



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 6, 2021 – 02:24 pm BST

PDB ID : 7B2C  
Title : Crystal structure of the ethyl-coenzyme M reductase from *Candidatus*  
Ethanoperedens thermophilum gassed with xenon  
Authors : Wagner, T.; Lemaire, O.N.; Engilberge, S.  
Deposited on : 2020-11-26  
Resolution : 1.80 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](mailto: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) : 1.13  
EDS : 2.22  
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) : 2.22

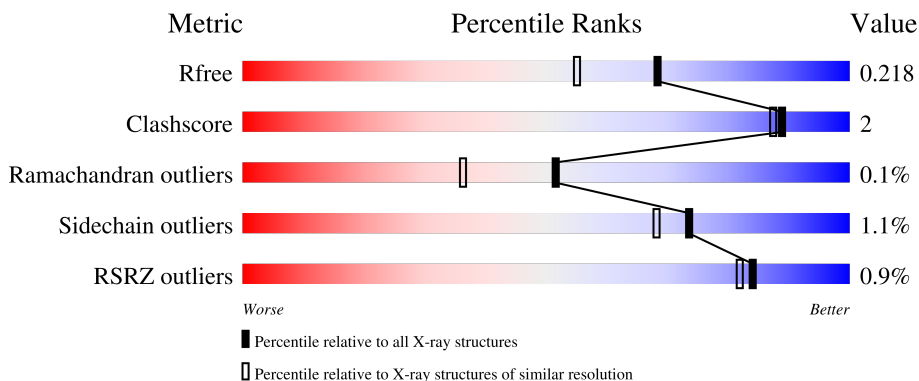
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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  $\geq 3$ , 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	A	595	 94% 5% ..
1	D	595	 95% ..
2	B	467	 96% .
2	E	467	 95% .
3	C	266	 94% 5%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	266	 93% 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	XE	B	511[B]	-	-	X	-
10	XE	F	1106[A]	-	-	X	-
4	USN	A	1201	X	-	-	-
4	USN	F	1101	X	-	-	-

## 2 Entry composition [i](#)

There are 15 unique types of molecules in this entry. The entry contains 23280 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 Ethyl-Coenzyme M reductase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	591	Total	C	N	O	S	0	5	0
			4669	2933	816	891	29			
1	D	593	Total	C	N	O	S	0	4	0
			4675	2942	816	888	29			

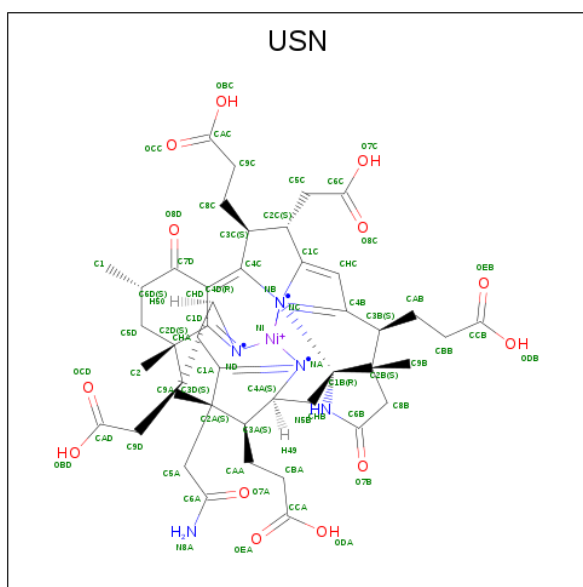
- Molecule 2 is a protein called Ethyl-Coenzyme M reductase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	466	Total	C	N	O	S	0	2	0
			3517	2222	609	670	16			
2	E	466	Total	C	N	O	S	0	1	0
			3503	2215	604	668	16			

- Molecule 3 is a protein called Ethyl-Coenzyme M reductase gamma subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	265	Total	C	N	O	S	0	2	0
			2134	1322	387	408	17			
3	F	265	Total	C	N	O	S	0	3	0
			2140	1325	388	410	17			

- Molecule 4 is Dimethylated-F430 cofactor (three-letter code: USN) (formula:  $C_{44}H_{55}N_6NiO_{13}$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	Ni			O
4	A	1	64	44	6	1	13	0	0
4	F	1	64	44	6	1	13	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	6	3	3	0	0
5	A	1	6	3	3	0	0

*Continued on next page...*

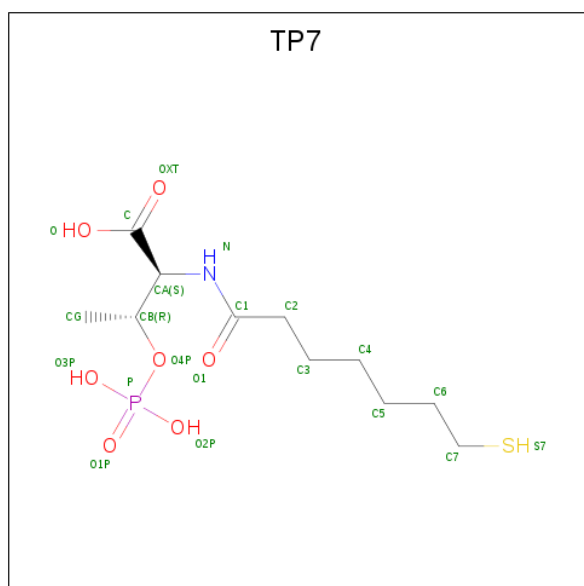
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

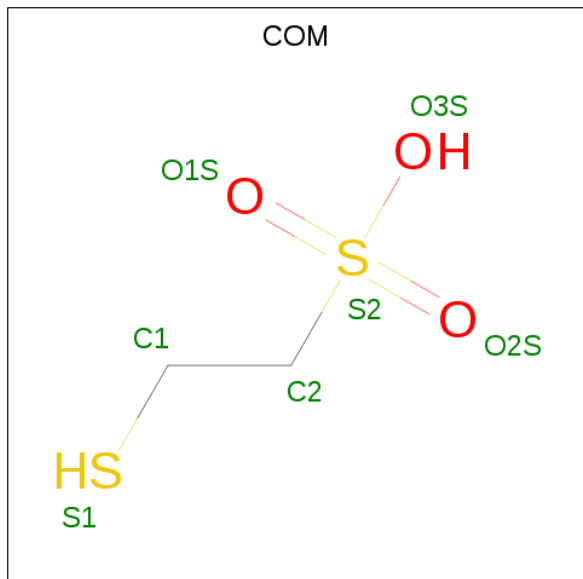
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		
6	C	1	Total	Na	0	0
			1	1		
6	D	3	Total	Na	0	0
			3	3		

- Molecule 7 is Coenzyme B (three-letter code: TP7) (formula: C<sub>11</sub>H<sub>22</sub>NO<sub>7</sub>PS) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
7	A	1	Total	C	N	O	P	S	0	0
			21	11	1	7	1	1		
7	D	1	Total	C	N	O	P	S	0	0
			21	11	1	7	1	1		

- Molecule 8 is 1-THIOETHANESULFONIC ACID (three-letter code: COM) (formula:  $C_2H_6O_3S_2$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	O	S	0	0
			7	2	3	2		
8	F	1	Total	C	O	S	0	0
			7	2	3	2		

- Molecule 9 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	2	Total	K	0	0
			2	2		
9	B	5	Total	K	0	0
			5	5		
9	C	1	Total	K	0	0
			1	1		
9	D	1	Total	K	0	0
			1	1		
9	E	7	Total	K	0	0
			7	7		
9	F	1	Total	K	0	0
			1	1		

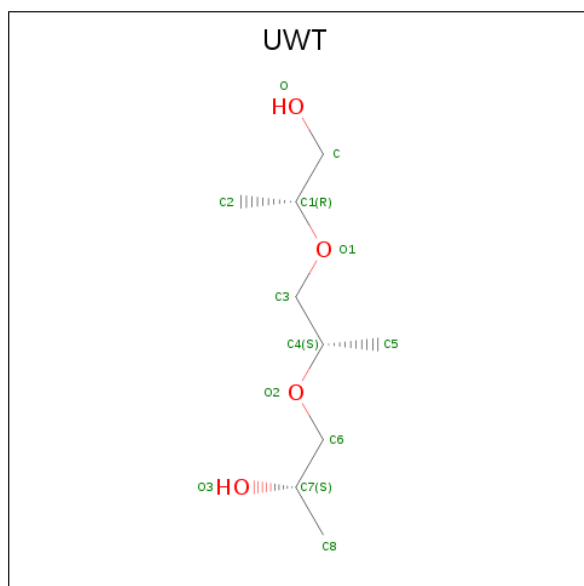
- Molecule 10 is XENON (three-letter code: XE) (formula: Xe) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	3	Total Xe 3 3	0	0
10	B	3	Total Xe 3 3	0	1
10	C	2	Total Xe 2 2	0	1
10	D	3	Total Xe 3 3	0	0
10	E	3	Total Xe 3 3	0	1
10	F	2	Total Xe 2 2	0	1

- Molecule 11 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	B	2	Total Cl 2 2	0	0
11	E	3	Total Cl 3 3	0	0

- Molecule 12 is (2R)-2-[(2S)-2-[(2S)-2-oxidanylpropoxy]propoxy]propan-1-ol (three-letter code: UWT) (formula: C<sub>9</sub>H<sub>20</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	C	1	Total C O 11 8 3	0	0

*Continued on next page...*



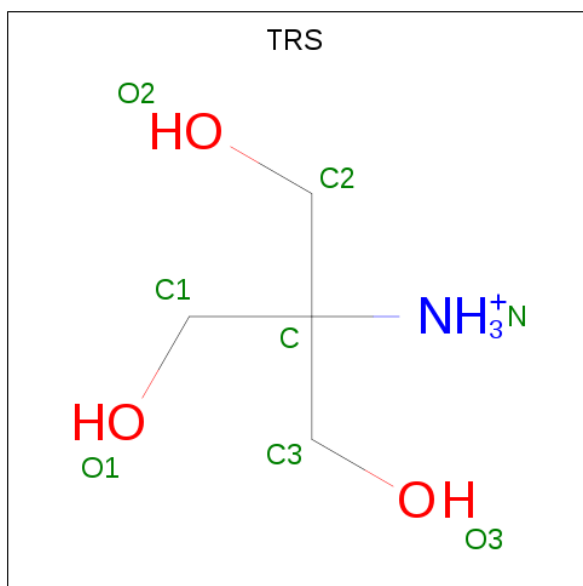
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
12	F	1	Total	C	O	0	0
			13	9	4		

- Molecule 13 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	D	1	Total	Mg	0	0
			1	1		

- Molecule 14 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
14	E	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 15 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	498	Total	O	0	4
			498	498		
15	B	406	Total	O	0	7
			406	406		
15	C	248	Total	O	0	2
			248	248		

*Continued on next page...*

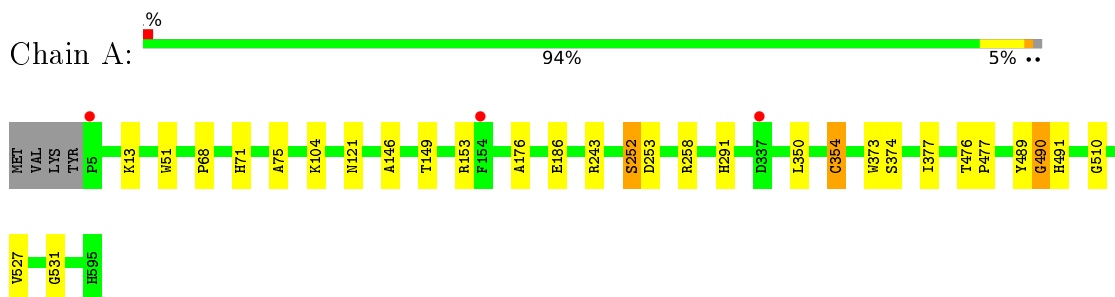
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
15	D	479	Total 479	O 479	0	5
15	E	436	Total 436	O 436	0	3
15	F	279	Total 279	O 279	0	2

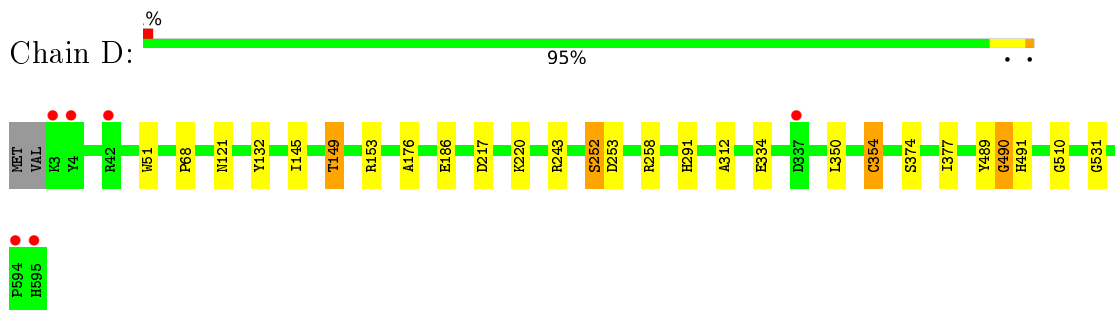
### 3 Residue-property plots [i](#)

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.

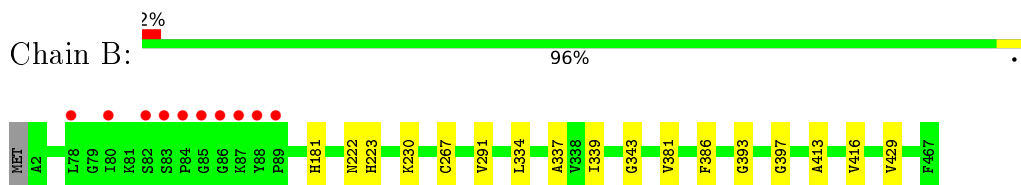
- Molecule 1: Ethyl-Coenzyme M reductase alpha subunit



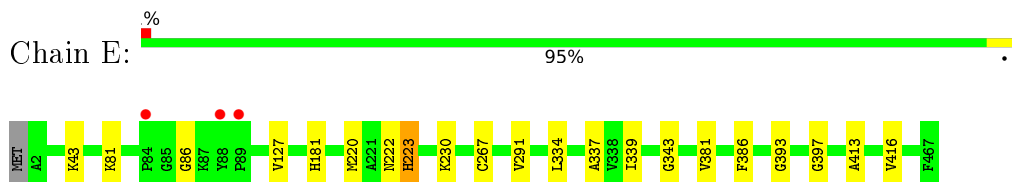
- Molecule 1: Ethyl-Coenzyme M reductase alpha subunit



- Molecule 2: Ethyl-Coenzyme M reductase beta subunit



- Molecule 2: Ethyl-Coenzyme M reductase beta subunit



- Molecule 3: Ethyl-Coenzyme M reductase gamma subunit

Chain C:  94% 5%



- Molecule 3: Ethyl-Coenzyme M reductase gamma subunit

Chain F:  93% 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.79Å 147.19Å 113.38Å 90.00° 107.20° 90.00°	Depositor
Resolution (Å)	35.90 – 1.80 67.53 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (35.90-1.80) 96.0 (67.53-1.80)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 1.80Å)	Xtriage
Refinement program	BUSTER 2.10.3 (19-MAR-2020)	Depositor
R, $R_{free}$	0.174 , 0.201 0.189 , 0.218	Depositor DCC
$R_{free}$ test set	11695 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.1	Xtriage
Anisotropy	0.007	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 52.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	23280	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: MHS, MGN, K, GL3, GOL, HIC, I2M, XE, COM, TP7, TRS, MG, USN, NA, AGM, SMC, UWT, CL

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	0.41	0/4713	0.61	0/6373
1	D	0.40	0/4723	0.61	0/6388
2	B	0.40	0/3589	0.59	0/4872
2	E	0.40	0/3578	0.59	0/4857
3	C	0.38	0/2178	0.60	0/2944
3	F	0.38	0/2184	0.60	0/2952
All	All	0.40	0/20965	0.60	0/28386

There are no bond length outliers.

There are no bond angle outliers.

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	A	4669	0	4443	16	0
1	D	4675	0	4462	20	0
2	B	3517	0	3485	14	0
2	E	3503	0	3473	18	0
3	C	2134	0	2088	5	0
3	F	2140	0	2092	9	0
4	A	64	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	64	0	0	1	0
5	A	12	0	16	1	0
5	B	6	0	8	1	0
5	D	12	0	16	0	0
5	F	6	0	8	0	0
6	A	1	0	0	0	0
6	C	1	0	0	0	0
6	D	3	0	0	0	0
7	A	21	0	19	0	0
7	D	21	0	19	0	0
8	A	7	0	5	0	0
8	F	7	0	5	1	0
9	A	2	0	0	0	0
9	B	5	0	0	0	0
9	C	1	0	0	0	0
9	D	1	0	0	0	0
9	E	7	0	0	0	0
9	F	1	0	0	0	0
10	A	3	0	0	0	0
10	B	3	0	0	3	0
10	C	2	0	0	0	0
10	D	3	0	0	1	0
10	E	3	0	0	1	0
10	F	2	0	0	2	0
11	B	2	0	0	0	0
11	E	3	0	0	0	0
12	C	11	0	0	0	0
12	F	13	0	0	0	0
13	D	1	0	0	0	0
14	E	8	0	12	1	0
15	A	498	0	0	1	0
15	B	406	0	0	2	0
15	C	248	0	0	1	0
15	D	479	0	0	2	0
15	E	436	0	0	2	0
15	F	279	0	0	5	0
All	All	23280	0	20151	76	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 76 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:B:267:CYS:SG	10:B:510:XE:XE	3.15	0.82
2:B:429:VAL:HG21	1:D:149[B]:THR:HG22	1.66	0.78
2:E:267:CYS:SG	10:E:513:XE:XE	3.20	0.77
10:B:511[B]:XE:XE	1:D:153:ARG:HG3	2.72	0.68
2:E:413:ALA:O	2:E:416:VAL:HG22	1.98	0.64

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	587/595 (99%)	563 (96%)	22 (4%)	2 (0%)	41	27
1	D	588/595 (99%)	568 (97%)	18 (3%)	2 (0%)	41	27
2	B	466/467 (100%)	455 (98%)	11 (2%)	0	100	100
2	E	465/467 (100%)	457 (98%)	8 (2%)	0	100	100
3	C	264/266 (99%)	257 (97%)	7 (3%)	0	100	100
3	F	265/266 (100%)	258 (97%)	7 (3%)	0	100	100
All	All	2635/2656 (99%)	2558 (97%)	73 (3%)	4 (0%)	51	33

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252[A]	SER
1	A	252[B]	SER
1	D	252[A]	SER
1	D	252[B]	SER



### 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	471/470 (100%)	468 (99%)	3 (1%)	86	84
1	D	472/470 (100%)	466 (99%)	6 (1%)	69	62
2	B	365/364 (100%)	364 (100%)	1 (0%)	92	91
2	E	364/364 (100%)	363 (100%)	1 (0%)	92	91
3	C	236/236 (100%)	230 (98%)	6 (2%)	47	34
3	F	237/236 (100%)	229 (97%)	8 (3%)	37	22
All	All	2145/2140 (100%)	2120 (99%)	25 (1%)	73	65

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	334	GLU
3	F	29	GLU
3	F	257	ASN
2	E	223	HIS
3	F	56	ASN

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](#)

14 non-standard protein/DNA/RNA residues 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
1	MHS	A	291	1	7,11,12	0.71	0	6,14,16	2.16	2 (33%)
1	MHS	D	291	1	7,11,12	0.60	0	6,14,16	2.15	2 (33%)
1	I2M	D	377	1	6,8,9	1.11	1 (16%)	6,11,13	0.77	0
1	HIC	D	491	1	8,11,12	1.07	1 (12%)	6,14,16	1.23	1 (16%)
1	AGM	D	305	1	10,11,12	0.48	0	6,13,15	0.67	0
1	AGM	A	305	1	10,11,12	0.45	0	6,13,15	0.81	0
1	MGN	A	445	1	6,9,10	0.72	0	5,12,14	0.15	0
1	SMC	D	354	1	5,6,7	1.11	1 (20%)	2,6,8	0.42	0
1	MGN	D	445	1	6,9,10	0.70	0	5,12,14	0.27	0
1	HIC	A	491	1	8,11,12	1.13	2 (25%)	6,14,16	1.38	1 (16%)
1	I2M	A	377	1	6,8,9	1.12	1 (16%)	6,11,13	0.61	0
1	SMC	A	354	1	5,6,7	1.13	1 (20%)	2,6,8	0.76	0
1	GL3	D	490	1	2,3,4	5.45	1 (50%)	1,2,4	0.30	0
1	GL3	A	490	1	2,3,4	4.73	1 (50%)	1,2,4	0.40	0

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	MHS	A	291	1	-	0/5/6/8	0/1/1/1
1	MHS	D	291	1	-	0/5/6/8	0/1/1/1
1	I2M	D	377	1	-	0/8/11/13	-
1	HIC	D	491	1	-	1/5/6/8	0/1/1/1
1	AGM	D	305	1	-	1/10/11/13	-
1	AGM	A	305	1	-	1/10/11/13	-
1	MGN	A	445	1	-	0/7/9/12	-
1	SMC	D	354	1	-	0/3/5/7	-
1	MGN	D	445	1	-	0/7/9/12	-
1	HIC	A	491	1	-	0/5/6/8	0/1/1/1
1	I2M	A	377	1	-	0/8/11/13	-
1	SMC	A	354	1	-	0/3/5/7	-
1	GL3	D	490	1	-	0/1/1/2	-
1	GL3	A	490	1	-	0/1/1/2	-

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	490	GL3	C-S	-7.70	1.54	1.80
1	A	490	GL3	C-S	-6.68	1.57	1.80
1	A	377	I2M	CB-CA	-2.49	1.53	1.57
1	A	354	SMC	CB-SG	-2.38	1.77	1.80
1	D	377	I2M	CB-CA	-2.38	1.53	1.57

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	291	MHS	CB-CA-C	-4.19	103.60	111.47
1	A	291	MHS	CB-CA-C	-4.12	103.75	111.47
1	A	491	HIC	CB-CA-C	-3.34	105.20	111.47
1	D	491	HIC	CB-CA-C	-3.00	105.85	111.47
1	A	291	MHS	NE2-CE1-ND1	-2.19	109.01	112.26

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	491	HIC	CA-CB-CG-ND1
1	A	305	AGM	NE1-CD-CG-CB
1	D	305	AGM	CE2-CD-NE1-CZ

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	354	SMC	1	0
1	A	354	SMC	1	0
1	D	490	GL3	1	0
1	A	490	GL3	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 59 ligands modelled in this entry, 44 are monoatomic - leaving 15 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
4	USN	F	1101	8	47,73,73	3.38	19 (40%)	51,123,123	2.34	16 (31%)
8	COM	F	1102	4	6,6,6	1.17	0	7,8,8	2.47	3 (42%)
14	TRS	E	501	-	7,7,7	0.15	0	9,9,9	0.34	0
4	USN	A	1201	8	47,73,73	2.87	16 (34%)	51,123,123	2.54	20 (39%)
12	UWT	F	1103	-	9,12,12	0.62	0	10,14,14	0.88	0
7	TP7	A	1204	-	16,20,20	0.59	0	18,26,26	0.86	0
5	GOL	A	1211	-	5,5,5	0.09	0	5,5,5	0.31	0
5	GOL	B	501	-	5,5,5	0.22	0	5,5,5	0.33	0
5	GOL	D	703	-	5,5,5	0.29	0	5,5,5	0.32	0
5	GOL	F	1107	-	5,5,5	0.09	0	5,5,5	0.32	0
8	COM	A	1205	4	6,6,6	1.47	2 (33%)	7,8,8	2.43	4 (57%)
5	GOL	A	1202	-	5,5,5	0.30	0	5,5,5	0.21	0
5	GOL	D	702	-	5,5,5	0.24	0	5,5,5	0.31	0
7	TP7	D	701	-	16,20,20	0.78	1 (6%)	18,26,26	0.96	0
12	UWT	C	301	-	8,10,12	0.59	0	8,11,14	0.56	0

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	USN	F	1101	8	1/1/31/33	0/18/193/193	-
8	COM	F	1102	4	-	0/4/4/4	-
14	TRS	E	501	-	-	5/9/9/9	-
4	USN	A	1201	8	1/1/31/33	0/18/193/193	-
12	UWT	F	1103	-	-	4/12/12/12	-
7	TP7	A	1204	-	-	1/20/24/24	-
5	GOL	A	1211	-	-	2/4/4/4	-
5	GOL	B	501	-	-	0/4/4/4	-
5	GOL	D	703	-	-	2/4/4/4	-
5	GOL	F	1107	-	-	2/4/4/4	-
8	COM	A	1205	4	-	0/4/4/4	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	1202	-	-	2/4/4/4	-
5	GOL	D	702	-	-	0/4/4/4	-
7	TP7	D	701	-	-	1/20/24/24	-
12	UWT	C	301	-	-	1/9/9/12	-

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	1101	USN	NI-NA	11.60	2.14	1.89
4	A	1201	USN	NI-NA	10.24	2.11	1.89
4	A	1201	USN	NI-ND	8.56	2.08	1.89
4	F	1101	USN	CHD-C1D	-8.50	1.32	1.43
4	F	1101	USN	NI-NB	8.29	2.07	1.89

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1201	USN	O7B-C6B-C8B	-7.86	117.00	126.59
4	F	1101	USN	CAA-CBA-CCA	-7.68	97.08	113.59
4	A	1201	USN	C9A-C2A-C3A	5.72	121.42	112.98
4	F	1101	USN	C9A-C2A-C3A	5.03	120.41	112.98
4	A	1201	USN	C9B-C2B-C8B	-4.99	97.86	110.45

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1201	USN	NC
4	F	1101	USN	NC

5 of 20 torsion outliers are listed below:

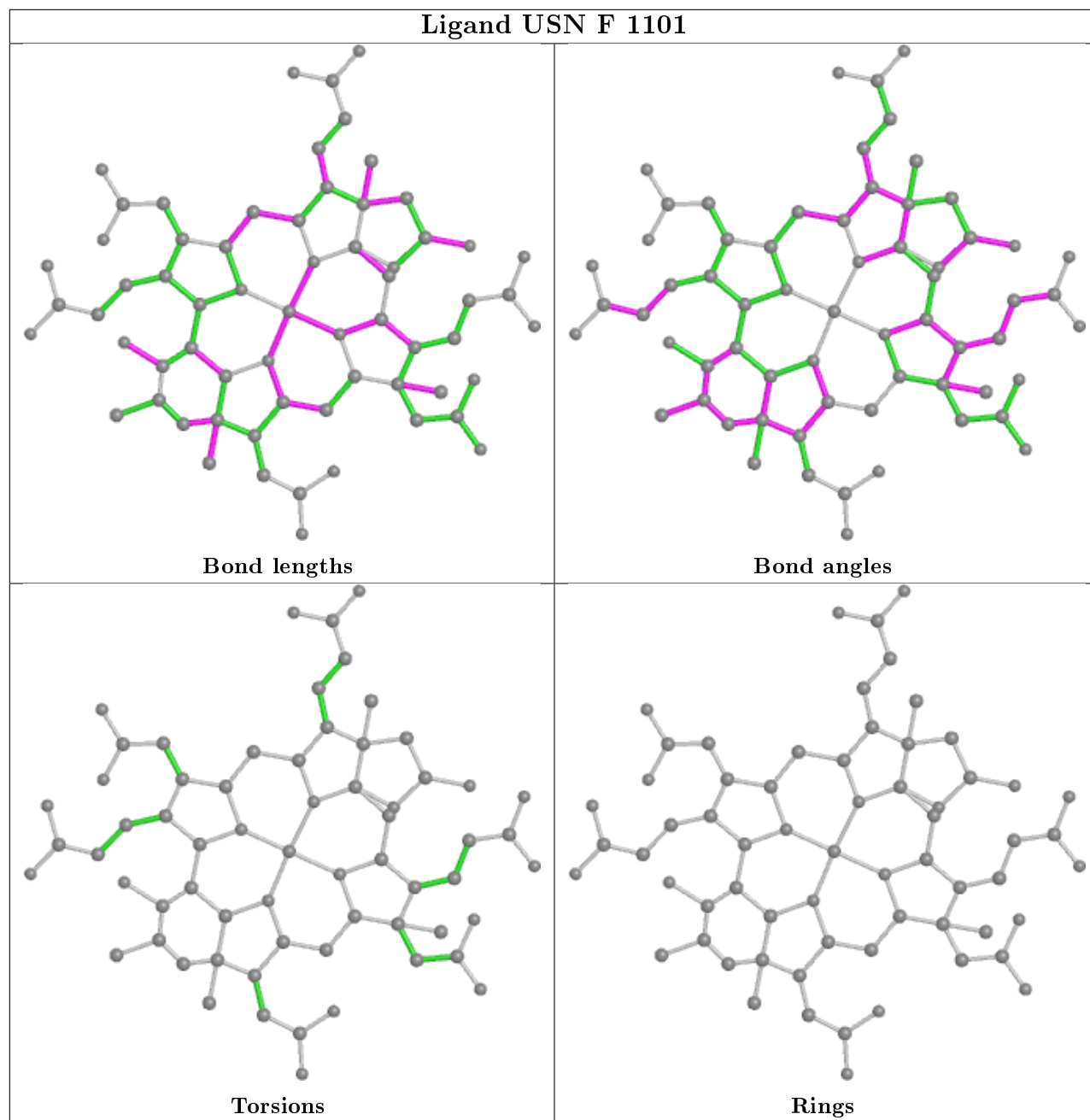
Mol	Chain	Res	Type	Atoms
5	A	1202	GOL	O1-C1-C2-C3
5	D	703	GOL	O1-C1-C2-C3
5	F	1107	GOL	C1-C2-C3-O3
12	F	1103	UWT	O1-C3-C4-O2
12	F	1103	UWT	C2-C1-O1-C3

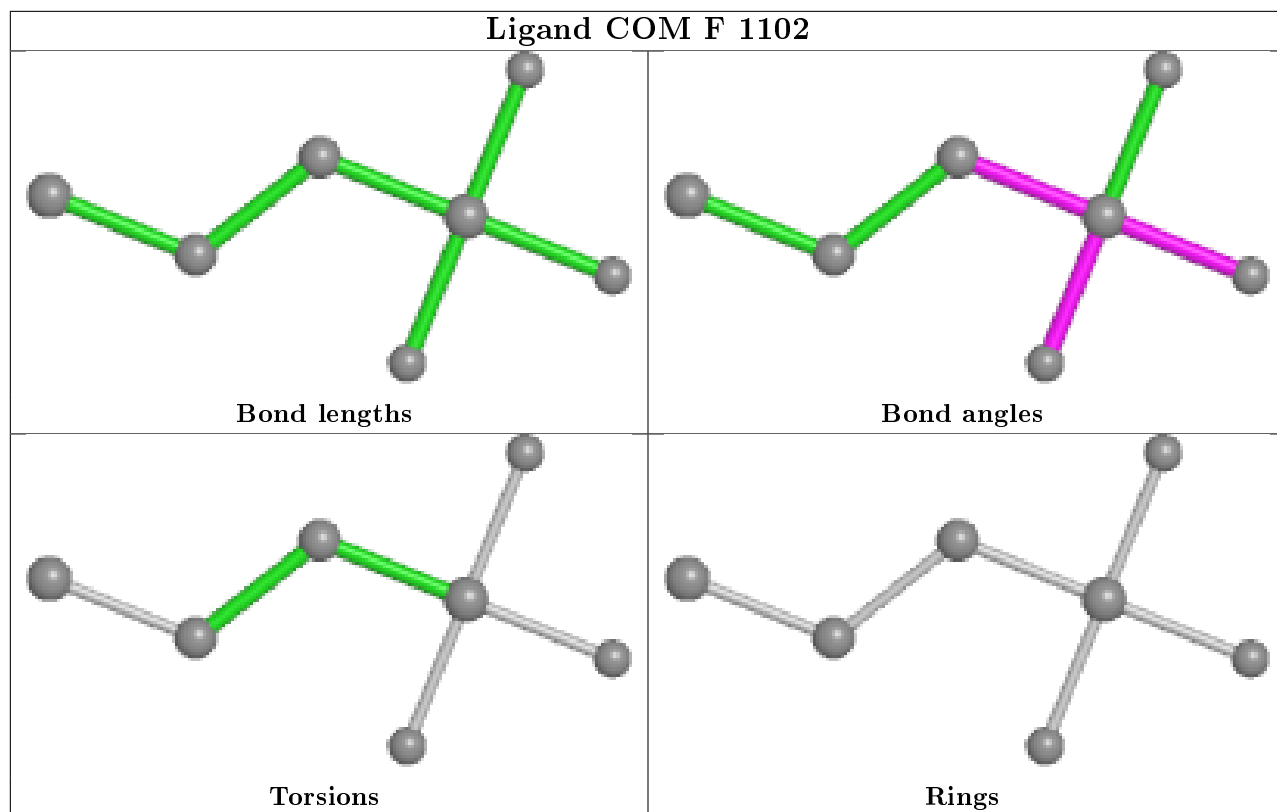
There are no ring outliers.

5 monomers are involved in 4 short contacts:

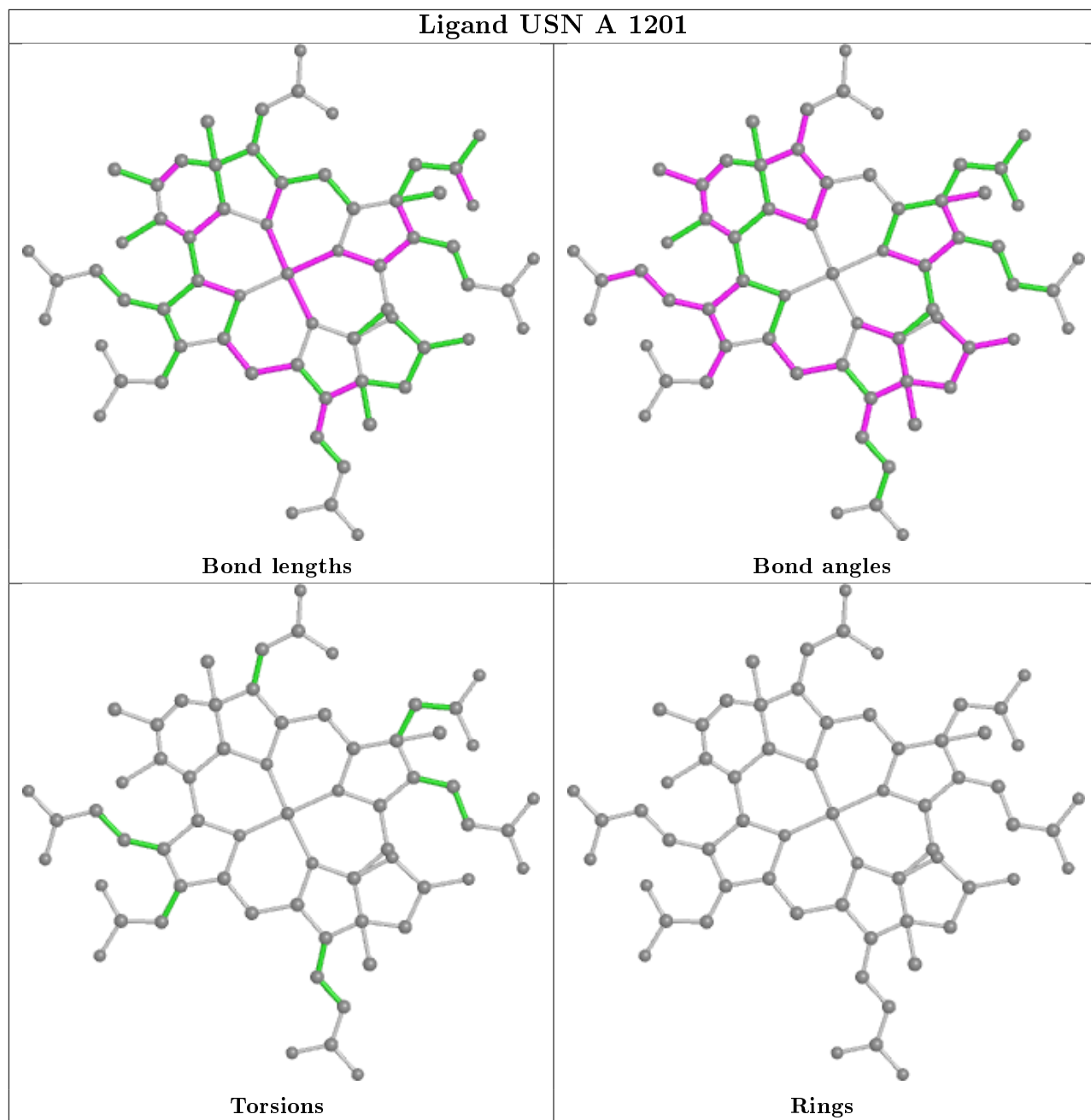
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	1101	USN	1	0
8	F	1102	COM	1	0
14	E	501	TRS	1	0
5	B	501	GOL	1	0
5	A	1202	GOL	1	0

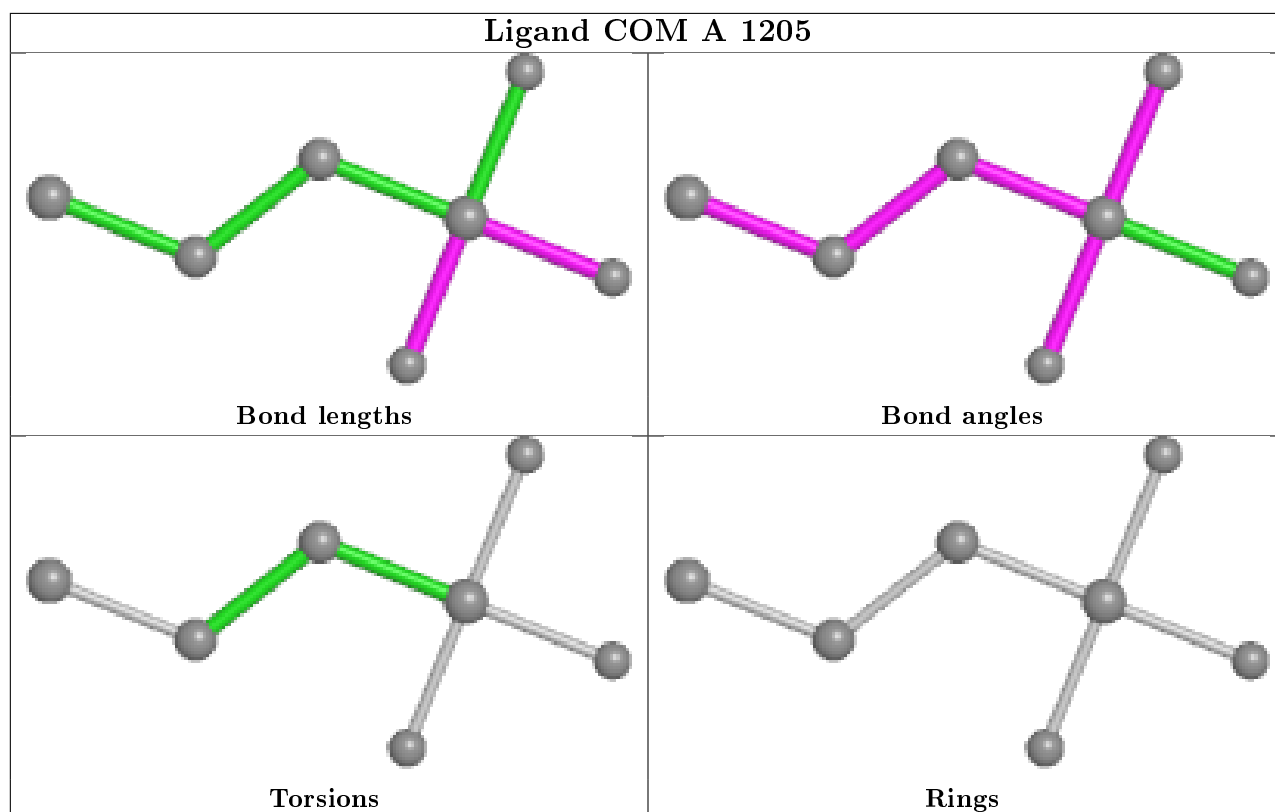
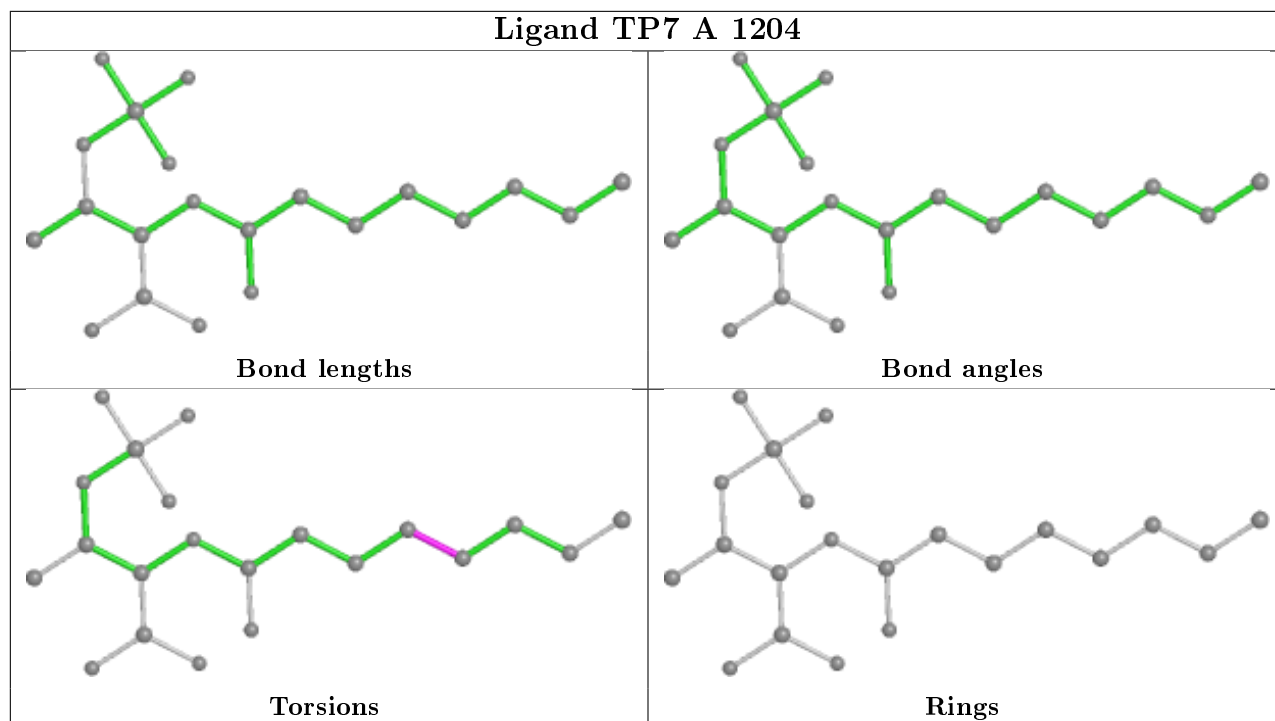
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

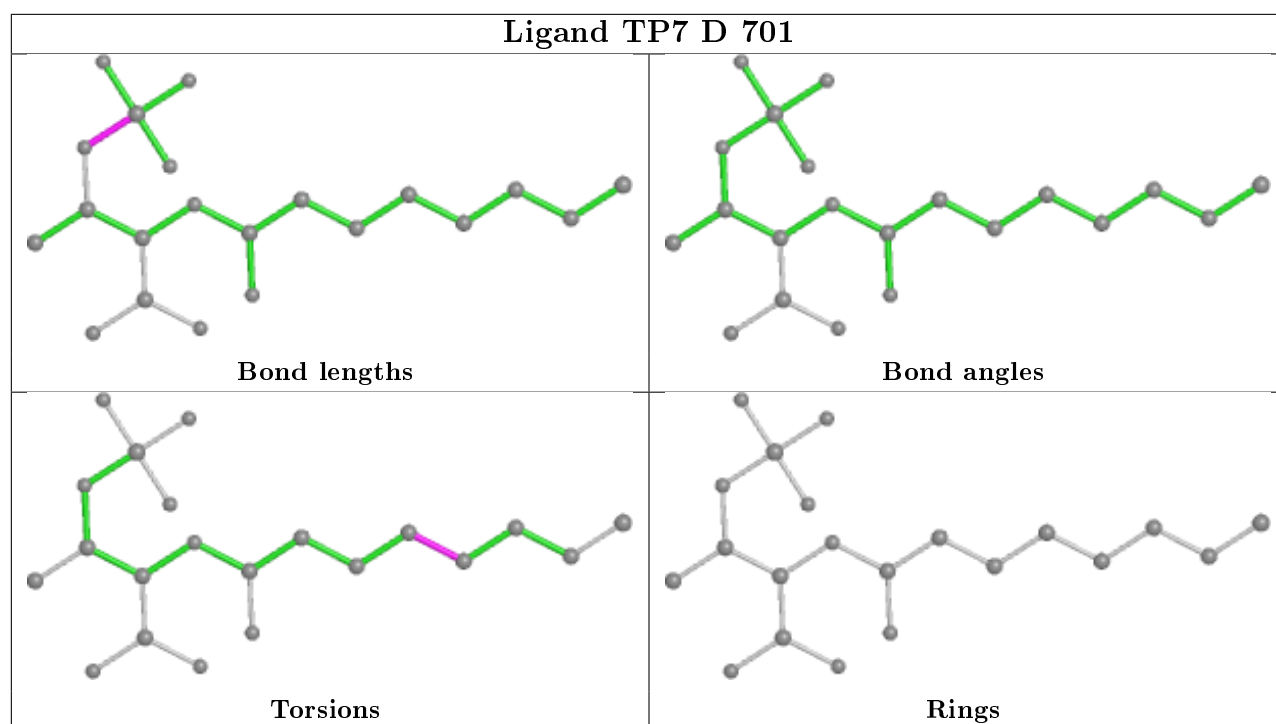












## 5.7 Other polymers [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	584/595 (98%)	-0.37	3 (0%) 91 89	7, 12, 25, 46	0
1	D	586/595 (98%)	-0.34	6 (1%) 82 80	8, 13, 26, 52	0
2	B	466/467 (99%)	-0.21	10 (2%) 63 59	7, 13, 32, 75	0
2	E	466/467 (99%)	-0.22	3 (0%) 89 87	8, 13, 28, 45	0
3	C	265/266 (99%)	-0.20	1 (0%) 92 90	9, 19, 35, 47	0
3	F	265/266 (99%)	-0.26	1 (0%) 92 90	9, 18, 31, 44	0
All	All	2632/2656 (99%)	-0.28	24 (0%) 84 82	7, 14, 29, 75	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	88	TYR	6.9
2	B	87	LYS	6.7
2	B	85	GLY	6.3
2	B	84	PRO	5.5
3	C	2	VAL	5.4

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	MHS	A	291	11/12	0.93	0.10	8,10,13,14	0
1	MGN	A	445	10/11	0.94	0.09	9,10,12,12	0
1	I2M	A	377	9/10	0.95	0.12	8,9,10,10	0
1	HIC	A	491	11/12	0.95	0.10	7,9,10,10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	HIC	D	491	11/12	0.95	0.10	8,8,9,10	0
1	AGM	A	305	12/13	0.96	0.11	6,8,9,11	0
1	SMC	A	354	7/8	0.96	0.09	7,7,9,11	0
1	MHS	D	291	11/12	0.97	0.08	10,12,13,16	0
1	MGN	D	445	10/11	0.97	0.07	8,9,11,12	0
1	AGM	D	305	12/13	0.97	0.14	7,9,10,11	0
1	I2M	D	377	9/10	0.97	0.09	8,9,12,12	0
1	SMC	D	354	7/8	0.98	0.08	8,8,10,10	0
1	GL3	D	490	4/5	0.98	0.08	8,8,9,9	0
1	GL3	A	490	4/5	0.99	0.06	8,8,8,9	0

### 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, 95<sup>th</sup> 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
5	GOL	D	703	6/6	0.68	0.18	46,48,48,49	0
5	GOL	F	1107	6/6	0.69	0.20	45,45,45,45	0
5	GOL	A	1202	6/6	0.72	0.17	46,47,48,48	0
10	XE	E	514[B]	1/1	0.83	0.08	46,46,46,46	1
5	GOL	A	1211	6/6	0.84	0.19	50,50,50,50	0
12	UWT	F	1103	13/13	0.84	0.15	21,29,40,41	0
14	TRS	E	501	8/8	0.84	0.22	33,35,37,38	0
13	MG	D	711	1/1	0.85	0.19	38,38,38,38	0
12	UWT	C	301	11/13	0.86	0.14	23,29,36,37	0
5	GOL	D	702	6/6	0.90	0.13	52,53,53,53	0
9	K	C	303	1/1	0.91	0.07	41,41,41,41	0
9	K	E	507	1/1	0.92	0.08	32,32,32,32	0
5	GOL	B	501	6/6	0.92	0.23	23,25,26,28	0
9	K	D	707	1/1	0.92	0.06	42,42,42,42	0
10	XE	B	510	1/1	0.93	0.06	46,46,46,46	1
9	K	B	502	1/1	0.93	0.05	32,32,32,32	0
9	K	E	508	1/1	0.93	0.06	68,68,68,68	0
9	K	E	510	1/1	0.94	0.06	57,57,57,57	0

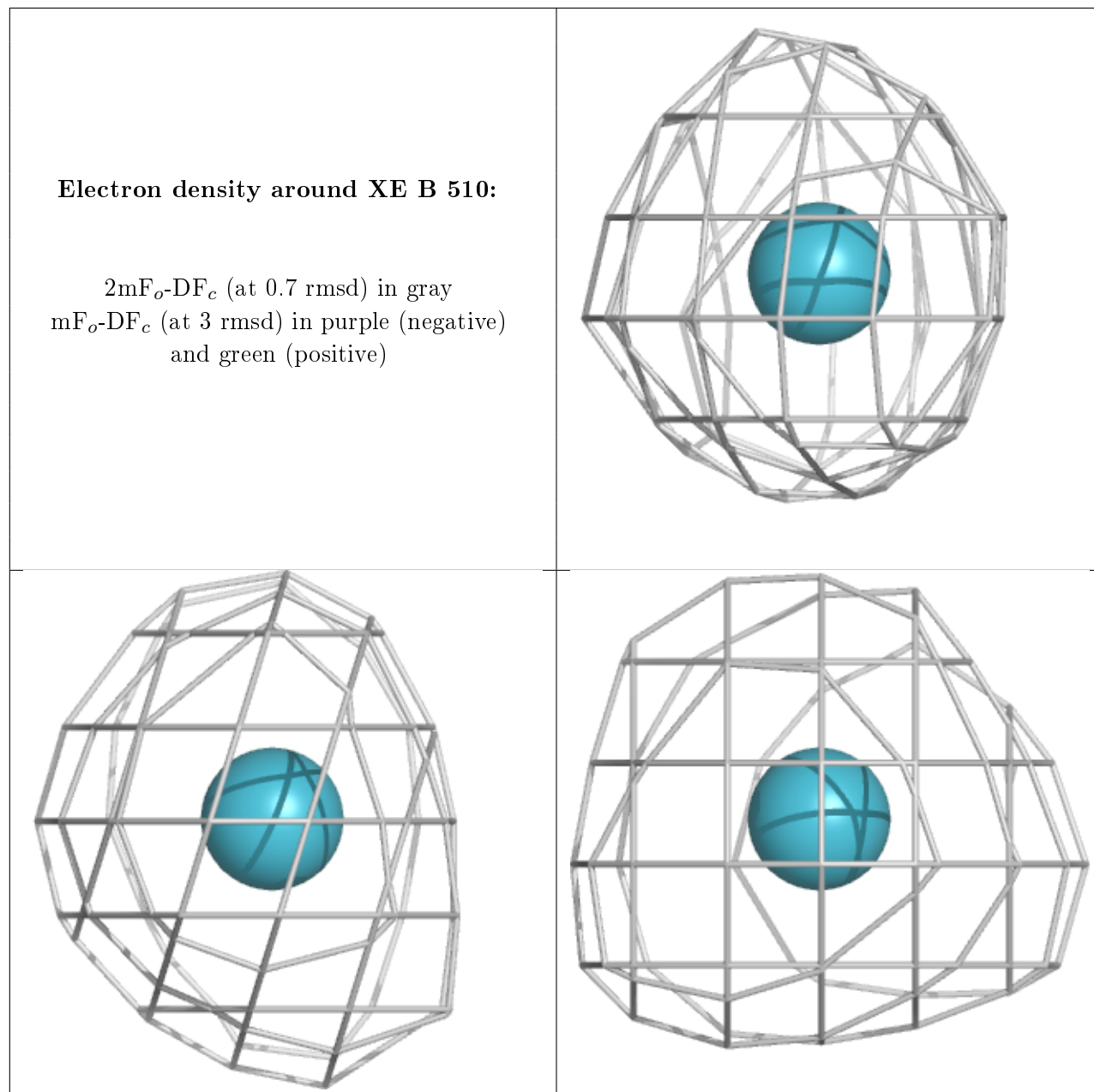
*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
11	CL	E	504	1/1	0.94	0.26	42,42,42,42	0
10	XE	A	1210	1/1	0.95	0.07	32,32,32,32	1
9	K	B	503	1/1	0.95	0.07	29,29,29,29	0
10	XE	E	513	1/1	0.95	0.05	47,47,47,47	1
9	K	E	502	1/1	0.96	0.14	42,42,42,42	0
9	K	E	509	1/1	0.96	0.07	33,33,33,33	0
7	TP7	A	1204	21/21	0.96	0.10	7,10,14,16	0
7	TP7	D	701	21/21	0.97	0.11	7,10,14,17	0
10	XE	F	1106[A]	1/1	0.97	0.10	30,30,30,30	1
11	CL	B	505	1/1	0.97	0.04	26,26,26,26	0
8	COM	F	1102	7/7	0.97	0.07	8,9,10,11	0
9	K	A	1206	1/1	0.97	0.11	38,38,38,38	0
4	USN	F	1101	64/64	0.97	0.08	8,9,13,15	0
10	XE	C	305[A]	1/1	0.97	0.09	34,34,34,34	1
4	USN	A	1201	64/64	0.97	0.08	7,10,14,16	0
10	XE	D	710	1/1	0.98	0.05	25,25,25,25	1
6	NA	D	706	1/1	0.98	0.05	18,18,18,18	0
6	NA	A	1203	1/1	0.98	0.06	15,15,15,15	0
10	XE	F	1105	1/1	0.98	0.07	15,15,15,15	1
9	K	B	507	1/1	0.98	0.05	32,32,32,32	0
9	K	B	508	1/1	0.98	0.05	20,20,20,20	0
9	K	F	1104	1/1	0.98	0.11	26,26,26,26	0
6	NA	C	302	1/1	0.98	0.06	10,10,10,10	0
6	NA	D	704	1/1	0.98	0.07	9,9,9,9	0
10	XE	B	511[B]	1/1	0.98	0.10	41,41,41,41	1
6	NA	D	705	1/1	0.98	0.04	14,14,14,14	0
10	XE	A	1209	1/1	0.99	0.04	11,11,11,11	1
8	COM	A	1205	7/7	0.99	0.06	7,8,9,9	0
10	XE	B	509	1/1	0.99	0.06	12,12,12,12	1
9	K	A	1207	1/1	0.99	0.06	23,23,23,23	0
11	CL	B	504	1/1	0.99	0.02	19,19,19,19	0
9	K	E	506	1/1	0.99	0.10	31,31,31,31	0
11	CL	E	503	1/1	0.99	0.03	16,16,16,16	0
10	XE	C	304	1/1	0.99	0.06	18,18,18,18	1
11	CL	E	505	1/1	0.99	0.05	22,22,22,22	0
9	K	E	511	1/1	0.99	0.04	19,19,19,19	0
10	XE	D	709	1/1	0.99	0.05	13,13,13,13	1
9	K	B	506	1/1	0.99	0.06	25,25,25,25	0
10	XE	E	512	1/1	0.99	0.05	13,13,13,13	1
10	XE	A	1208	1/1	1.00	0.03	13,13,13,13	1
10	XE	D	708	1/1	1.00	0.06	10,10,10,10	1

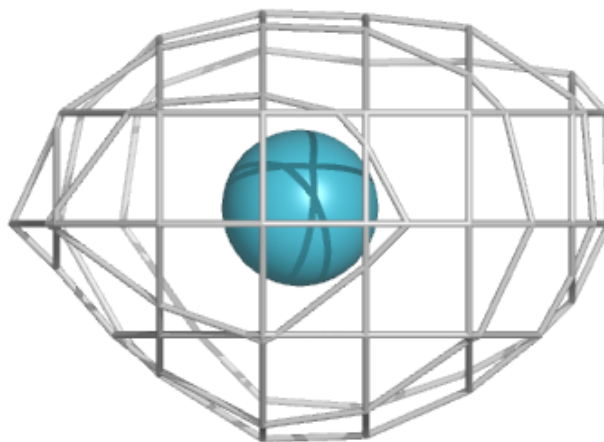
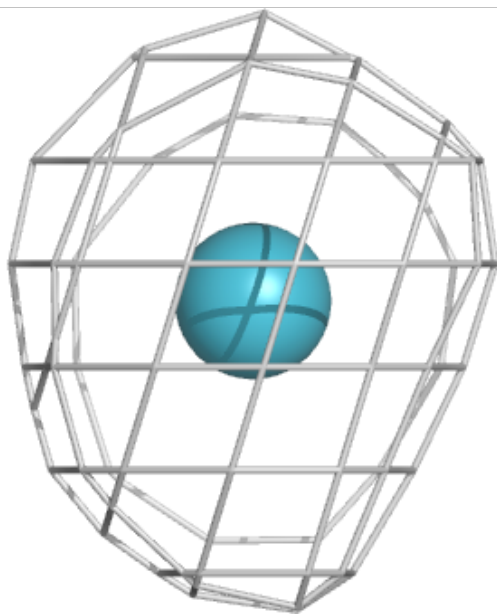
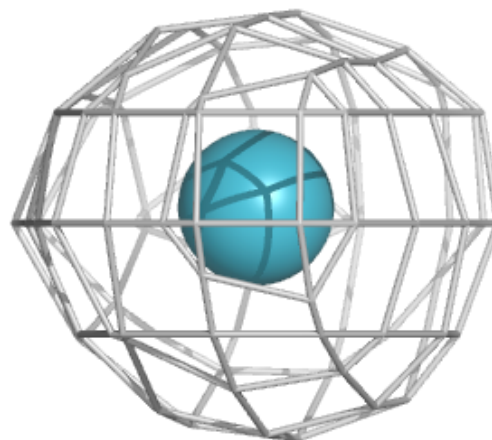
The following is a graphical depiction of the model fit to experimental electron density of all

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 XE E 513:**

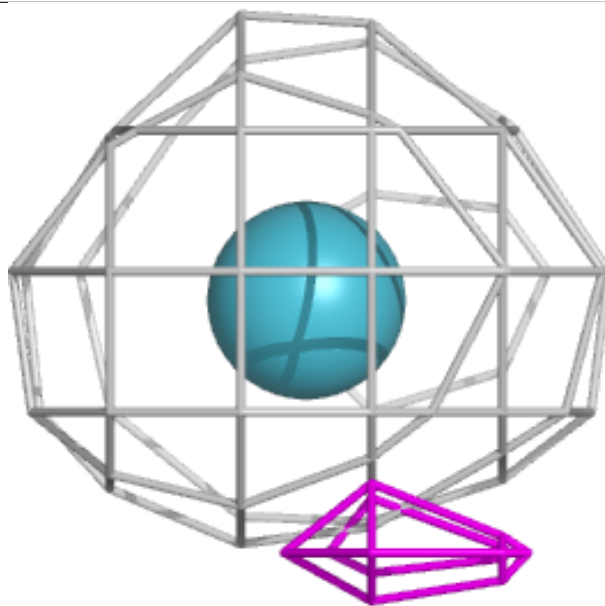
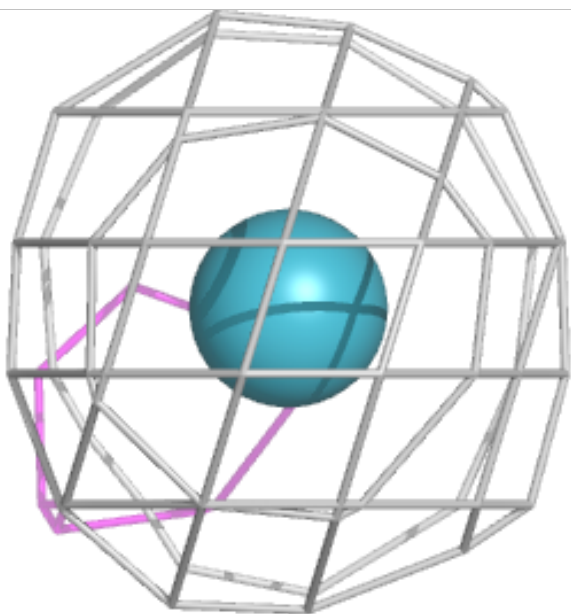
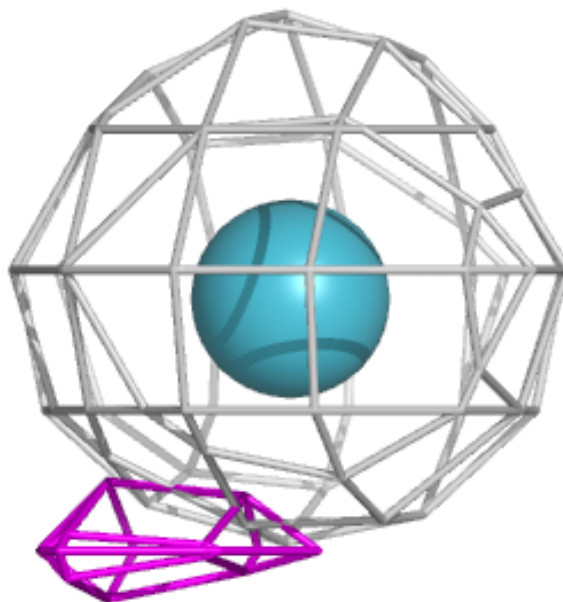
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





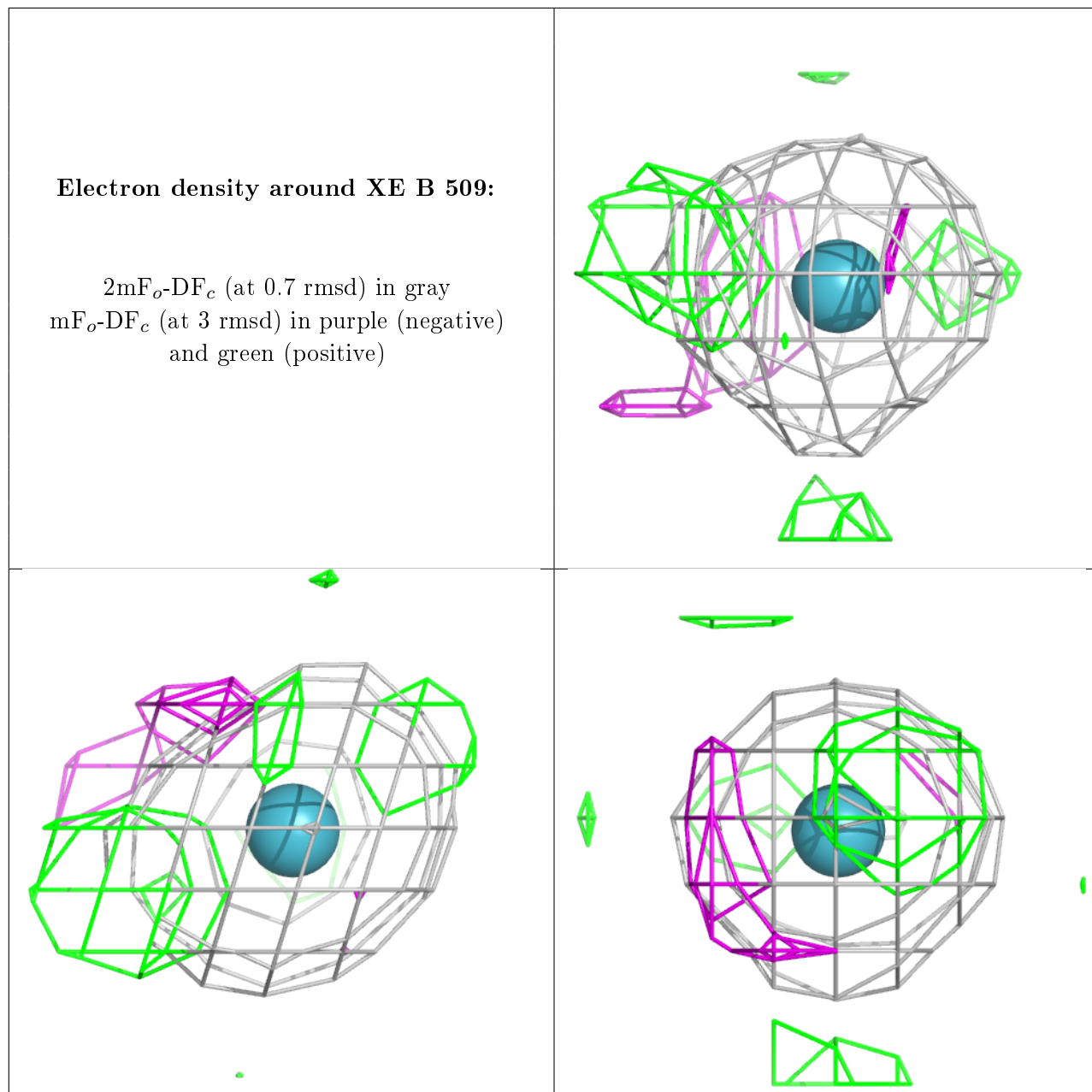
**Electron density around XE D 710:**

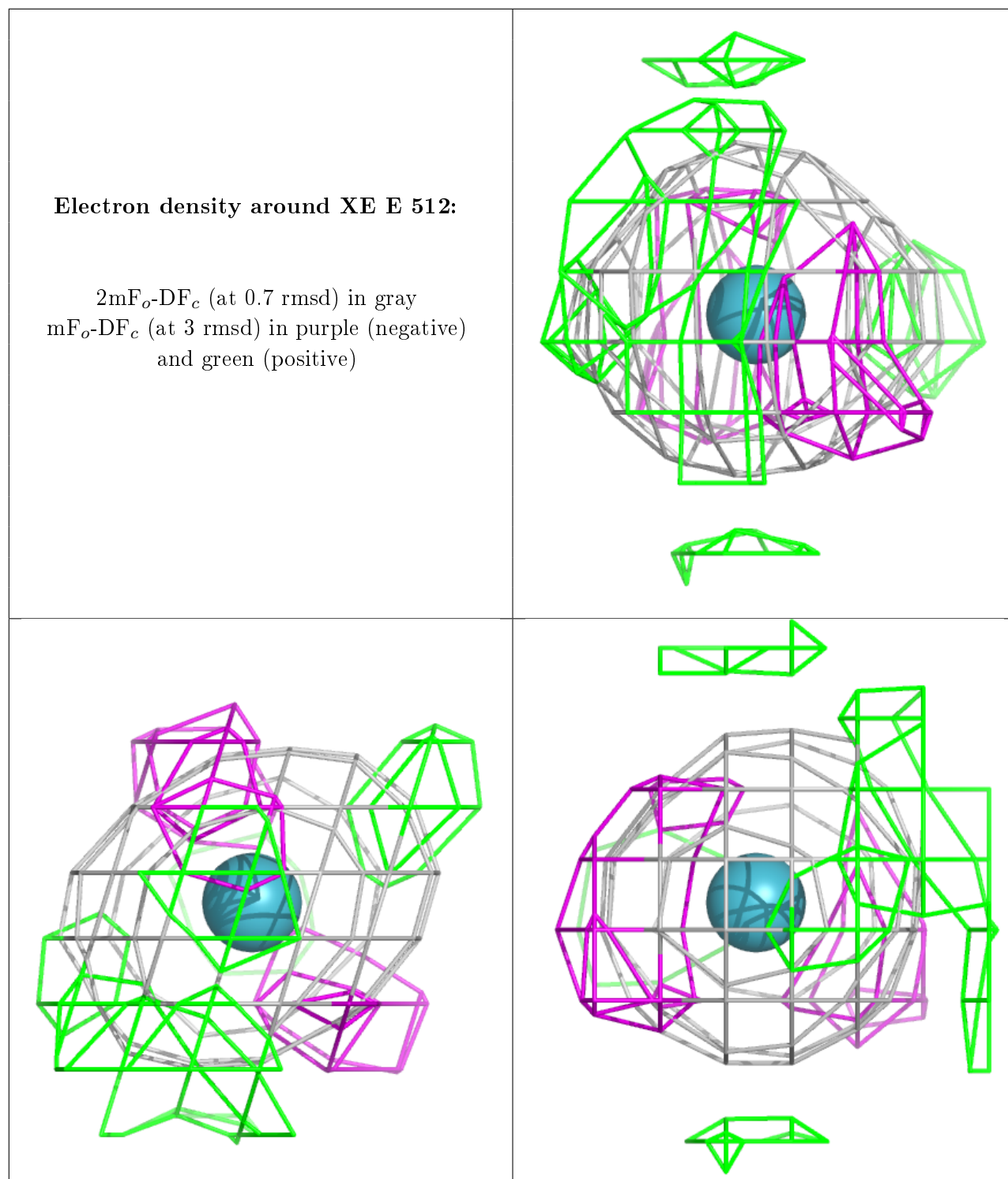
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around XE B 509:**

$2mF_o-DF_c$  (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.