

# wwPDB X-ray Structure Validation Summary Report (i)

#### Apr 15, 2024 – 08:08 PM JST

PDB ID	:	8YL9
Title	:	Crystal structures of terpene synthases complexed with a substrate mimic
Authors	:	Xu, M.; Ma, M.
Deposited on		
Resolution	:	2.75  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:

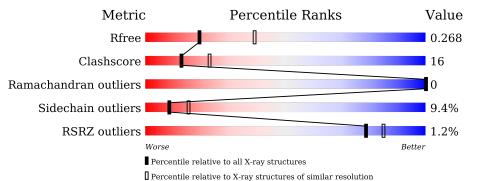
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.36
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.75 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	$\begin{array}{l} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	1235(2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	А	352	% • 66%	29%	•••		
1	В	352	% • 68%	25%	<del>.</del>		
1	Ι	352	.% 62%	31%	5% •		

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
4	BTM	А	405	-	-	Х	Х
4	BTM	В	405	-	-	Х	-
4	BTM	Ι	405	-	-	Х	-



#### $\mathbf{2}$ Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8621 atoms, of which 66 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.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Δ	346	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
	А	340	2791	1760	487	530	14	0		
1	Р	337	Total	С	Ν	0	S	0	0	0
	D	557	2729	1722	476	518	13	0		
1	т	346	Total	С	Ν	0	S	0	0	0
	1	340	2792	1759	487	532	14	0	0	0

• Molecule 1 is a protein called Sesterbrasiliatriene synthase PbSS.

• Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

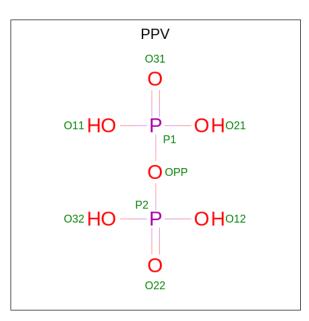
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	3	Total Mg 3 3	0	0
2	В	3	Total Mg 3 3	0	0
2	Ι	3	Total Mg 3 3	0	0

• Molecule 3 is PYROPHOSPHATE (three-letter code: PPV) (formula: H<sub>4</sub>O<sub>7</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



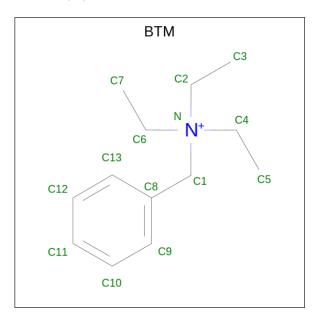
8YL9





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	TotalOP972	0	0
3	В	1	TotalOP972	0	0
3	Ι	1	TotalOP972	0	0

• Molecule 4 is N-benzyl-N,N-diethylethanaminium (three-letter code: BTM) (formula:  $C_{13}H_{22}N$ ) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
4	Δ	1	Total	С	Η	Ν	0	0	
4	Л	T	36	13	22	1	0	0	
4	В	1	Total	С	Η	Ν	0	0	
4	D	1	36	13	22	1	0		
4	Т	1	Total	С	Η	Ν	0	0	
4			36	13	22	1	0	U	

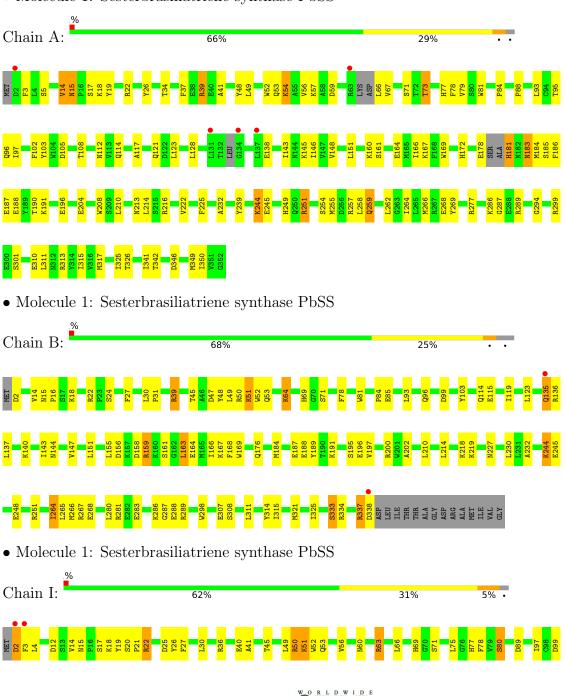
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	64	$\begin{array}{cc} \text{Total} & \text{O} \\ 64 & 64 \end{array}$	0	0
5	В	69	Total O 69 69	0	0
5	Ι	32	$\begin{array}{cc} \text{Total} & \text{O} \\ 32 & 32 \end{array}$	0	0



# 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: Sesterbrasiliatriene synthase PbSS

#### A126 L127 L128 S129 V111 N112 R159 K160 S161 G162 T120 Q121 M184 S185 ASN 1146 V147 Q11, N243 K244 E245 F246 D247 E248 E248 S254 M255 M255 D256 R257 L258 L258 L262 C263 C263 Y 240 3241 7242 E268 Y 269 L280 R281 325 A27 R27 M29 127 L340 ILE T342 T343 A344 R347 A348 M349 ILE VAL GLY S333 R334 Y335 <mark>H336</mark> R337 Y314 I315 V316 M317 I318 I318 I319 E307 S308 H309 1325 T326



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Depositor
Resolution (Å)	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor EDS
% Data completeness	99.6(44.32 - 2.75)	Depositor
(in resolution range)	99.6(44.32 - 2.75)	EDS
$R_{merge}$	0.14	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$11.81 (at 2.77 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
$R, R_{free}$	0.187 , $0.268$	Depositor
It, It <sub>free</sub>	0.187 , $0.268$	DCC
$R_{free}$ test set	1994 reflections $(7.21\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	27.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.31, $36.1$	EDS
L-test for twinning <sup>2</sup>	$<  L  > = 0.52, < L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	8621	wwPDB-VP
Average B, all atoms $(Å^2)$	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 2.08% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: MG, PPV, BTM

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		
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.75	0/2853	0.72	0/3850	
1	В	0.54	0/2794	0.65	0/3774	
1	Ι	0.63	0/2855	0.69	0/3853	
All	All	0.65	0/8502	0.69	0/11477	

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	А	2791	0	2681	91	0
1	В	2729	0	2619	76	0
1	Ι	2792	0	2679	91	0
2	А	3	0	0	0	0
2	В	3	0	0	0	0
2	Ι	3	0	0	0	0
3	А	9	0	0	0	0
3	В	9	0	0	0	0
3	Ι	9	0	0	0	0
4	А	14	22	22	7	0

Continued on next page...



Mol	Chain	Non-H	H(model) H(added)		Clashes	Symm-Clashes
4	В	14	22	22	12	0
4	Ι	14	22	22	7	0
5	А	64	0	0	1	0
5	В	69	0	0	3	0
5	Ι	32	0	0	1	0
All	All	8555	66	8045	271	0

Continued from previous page...

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 16.

The worst 5 of 271 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:66:LEU:HD13	1:A:67:VAL:H	1.05	1.08
1:A:117:ALA:O	1:A:121:GLN:HG3	1.56	1.05
1:A:66:LEU:HD13	1:A:67:VAL:N	1.70	1.04
1:I:272:THR:OG1	1:I:275:GLU:HG3	1.58	1.03
1:I:244:LYS:HE3	1:I:245:GLU:HG3	1.39	1.01

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	Favoured Allowed		Perce	Percentiles	
1	А	338/352~(96%)	314 (93%)	24 (7%)	0	100	100	
1	В	335/352~(95%)	317~(95%)	18 (5%)	0	100	100	
1	Ι	340/352~(97%)	316 (93%)	24 (7%)	0	100	100	
All	All	1013/1056~(96%)	947 (94%)	66~(6%)	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	А	297/302~(98%)	276~(93%)	21 (7%)	14	26	
1	В	291/302~(96%)	268~(92%)	23~(8%)	12	22	
1	Ι	297/302~(98%)	258~(87%)	39 (13%)	4	6	
All	All	885/906~(98%)	802 (91%)	83 (9%)	8	15	

 $5~{\rm of}~83$  residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
1	Ι	147	VAL
1	Ι	257	ARG
1	Ι	157	LYS
1	Ι	218	LYS
1	Ι	280	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	В	53	GLN
1	В	96	GLN
1	В	135	GLN
1	Ι	309	HIS
1	Ι	336	HIS

#### 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)

Of 15 ligands modelled in this entry, 9 are monoatomic - leaving 6 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	Turne	Chain	Res	es Link	Bo	ond leng	$\mathbf{ths}$	В	ond ang	les
MIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
3	PPV	А	404	2	6,8,8	0.71	0	$13,\!13,\!13$	0.89	1 (7%)
4	BTM	А	405	-	14,14,14	0.91	1 (7%)	18,18,18	0.55	0
4	BTM	Ι	405	-	14,14,14	0.81	0	18,18,18	0.56	0
4	BTM	В	405	-	14,14,14	0.92	1 (7%)	18,18,18	0.90	1 (5%)
3	PPV	Ι	404	2	6,8,8	0.77	0	$13,\!13,\!13$	1.57	1 (7%)
3	PPV	В	404	2	6,8,8	0.71	0	$13,\!13,\!13$	1.24	1 (7%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PPV	А	404	2	-	0/6/6/6	-
4	BTM	А	405	-	-	8/14/14/14	0/1/1/1
4	BTM	Ι	405	-	-	6/14/14/14	0/1/1/1
4	BTM	В	405	-	-	8/14/14/14	0/1/1/1
3	PPV	Ι	404	2	-	0/6/6/6	-
3	PPV	В	404	2	-	3/6/6/6	-

All (2) bond length outliers are listed below:

Mol	Chain	$\operatorname{Res}$	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\operatorname{\AA})$	$\mathrm{Ideal}(\mathrm{\AA})$
4	В	405	BTM	C1-N	-2.52	1.47	1.53

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	А	405	BTM	C1-N	-2.45	1.47	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	Ι	404	PPV	P2-OPP-P1	-5.11	115.29	132.83
3	В	404	PPV	P2-OPP-P1	-3.35	121.33	132.83
3	А	404	PPV	P2-OPP-P1	-2.54	124.10	132.83
4	В	405	BTM	C2-N-C1	-2.51	101.85	109.16

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	$\operatorname{Res}$	Type	Atoms
3	В	404	PPV	P1-OPP-P2-O32
4	А	405	BTM	C8-C1-N-C2
4	В	405	BTM	C8-C1-N-C4
4	В	405	BTM	C8-C1-N-C6
4	Ι	405	BTM	C8-C1-N-C2

There are no ring outliers.

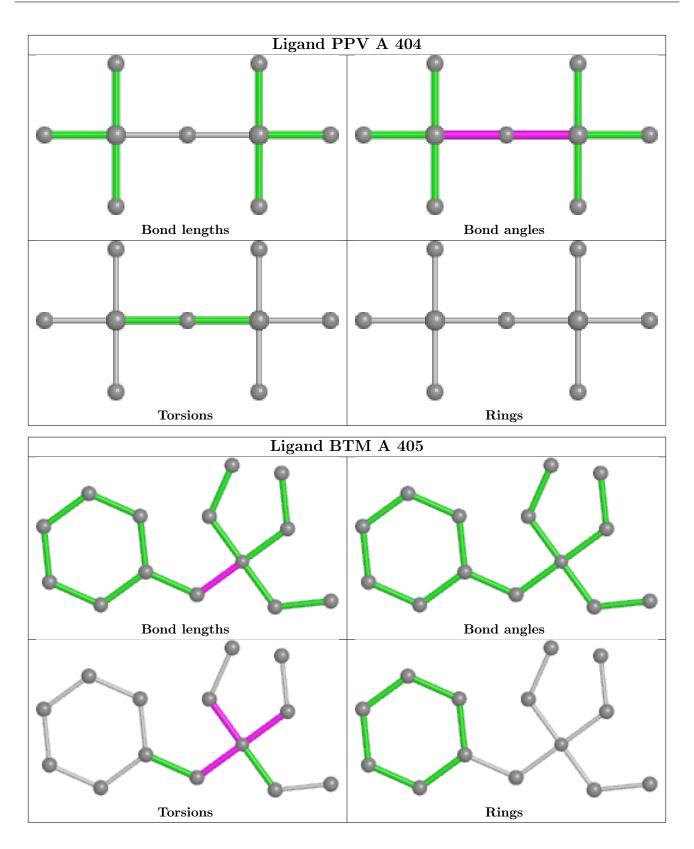
3 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	А	405	BTM	7	0
4	Ι	405	BTM	7	0
4	В	405	BTM	12	0

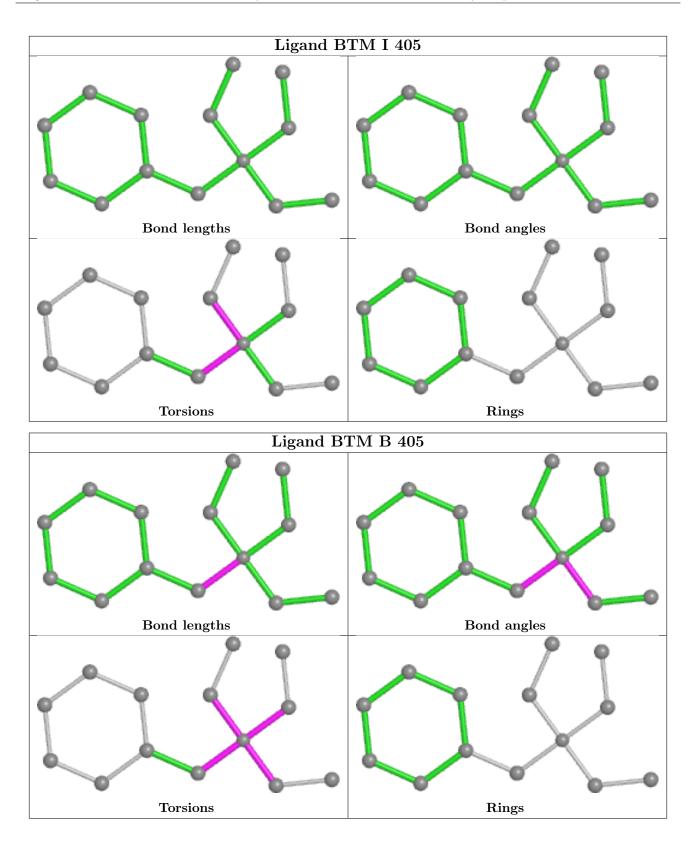
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



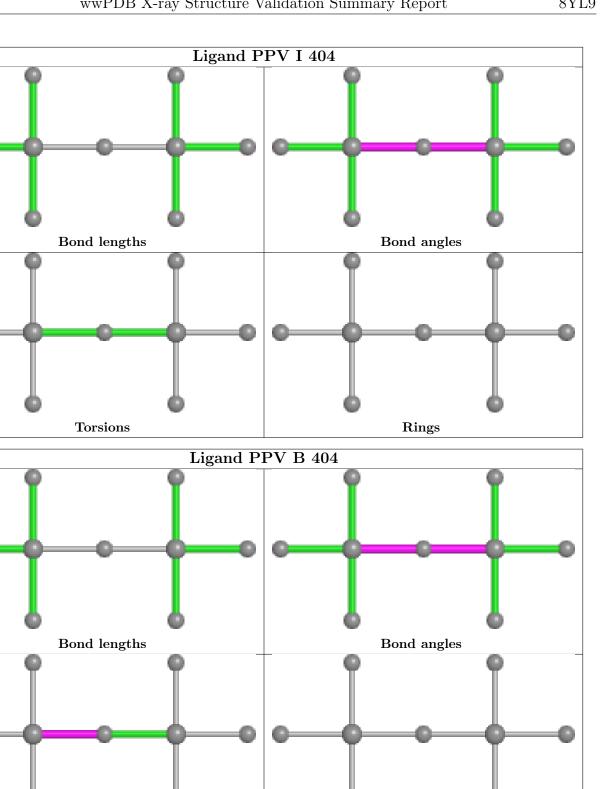












#### Other polymers (i) 5.7

There are no such residues in this entry.

Torsions





Rings

## 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,  $95^{th}$  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	А	346/352~(98%)	-0.30	5 (1%) 75 82	7, 20, 46, 73	0
1	В	337/352~(95%)	-0.30	2 (0%) 89 92	10, 19, 40, 69	0
1	Ι	346/352~(98%)	-0.09	5 (1%) 75 82	13, 29, 59, 79	0
All	All	1029/1056~(97%)	-0.23	12 (1%) 79 85	7, 23, 51, 79	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	134	GLY	4.3
1	Ι	2	ASP	3.9
1	В	135	GLN	3.0
1	Ι	3	PHE	2.9
1	В	338	ASP	2.8

### 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.

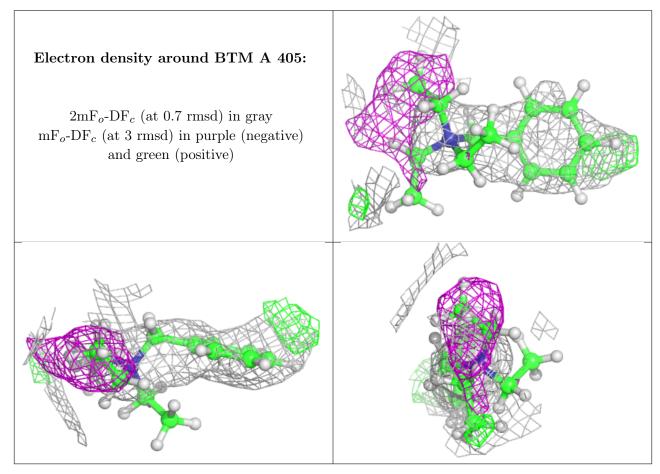
### 6.4 Ligands (i)

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,  $95^{th}$  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.

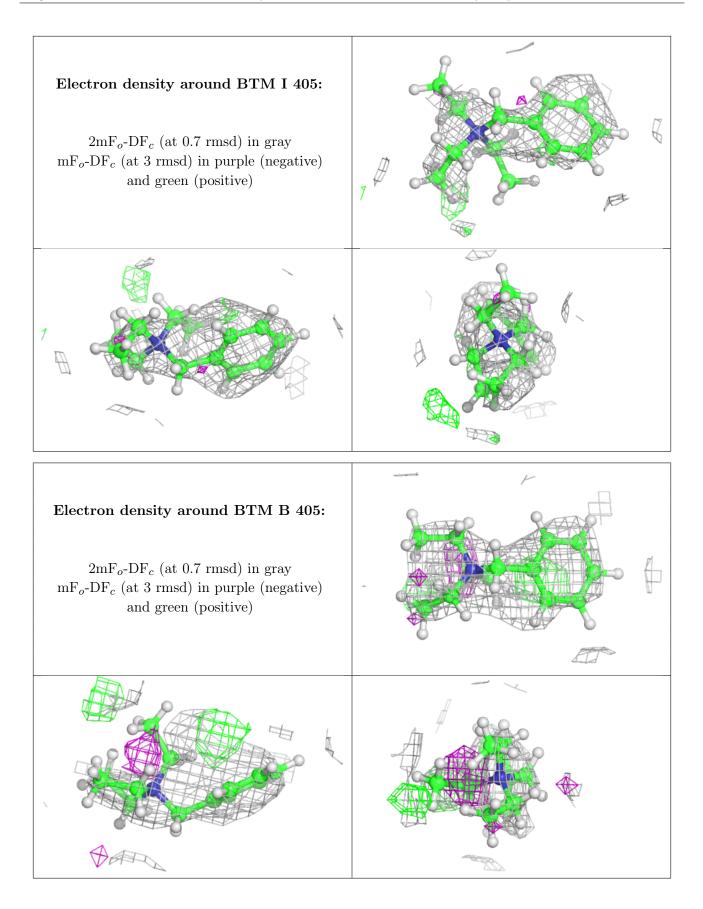


Mol	Type	Chain	Res	Atoms	RSCC	$\mathbf{RSR}$	$\operatorname{B-factors}(\operatorname{\AA}^2)$	Q < 0.9
4	BTM	А	405	14/14	0.79	0.48	32,41,54,64	0
4	BTM	Ι	405	14/14	0.81	0.37	27,39,46,48	0
4	BTM	В	405	14/14	0.83	0.41	20,26,35,40	0
2	MG	Ι	402	1/1	0.88	0.14	30,30,30,30	0
2	MG	В	401	1/1	0.89	0.08	16,16,16,16	0
2	MG	А	403	1/1	0.90	0.08	19,19,19,19	0
2	MG	В	402	1/1	0.91	0.10	13,13,13,13	0
2	MG	А	402	1/1	0.93	0.19	18,18,18,18	0
2	MG	В	403	1/1	0.93	0.04	12,12,12,12	0
2	MG	Ι	401	1/1	0.94	0.17	$22,\!22,\!22,\!22$	0
2	MG	Ι	403	1/1	0.95	0.05	29,29,29,29	0
2	MG	А	401	1/1	0.95	0.20	18,18,18,18	0
3	PPV	Ι	404	9/9	0.97	0.10	19,23,27,29	0
3	PPV	А	404	9/9	0.97	0.12	12,17,21,23	0
3	PPV	В	404	9/9	0.98	0.11	$9,\!15,\!17,\!19$	0

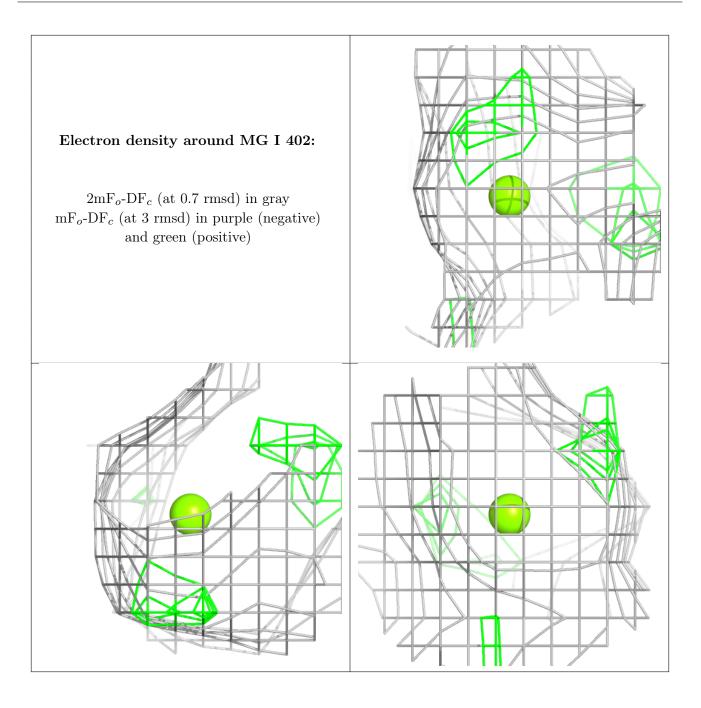
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



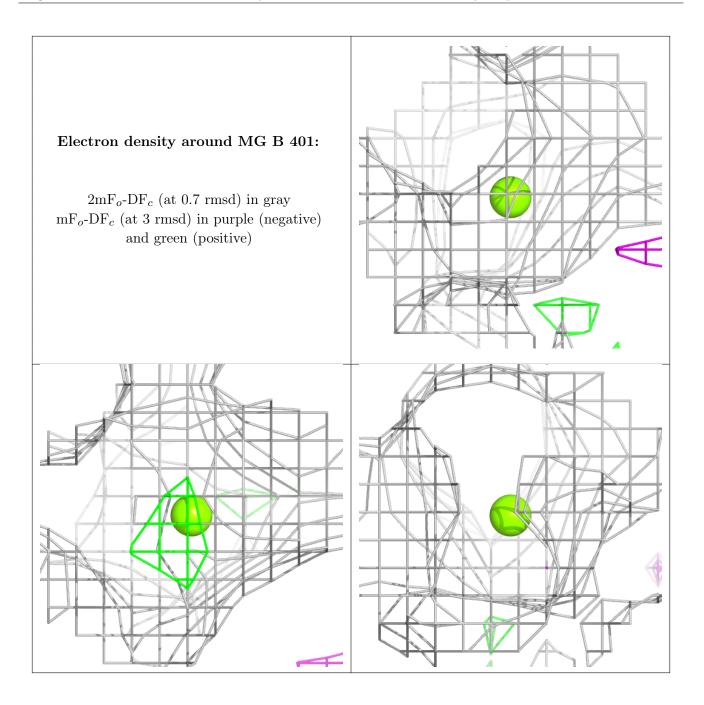




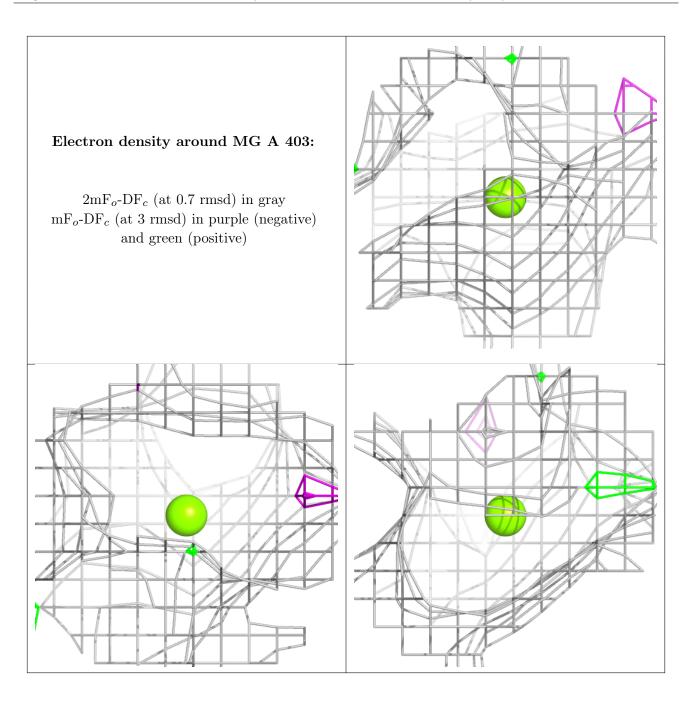




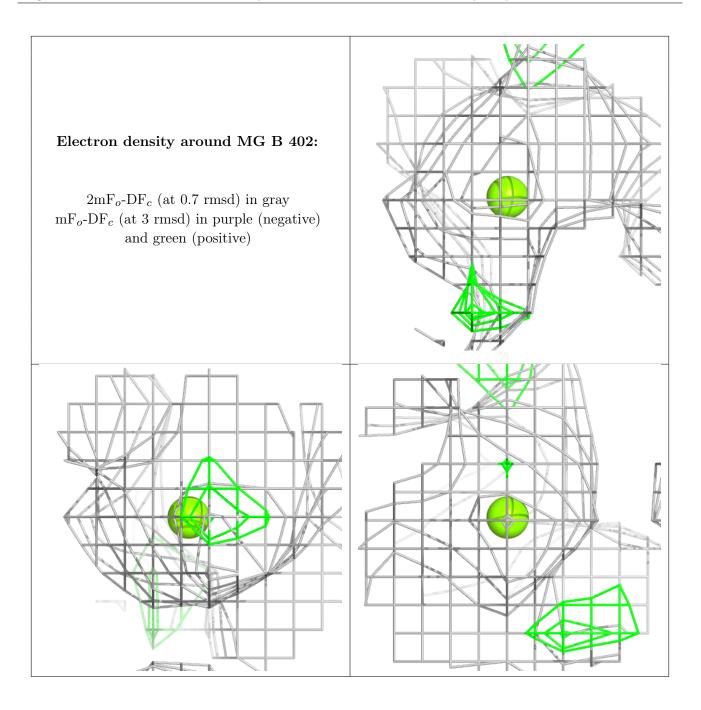




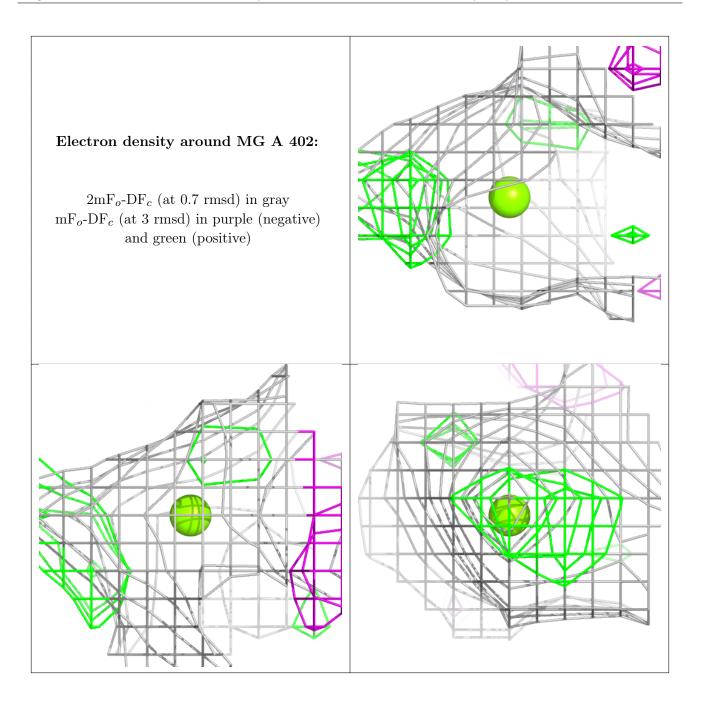




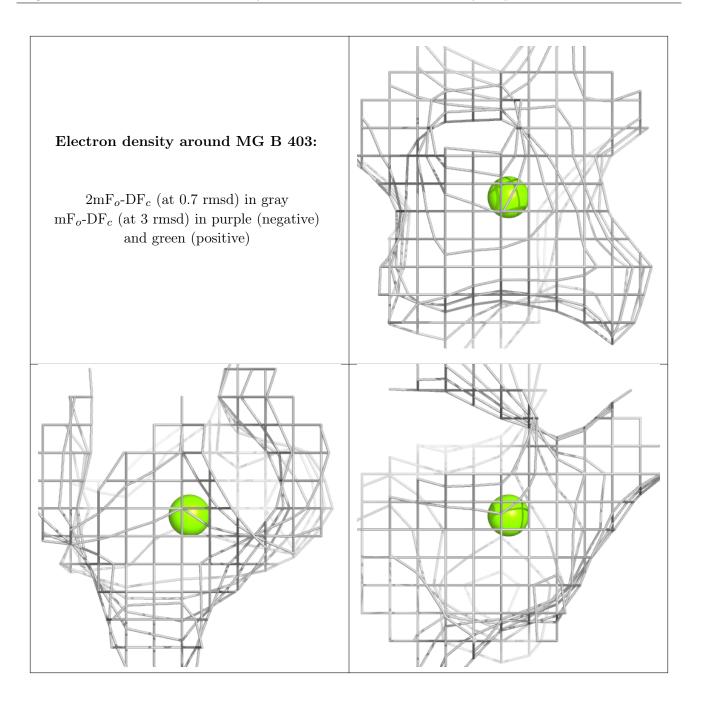




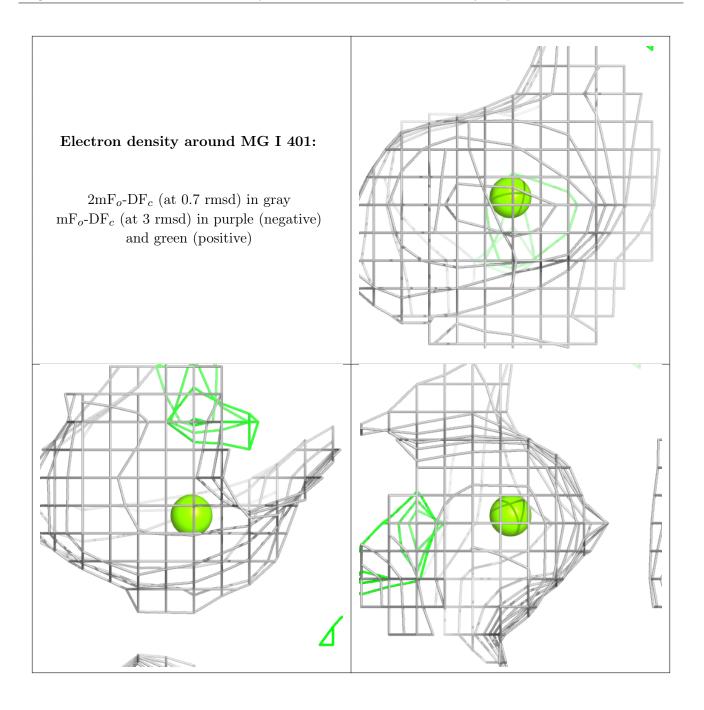




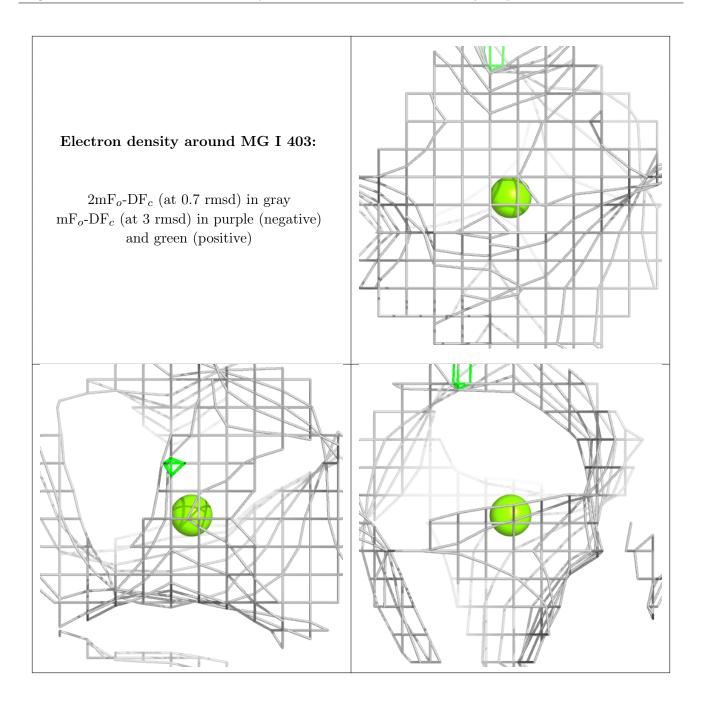




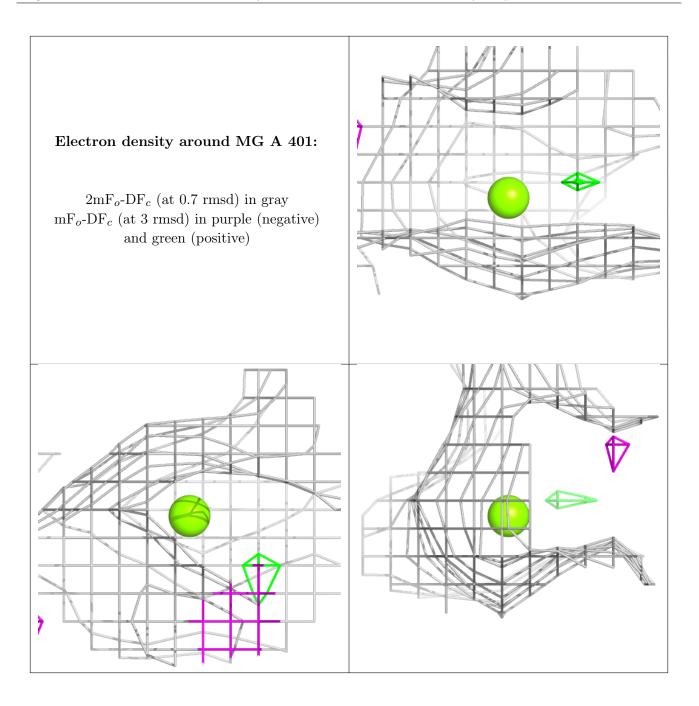




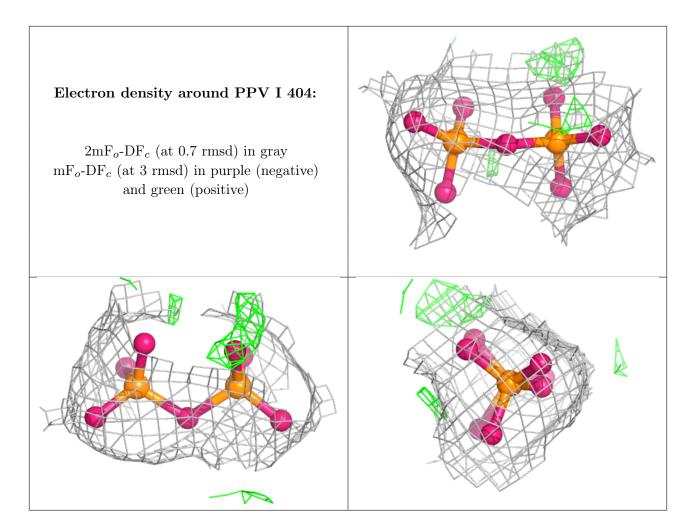




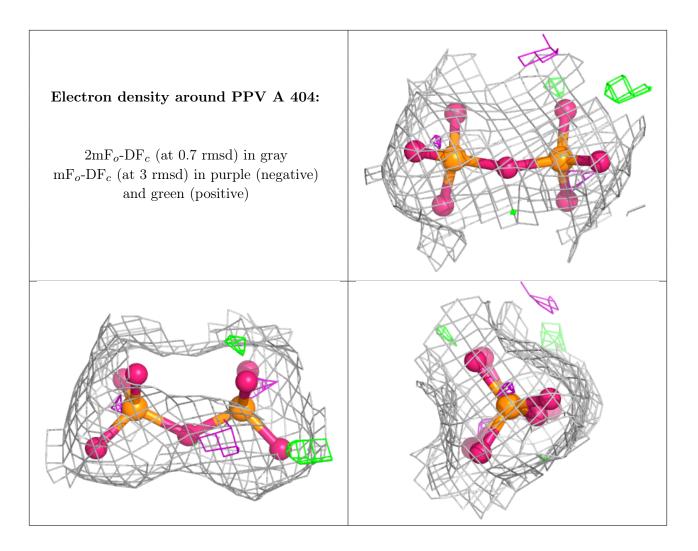




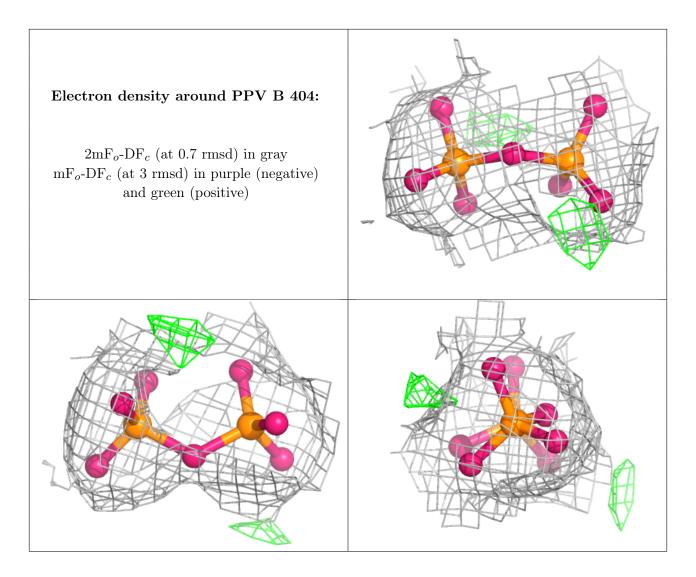












## 6.5 Other polymers (i)

There are no such residues in this entry.

