



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 02:53 PM BST

PDB ID : 6AT4  
Title : E. coli phosphoenolpyruvate carboxykinase bound to thiosulfate  
Authors : Tang, H.Y.H.; Shin, D.S.; Tainer, J.A.  
Deposited on : 2017-08-27  
Resolution : 1.33 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](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 : 2.13.1  
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) : 2.13.1

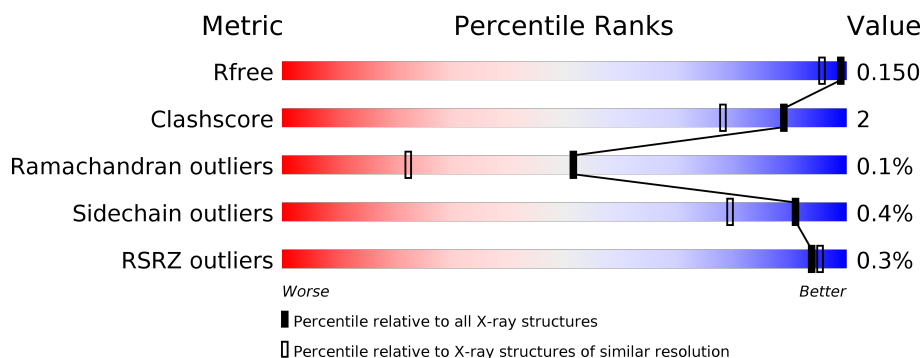
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.33 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	1385 (1.36-1.32)
Clashscore	141614	1417 (1.36-1.32)
Ramachandran outliers	138981	1397 (1.36-1.32)
Sidechain outliers	138945	1397 (1.36-1.32)
RSRZ outliers	127900	1369 (1.36-1.32)

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  $\geq 3$ , 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	546	<div> <div style="width: 95%;"></div> <div>95%</div> </div>
1	B	546	<div> <div style="width: 92%;"></div> <div>92%</div> <div style="width: 5%;"></div> <div>5%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17105 atoms, of which 7884 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.

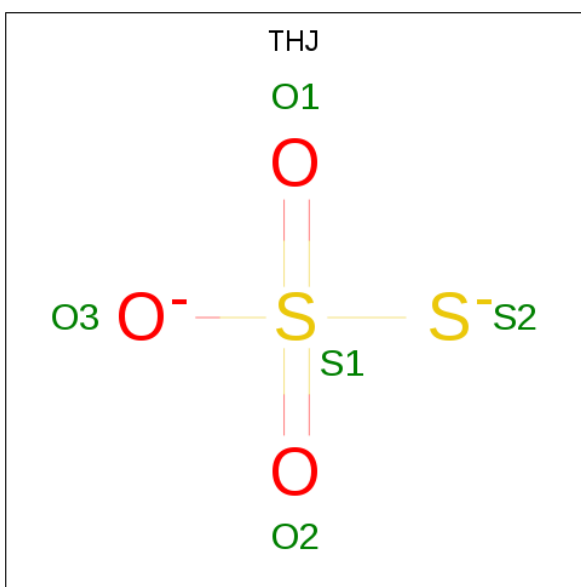
- Molecule 1 is a protein called Phosphoenolpyruvate carboxykinase (ATP).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	535	Total	C	H	N	O	S	0	6	0
			8140	2656	3971	700	797	16			
1	B	534	Total	C	H	N	O	S	0	3	0
			8014	2615	3913	692	779	15			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	541	HIS	-	expression tag	UNP P22259
A	542	HIS	-	expression tag	UNP P22259
A	543	HIS	-	expression tag	UNP P22259
A	544	HIS	-	expression tag	UNP P22259
A	545	HIS	-	expression tag	UNP P22259
A	546	HIS	-	expression tag	UNP P22259
B	541	HIS	-	expression tag	UNP P22259
B	542	HIS	-	expression tag	UNP P22259
B	543	HIS	-	expression tag	UNP P22259
B	544	HIS	-	expression tag	UNP P22259
B	545	HIS	-	expression tag	UNP P22259
B	546	HIS	-	expression tag	UNP P22259

- Molecule 2 is THIOSULFATE (three-letter code: THJ) (formula: O<sub>3</sub>S<sub>2</sub>) (labeled as "Ligand of Interest" by author).



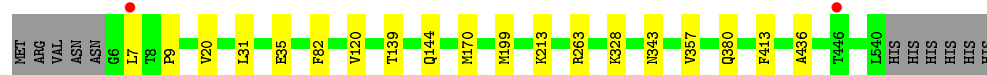
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	3	2		
2	A	1	Total	O	S	0	0
			5	3	2		
2	A	1	Total	O	S	0	0
			5	3	2		
2	B	1	Total	O	S	0	0
			5	3	2		
2	B	1	Total	O	S	0	0
			5	3	2		

- Molecule 3 is water.

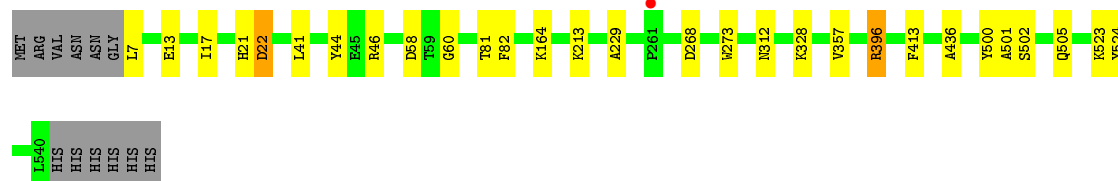
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	473	Total	O	0	0
			473	473		
3	B	453	Total	O	0	0
			453	453		



- Molecule 1: Phosphoenolpyruvate carboxykinase (ATP)



- Chain B:  92% 5% .



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.72Å 103.61Å 119.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	78.19 – 1.33 78.19 – 1.33	Depositor EDS
% Data completeness (in resolution range)	90.7 (78.19-1.33) 84.6 (78.19-1.33)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.58 (at 1.33Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, $R_{free}$	0.135 , 0.151 0.134 , 0.150	Depositor DCC
$R_{free}$ test set	1969 reflections (0.82%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.2	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.45 , 51.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	17105	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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: THJ

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	0.31	0/4277	0.56	0/5815
1	B	0.31	0/4207	0.57	2/5725 (0.0%)
All	All	0.31	0/8484	0.56	2/11540 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	396	ARG	NE-CZ-NH2	6.30	123.45	120.30
1	B	396	ARG	NE-CZ-NH1	-5.69	117.45	120.30

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	A	4169	3971	4034	12	0
1	B	4101	3913	3962	19	0
2	A	15	0	0	0	0
2	B	10	0	0	0	0
3	A	473	0	0	3	0
3	B	453	0	0	7	1
All	All	9221	7884	7996	31	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (31) 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:B:46:ARG:NH1	3:B:702:HOH:O	2.02	0.92
1:B:500:TYR:O	3:B:701:HOH:O	2.00	0.79
1:B:502:SER:N	1:B:505:GLN:OE1	2.17	0.75
1:B:7:LEU:N	3:B:703:HOH:O	2.21	0.73
1:B:501:ALA:N	1:B:505:GLN:OE1	2.22	0.73
3:A:1186:HOH:O	1:B:396:ARG:HD2	1.97	0.65
1:B:44:TYR:CZ	1:B:60:GLY:HA3	2.40	0.55
1:B:213:LYS:HE2	1:B:413:PHE:CZ	2.42	0.54
1:A:213:LYS:HE2	1:A:413:PHE:CZ	2.45	0.51
1:B:164:LYS:NZ	3:B:707:HOH:O	2.42	0.50
1:A:139[B]:THR:OG1	1:A:144:GLN:OE1	2.24	0.50
1:A:9:PRO:HB3	1:A:20:VAL:O	2.12	0.49
1:B:312:ASN:OD1	1:B:328:LYS:HE2	2.13	0.47
1:B:21:HIS:O	1:B:22:ASP:HB2	2.15	0.47
1:B:81:THR:HG23	3:B:730:HOH:O	2.15	0.46
1:A:263:ARG:O	1:A:343:ASN:ND2	2.47	0.45
1:B:13:GLU:HA	1:B:17:ILE:O	2.18	0.43
1:A:328:LYS:N	3:A:913:HOH:O	2.51	0.43
1:A:7:LEU:HD22	1:A:120:VAL:HG11	2.01	0.43
1:A:199:MET:HE2	1:A:199:MET:HB2	1.92	0.43
1:A:357:VAL:O	1:A:436:ALA:HA	2.19	0.42
1:B:357:VAL:O	1:B:436:ALA:HA	2.19	0.42
1:A:380:GLN:OE1	3:A:901:HOH:O	2.21	0.42
1:A:170:MET:HE2	1:A:199:MET:SD	2.60	0.42
1:B:523:LYS:HE2	1:B:524:TYR:CZ	2.55	0.42
1:A:31:LEU:HD23	1:A:35:GLU:HG3	2.02	0.41
1:B:41:LEU:O	1:B:46:ARG:HG2	2.20	0.41
1:B:58:ASP:CB	3:B:704:HOH:O	2.68	0.41
1:A:170:MET:CE	1:A:199:MET:SD	3.08	0.41
1:B:164:LYS:NZ	3:B:717:HOH:O	2.53	0.41
1:B:229:ALA:HB3	1:B:273:TRP:HB3	2.03	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:1005:HOH:O	3:B:1095:HOH:O[2_544]	2.02	0.18

## 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	539/546 (99%)	527 (98%)	12 (2%)	0	100	100
1	B	535/546 (98%)	522 (98%)	12 (2%)	1 (0%)	47	21
All	All	1074/1092 (98%)	1049 (98%)	24 (2%)	1 (0%)	51	22

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	22	ASP

### 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	437/454 (96%)	436 (100%)	1 (0%)	93	82
1	B	426/454 (94%)	424 (100%)	2 (0%)	88	71
All	All	863/908 (95%)	860 (100%)	3 (0%)	91	81

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	PHE
1	B	82	PHE
1	B	268	ASP

Some 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](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

5 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	$\# Z  > 2$
2	THJ	B	602	-	2,4,4	1.63	0	2,6,6	1.00	0
2	THJ	B	601	-	2,4,4	2.10	1 (50%)	2,6,6	1.07	0
2	THJ	A	801	-	2,4,4	1.73	0	2,6,6	0.50	0
2	THJ	A	803	-	2,4,4	2.35	1 (50%)	2,6,6	0.65	0
2	THJ	A	802	-	2,4,4	1.95	0	2,6,6	1.04	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	803	THJ	O1-S1	2.74	1.53	1.45
2	B	601	THJ	O1-S1	2.31	1.52	1.45

There are no bond angle outliers.

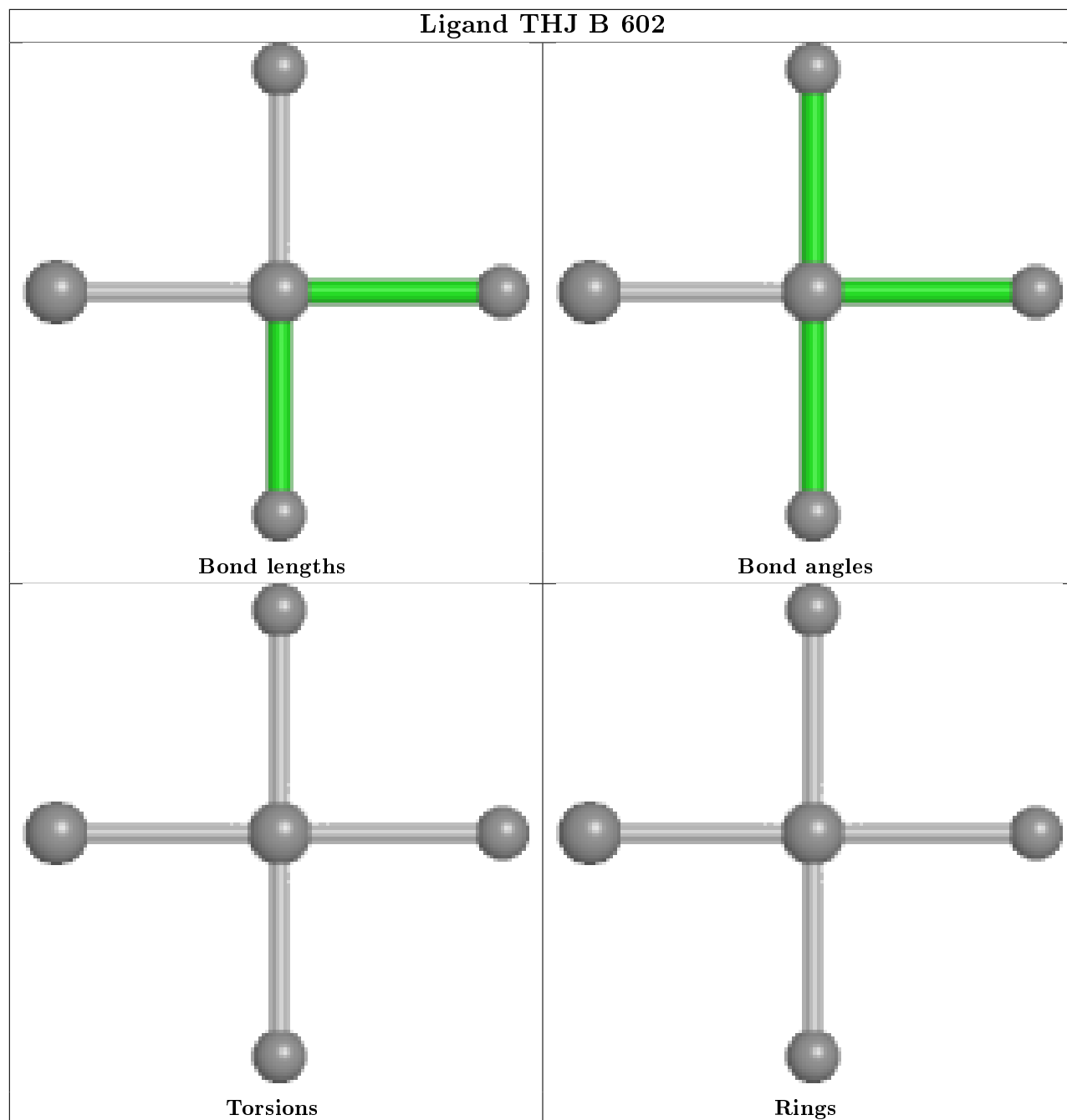
There are no chirality outliers.

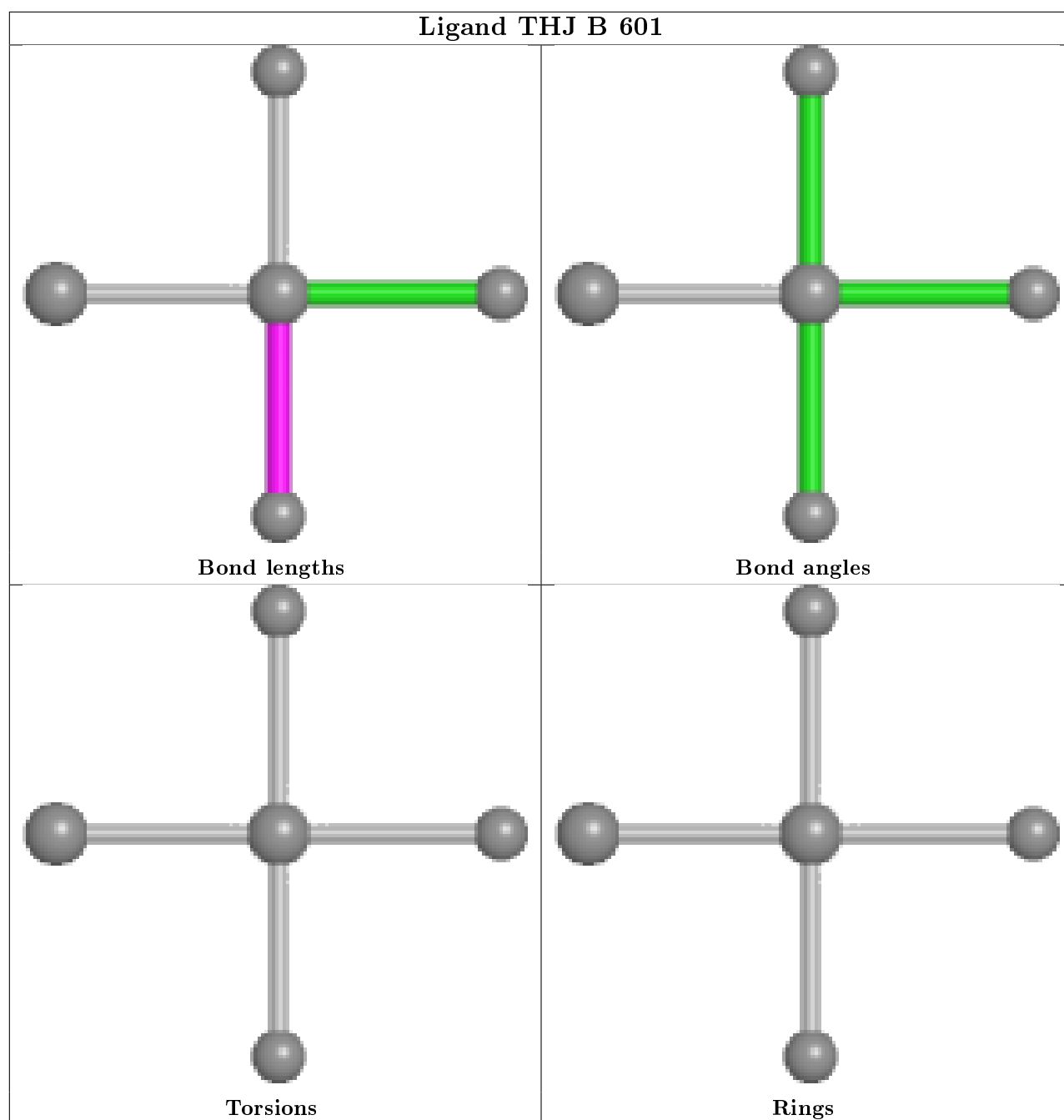
There are no torsion outliers.

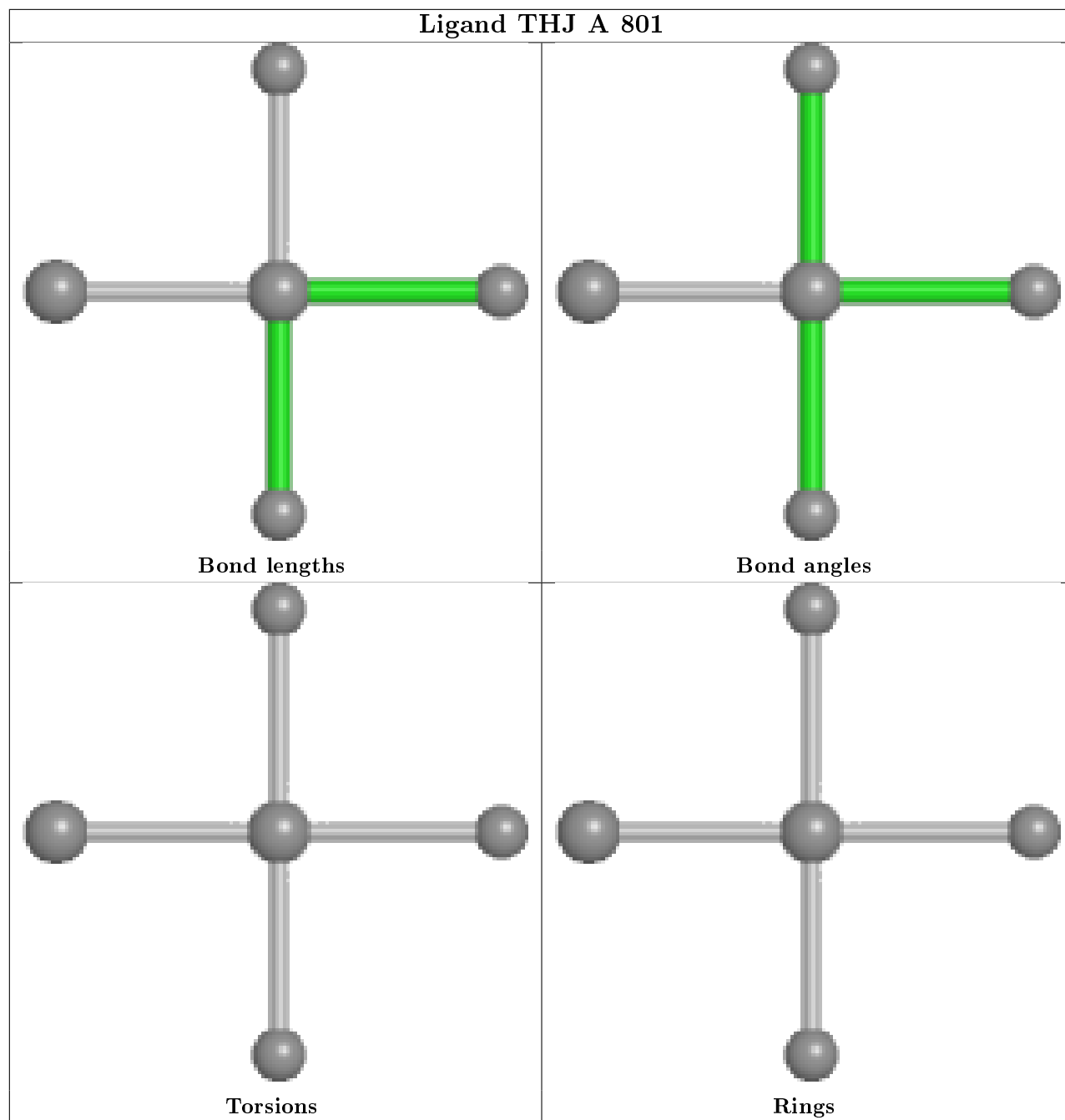
There are no ring outliers.

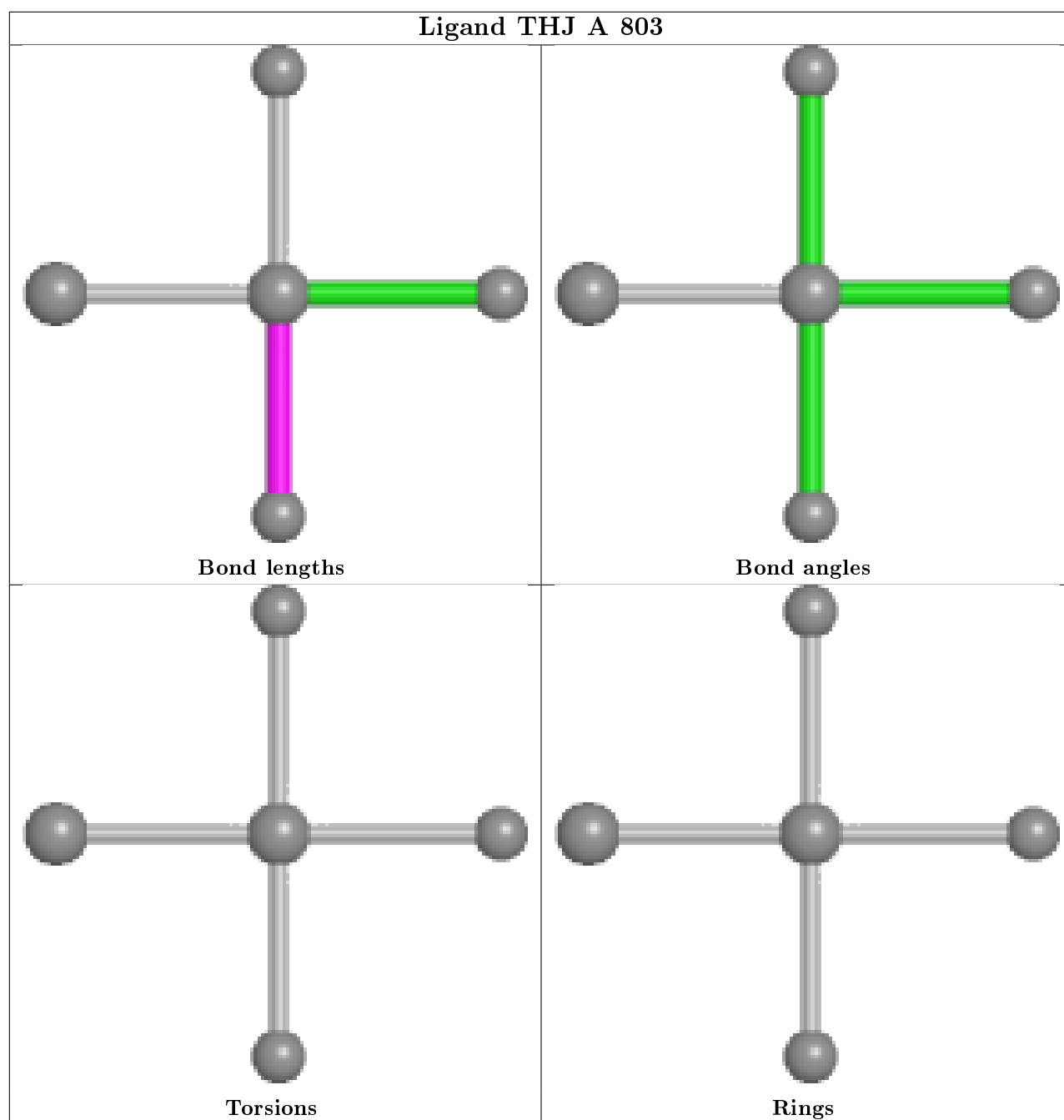
No monomer is involved in short contacts.

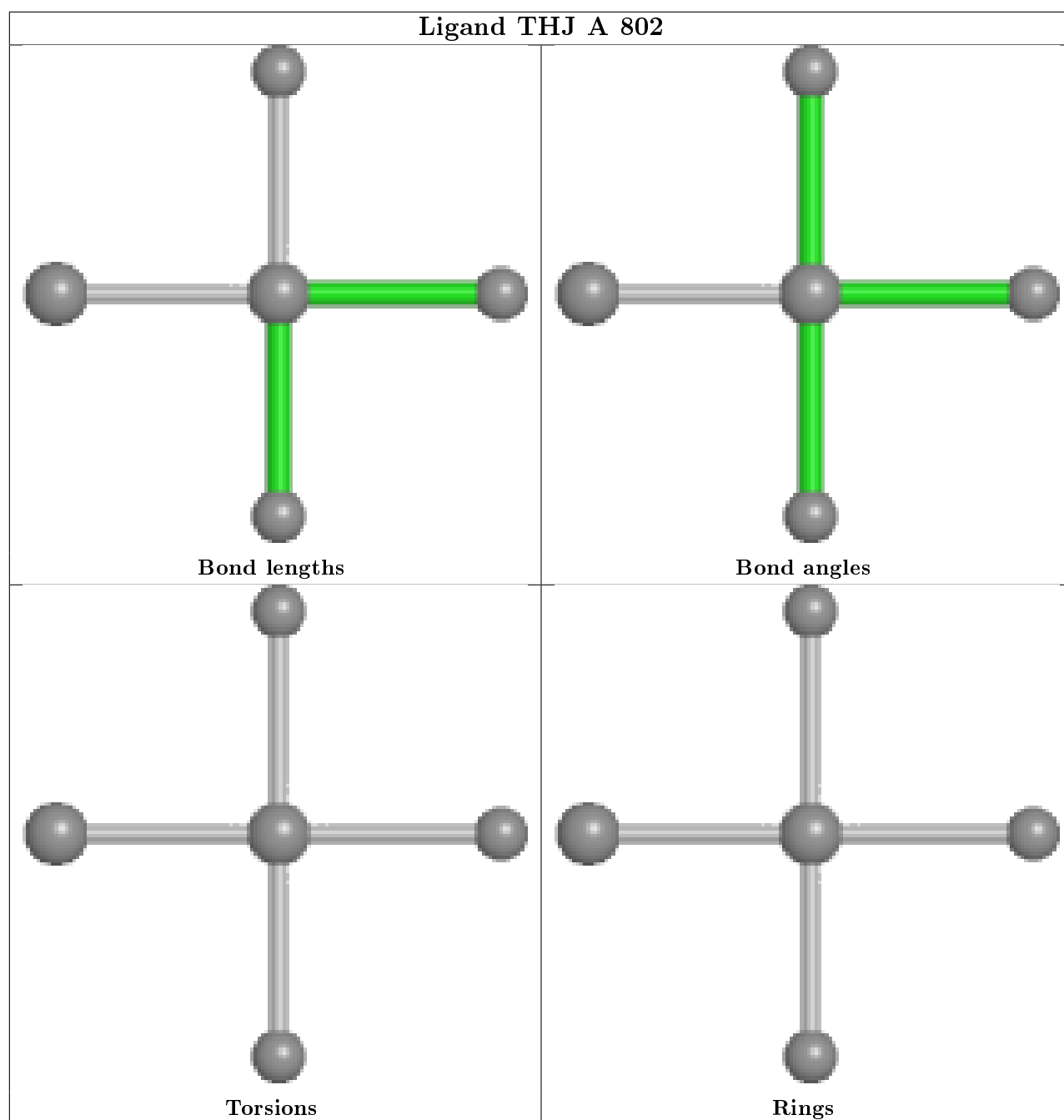
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	535/546 (97%)	-0.51	2 (0%) 92 93	11, 18, 29, 49	0
1	B	534/546 (97%)	-0.45	1 (0%) 95 95	12, 18, 36, 45	0
All	All	1069/1092 (97%)	-0.48	3 (0%) 94 95	11, 18, 34, 49	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	261	PRO	2.8
1	A	7	LEU	2.7
1	A	446	THR	2.1

### 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 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(Å <sup>2</sup> )	Q<0.9
2	THJ	B	602	5/5	0.91	0.12	22,23,23,24	5
2	THJ	A	803	5/5	0.94	0.09	18,20,21,21	5

*Continued on next page...*

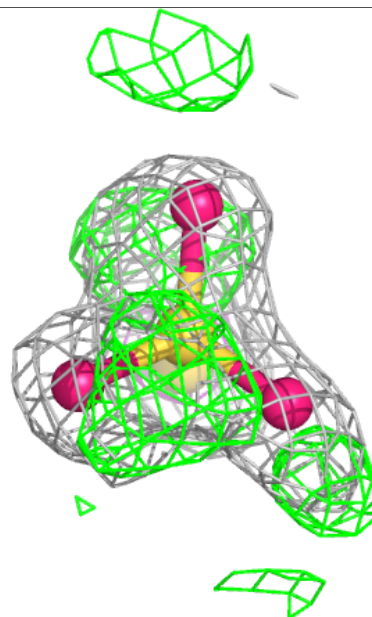
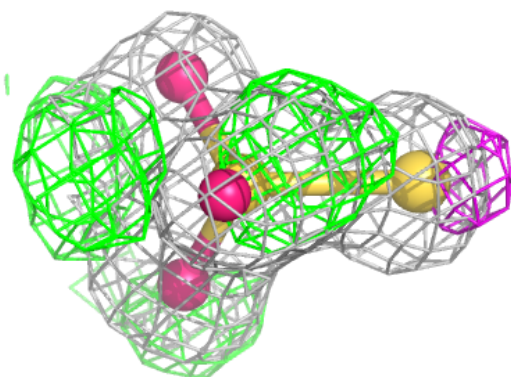
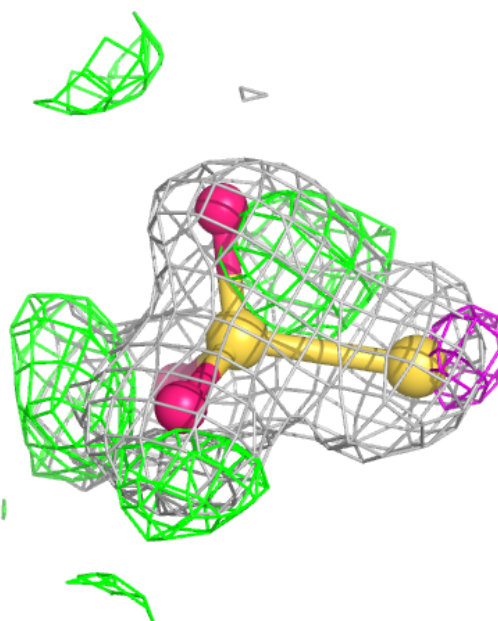
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	THJ	B	601	5/5	0.97	0.07	19,20,21,21	5
2	THJ	A	801	5/5	0.99	0.06	16,17,17,17	0
2	THJ	A	802	5/5	0.99	0.08	21,23,26,26	5

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.

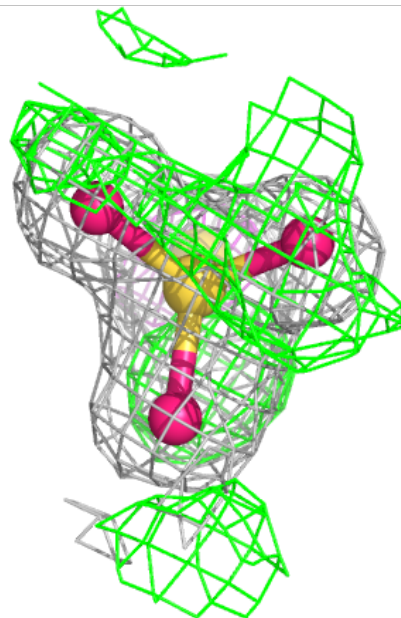
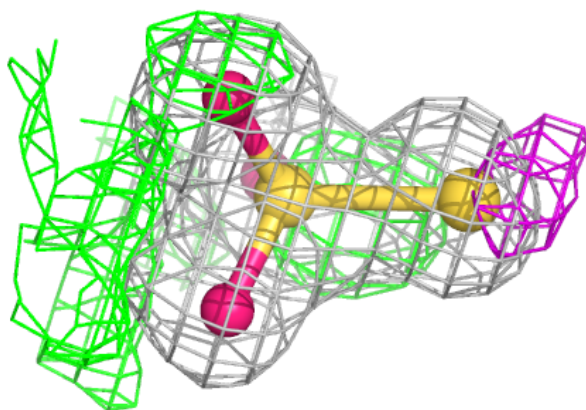
**Electron density around THJ B 602:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



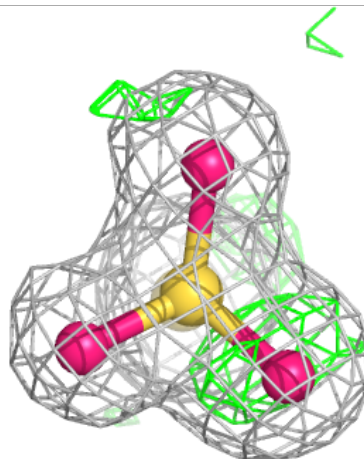
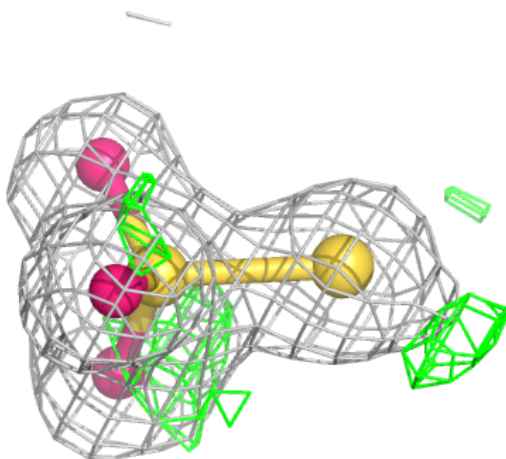
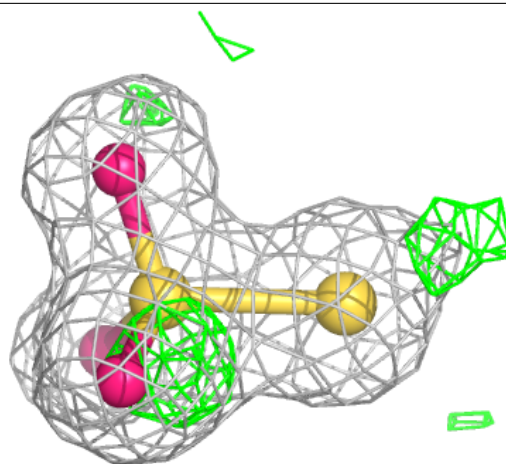
**Electron density around THJ A 803:**

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



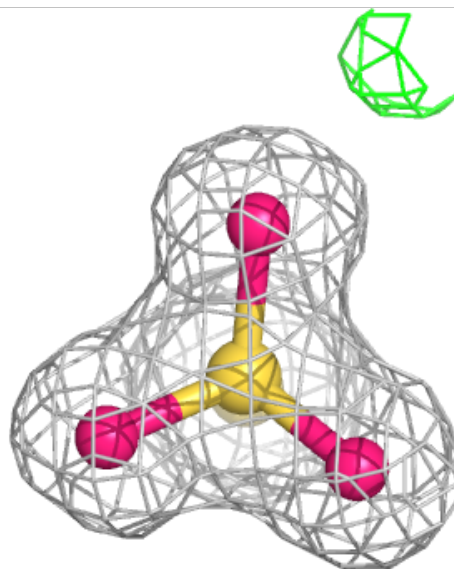
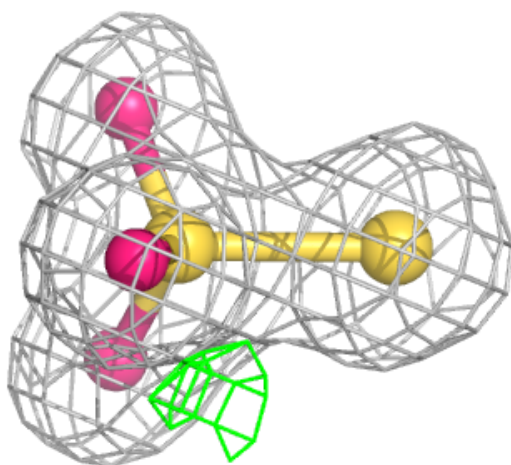
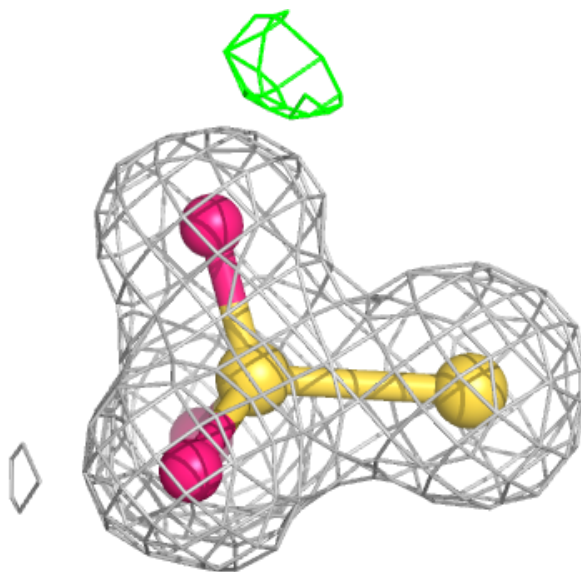
**Electron density around THJ B 601:**

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



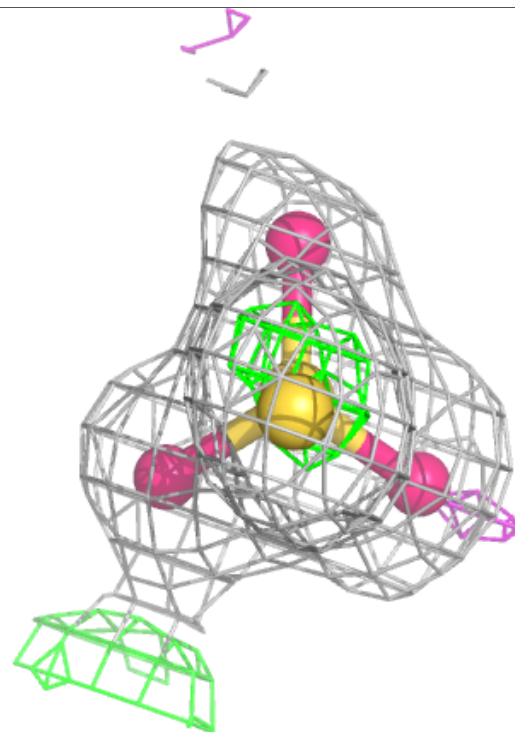
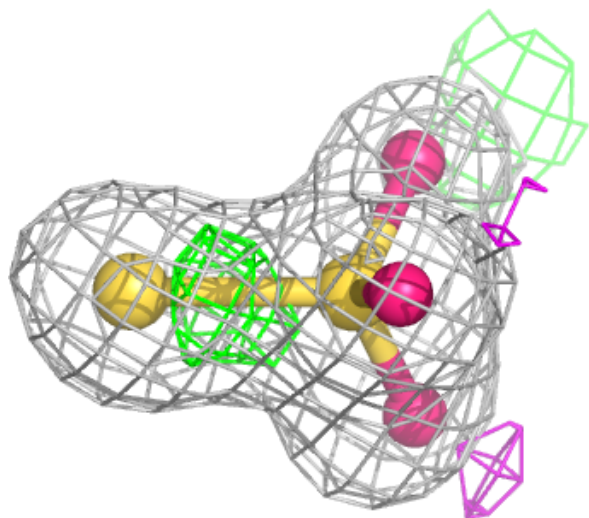
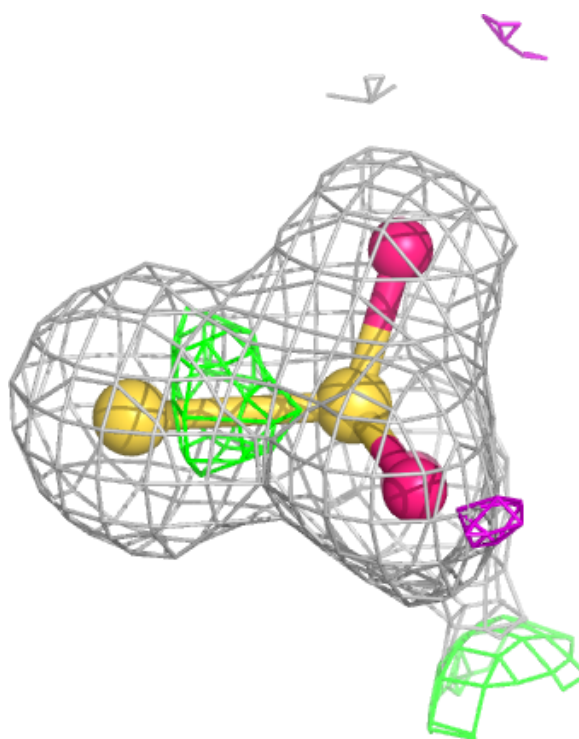
**Electron density around THJ A 801:**

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



**Electron density around THJ A 802:**

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



## 6.5 Other polymers ⓘ

There are no such residues in this entry.