...

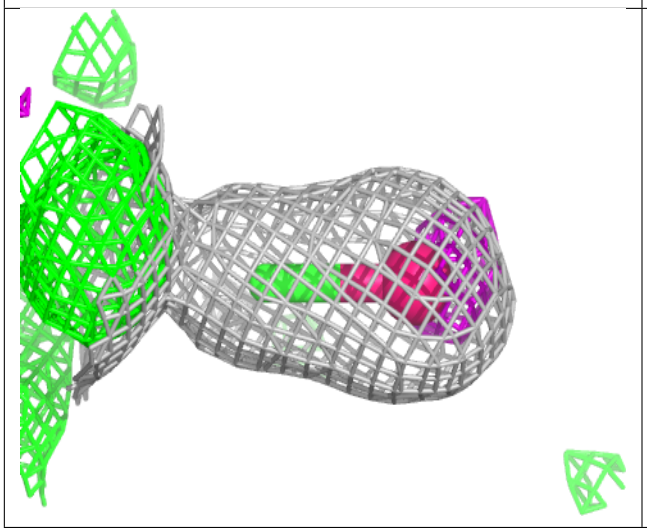
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	PG4	A	507	13/13	0.72	0.15	24,27,29,30	2
6	PG4	A	512	7/13	0.75	0.18	19,19,20,20	17
7	PEG	A	509	5/7	0.80	0.15	15,17,18,18	12
5	EPE	A	506	15/15	0.80	0.18	27,29,29,29	33
7	PEG	A	508	7/7	0.84	0.12	24,25,26,26	2
2	CMO	A	501	2/2	0.97	0.12	9,9,9,10	0
4	SF4	A	505	8/8	0.98	0.04	12,12,13,13	0
3	402	A	502	17/17	0.99	0.07	8,8,8,9	1
4	SF4	A	504	8/8	0.99	0.03	9,9,9,9	0
4	SF4	A	503	8/8	1.00	0.02	7,7,8,8	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 CMO A 501:

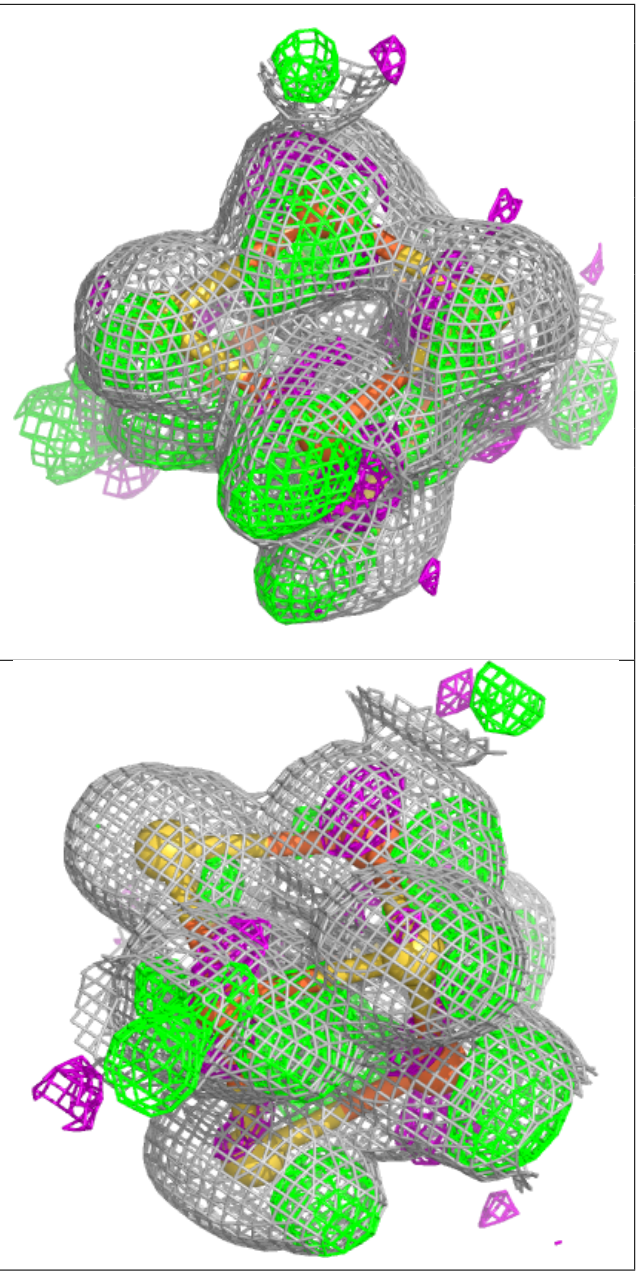
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)





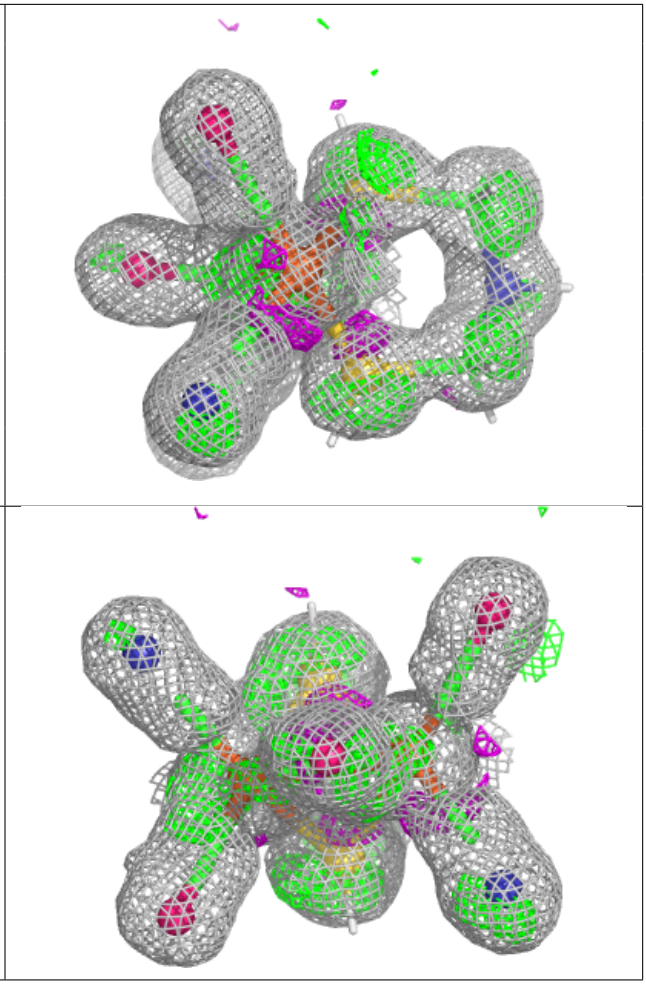
Electron density around SF4 A 505:

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



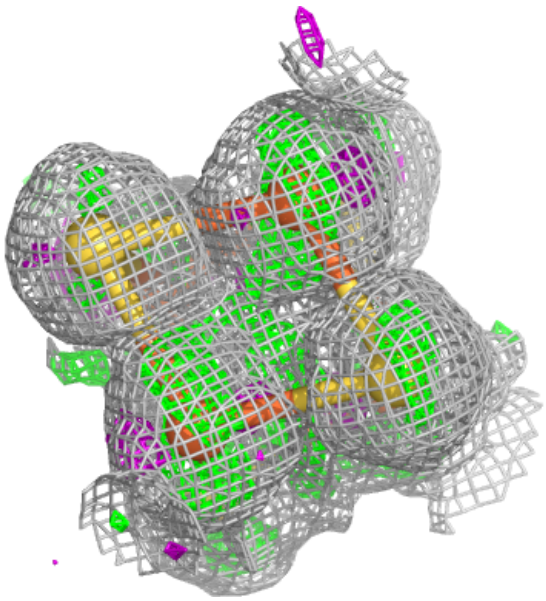
Electron density around 402 A 502:

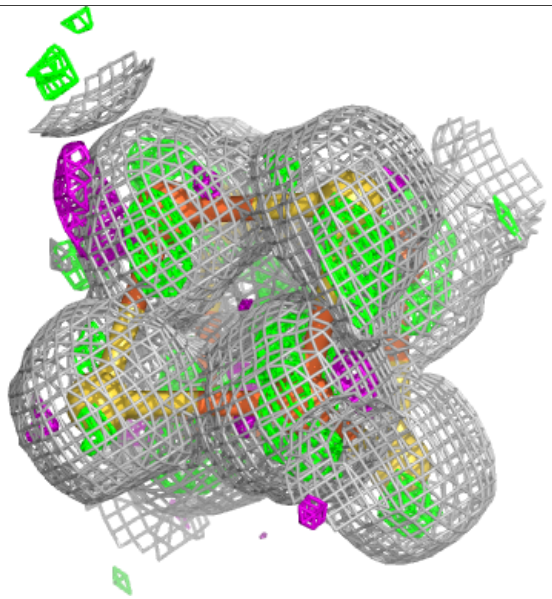
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

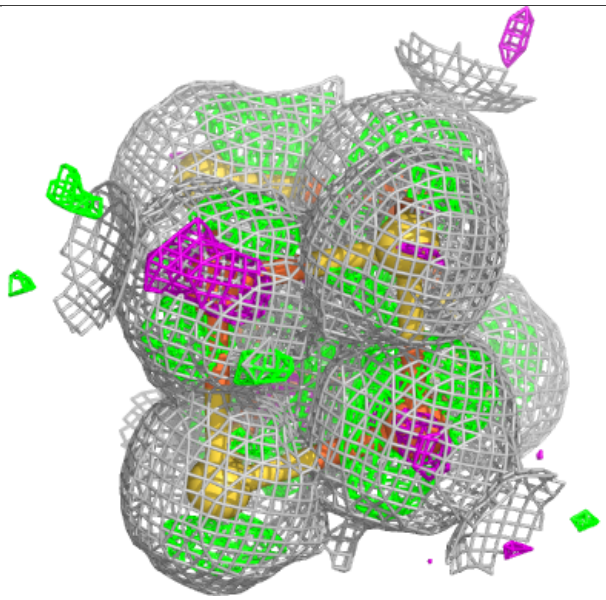


Electron density around SF4 A 504:

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

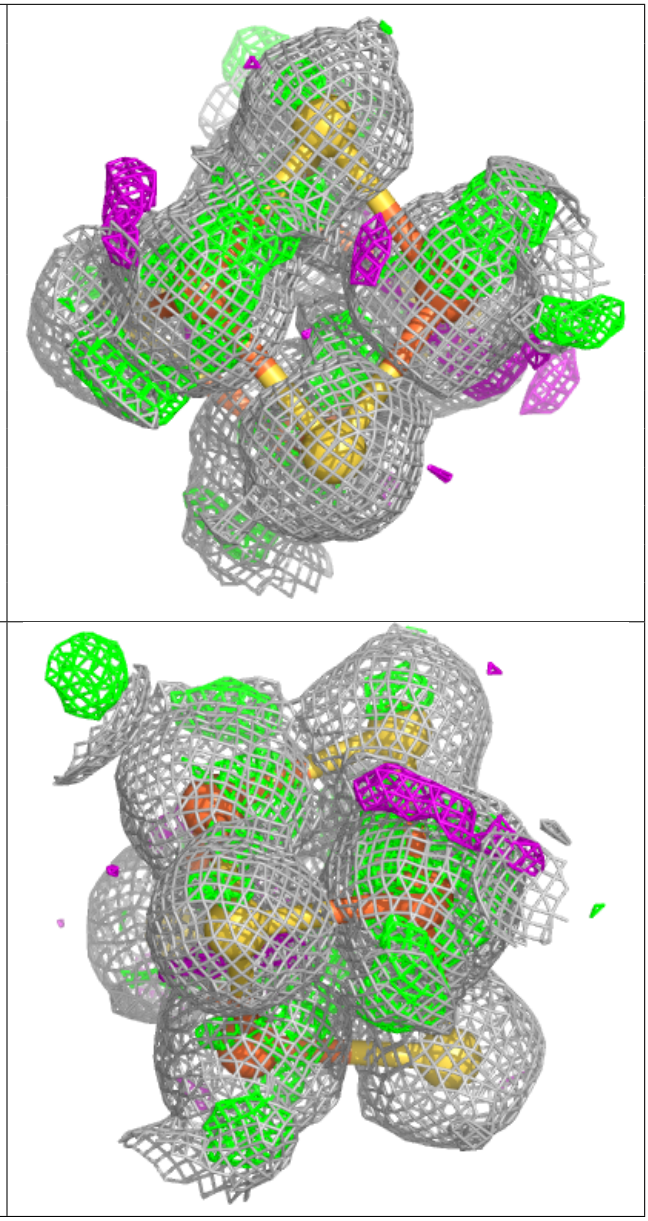






Electron density around SF4 A 503:

$2mF_o$-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.