WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 17, 2022 – 02:04 pm BST

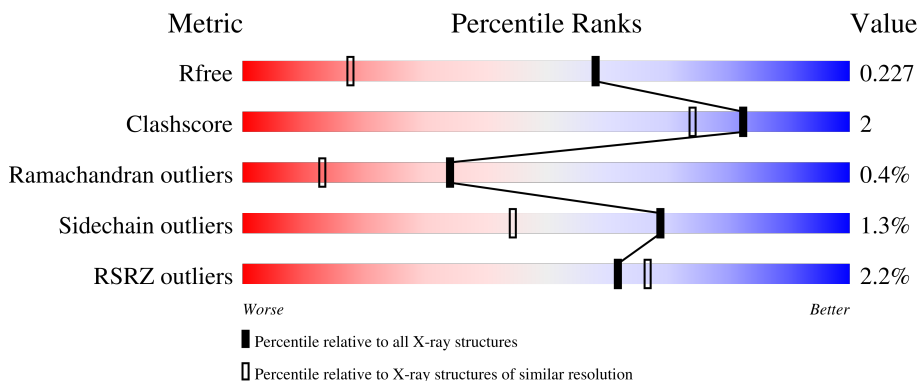|		
PDB ID	:	7O87
Title	:	Crystal structure of holo-F210W mutant of Hydroxy ketone aldolase (SwHKA) from Sphingomonas wittichii RW1 in complex with hydroxypyruvate
Authors	:	Laustsen, J.; Justo, I.; Marsden, S.R.; Hanefeld, U.; Bento, I.
Deposited on	:	2021-04-14
Resolution	:	1.50 Å(reported)

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 1.50 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	Percentile Ranks	Value
Rfree		0.227
Clashscore		2
Ramachandran outliers		0.4%
Sidechain outliers		1.3%
RSRZ outliers		2.2%

Worse — *Better*

▋ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	AAA	252	2% / 96%
1	BBB	252	2% / 92% / 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
2	3PY	BBB	301	-	-	X	-

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 8022 atoms, of which 3836 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 HpcH/HpaI aldolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	252	Total 3841	C 1208	H 1919	N 345	O 360	S 9	45	5	0
1	BBB	250	Total 3801	C 1201	H 1896	N 338	O 358	S 8	45	6	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	HIS	-	expression tag	UNP A5VH82
AAA	1	HIS	-	expression tag	UNP A5VH82
AAA	210	TRP	PHE	engineered mutation	UNP A5VH82
BBB	0	HIS	-	expression tag	UNP A5VH82
BBB	1	HIS	-	expression tag	UNP A5VH82
BBB	210	TRP	PHE	engineered mutation	UNP A5VH82

- Molecule 2 is 3-HYDROXYPYRUVIC ACID (three-letter code: 3PY) (formula: $C_3H_4O_4$) (labeled as "Ligand of Interest" by depositor).
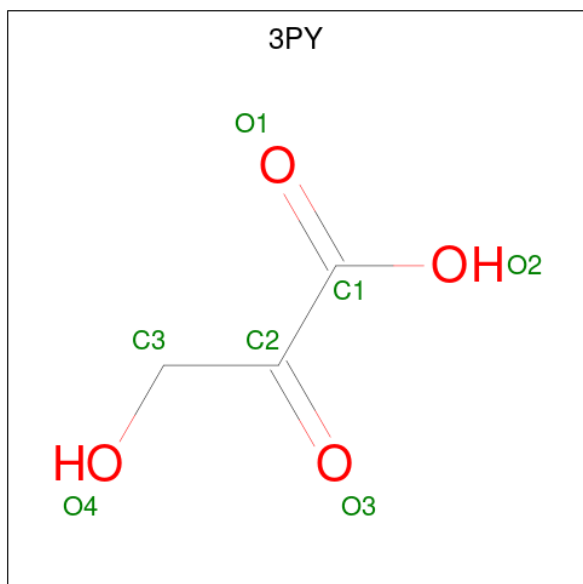
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total 10	C 3	H 3	O 4	1	0
2	BBB	1	Total 10	C 3	H 3	O 4	1	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).
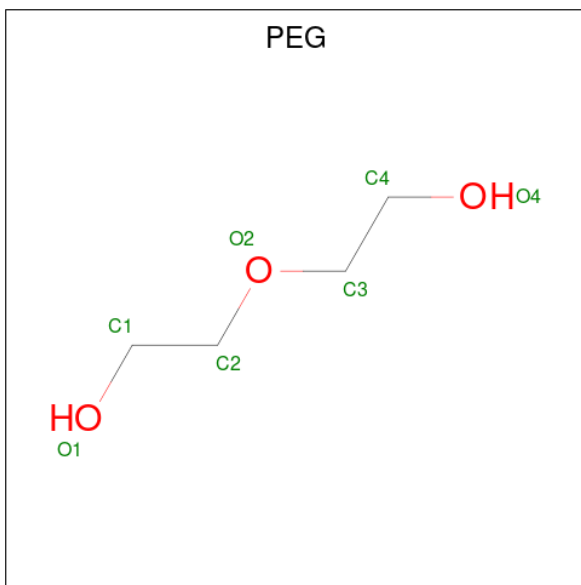
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total 9	C 2	H 5	O 2	1	0
3	BBB	1	Total 17	C 4	H 10	O 3	1	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	2	Total 2	Mg 2	0	0
4	BBB	1	Total 1	Mg 1	0	0

- Molecule 5 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	1	Total 1	K 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	BBB	2	Total 2	K 2	0	0

- Molecule 6 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	1	Total 1	Br 1	0	0
6	BBB	1	Total 1	Br 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	164	Total 165	O 165	0	2
7	BBB	157	Total 161	O 161	0	4

3 Residue-property plots ⓘ

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: HpcH/HpaI aldolase

Chain AAA:

• Molecule 1: HpcH/HpaI aldolase

Chain BBB:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	70.81Å 70.81Å 221.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	53.73 – 1.50 53.67 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.2 (53.73-1.50) 94.7 (53.67-1.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.184 , 0.225 0.184 , 0.227	Depositor DCC
R_{free} test set	3159 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	17.2	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.063 for -h-k,k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8022	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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: BR, MG, PEG, K, 3PY

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	AAA	0.77	0/1963	0.85	0/2668
1	BBB	0.74	0/1953	0.82	0/2658
All	All	0.75	0/3916	0.84	0/5326

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	AAA	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	161[A]	GLY	Peptide

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	AAA	1922	1919	1909	4	0
1	BBB	1905	1896	1892	13	0
2	AAA	7	3	3	0	0
2	BBB	7	3	3	6	0
3	AAA	4	5	5	0	0
3	BBB	7	10	10	0	0
4	AAA	2	0	0	0	0
4	BBB	1	0	0	0	0
5	AAA	1	0	0	0	0
5	BBB	2	0	0	0	0
6	AAA	1	0	0	1	0
6	BBB	1	0	0	0	0
7	AAA	165	0	0	0	0
7	BBB	161	0	0	4	0
All	All	4186	3836	3822	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.

All (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:BBB:97:ARG:NH2	1:BBB:158:SER:OG	2.09	0.86
6:AAA:306:BR:BR	7:BBB:479:HOH:O	2.59	0.76
1:BBB:242:ARG:HD3	7:BBB:498[A]:HOH:O	1.92	0.69
1:BBB:210[B]:TRP:CD1	2:BBB:301:3PY:C3	2.78	0.66
1:BBB:210[B]:TRP:CD1	2:BBB:301:3PY:H32	2.31	0.66
1:BBB:210[B]:TRP:CD1	2:BBB:301:3PY:H31	2.41	0.56
1:BBB:242:ARG:NH1	7:BBB:498[B]:HOH:O	2.37	0.55
1:BBB:143:MET:HA	1:BBB:166:PHE:O	2.08	0.54
1:AAA:178:HIS:HE1	1:AAA:188:GLU:OE1	1.93	0.51
1:AAA:143:MET:HA	1:AAA:166:PHE:O	2.12	0.50
1:BBB:199:ARG:HA	1:BBB:199:ARG:NE	2.27	0.50
1:AAA:208:GLY:HA2	1:AAA:227:PHE:O	2.13	0.48
1:BBB:210[B]:TRP:CE3	1:BBB:211:CYS:O	2.68	0.47
1:BBB:168:GLY:HA2	2:BBB:301:3PY:H31	1.96	0.46
1:AAA:167:ILE:O	1:AAA:169:PRO:HD3	2.16	0.45
2:BBB:301:3PY:H32	7:BBB:433:HOH:O	2.17	0.44
1:BBB:210[B]:TRP:HD1	2:BBB:301:3PY:H31	1.82	0.42
1:BBB:178:HIS:HE1	1:BBB:188:GLU:OE1	2.02	0.42
1:BBB:53:ALA:HB3	1:BBB:54:PRO:HD3	2.03	0.41

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	AAA	255/252 (101%)	248 (97%)	6 (2%)	1 (0%)	34	13
1	BBB	254/252 (101%)	246 (97%)	7 (3%)	1 (0%)	34	13
All	All	509/504 (101%)	494 (97%)	13 (3%)	2 (0%)	34	13

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	118	GLY
1	BBB	118	GLY

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	AAA	193/189 (102%)	192 (100%)	1 (0%)	88	78
1	BBB	192/189 (102%)	187 (97%)	5 (3%)	46	16
All	All	385/378 (102%)	379 (98%)	6 (2%)	69	36

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	114	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	BBB	114	LEU
1	BBB	145	GLU
1	BBB	147	ARG
1	BBB	160[A]	ASP
1	BBB	160[B]	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	PEG	BBB	302	-	6,6,6	0.19	0	5,5,5	0.13	0
3	PEG	AAA	302	-	3,3,6	0.07	0	2,2,5	0.35	0
2	3PY	BBB	301	4	5,6,6	6.43	1 (20%)	6,7,7	2.22	4 (66%)
2	3PY	AAA	301	4	5,6,6	6.46	2 (40%)	6,7,7	1.63	1 (16%)

Full wwPDB X-ray Structure Validation Report

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	PEG	BBB	302	-	-	3/4/4/4	-
3	PEG	AAA	302	-	-	0/1/1/4	-
2	3PY	BBB	301	4	-	0/5/6/6	-
2	3PY	AAA	301	4	-	2/5/6/6	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BBB	301	3PY	C2-C1	-14.21	1.34	1.53
2	AAA	301	3PY	C2-C1	-14.20	1.34	1.53
2	AAA	301	3PY	O1-C1	2.09	1.28	1.22

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	301	3PY	O1-C1-C2	-3.43	117.14	121.72
2	AAA	301	3PY	O3-C2-C3	-3.42	115.22	120.50
2	BBB	301	3PY	O3-C2-C3	-2.55	116.56	120.50
2	BBB	301	3PY	O2-C1-C2	2.37	120.46	113.97
2	BBB	301	3PY	O3-C2-C1	2.27	122.72	119.43

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	301	3PY	C1-C2-C3-O4
2	AAA	301	3PY	O3-C2-C3-O4
3	BBB	302	PEG	O2-C3-C4-O4
3	BBB	302	PEG	C1-C2-O2-C3
3	BBB	302	PEG	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	BBB	301	3PY	6	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 3PY BBB 301

Bond lengths

Bond angles

Torsions

Rings

Ligand 3PY AAA 301

Bond lengths

Bond angles

Torsions

Rings
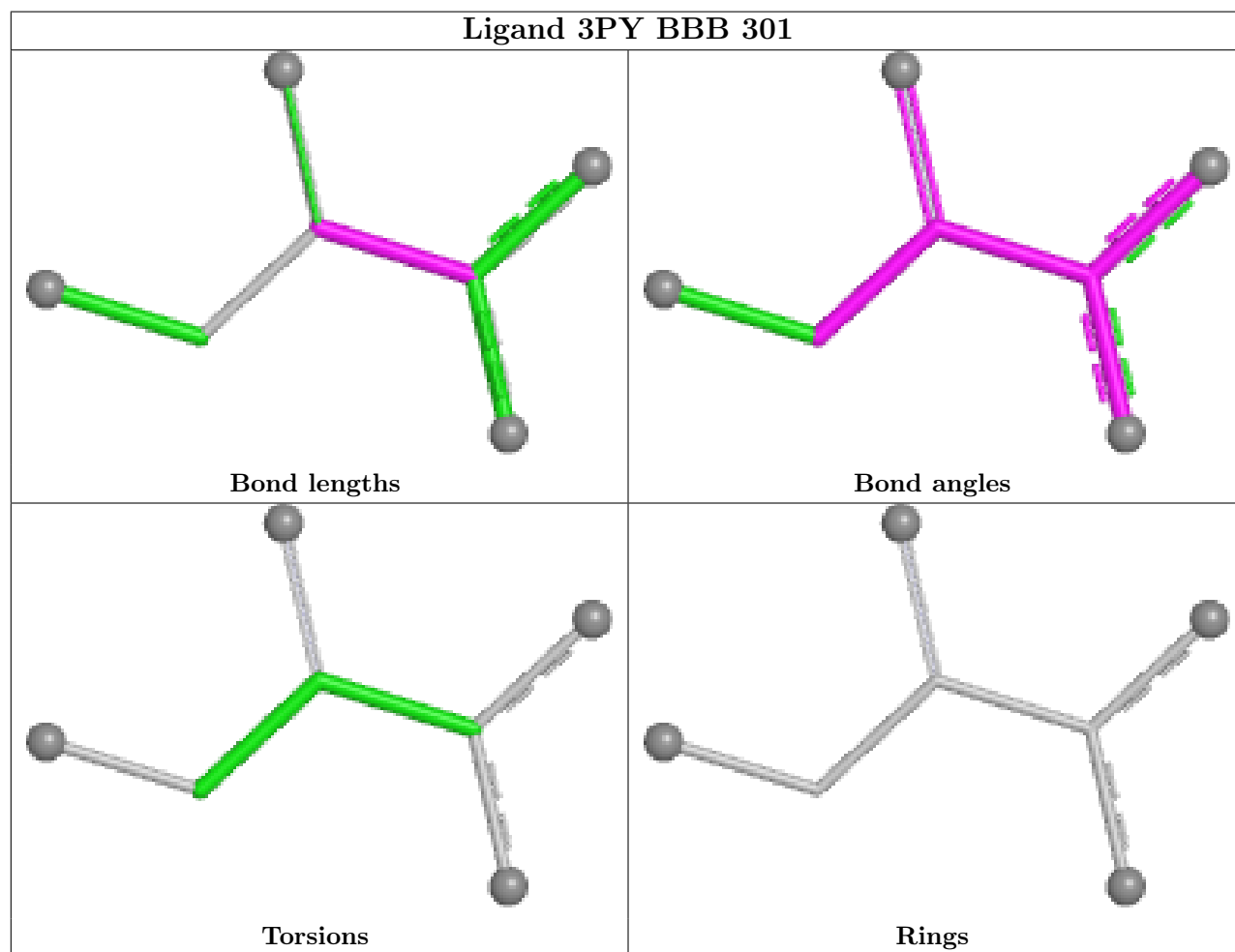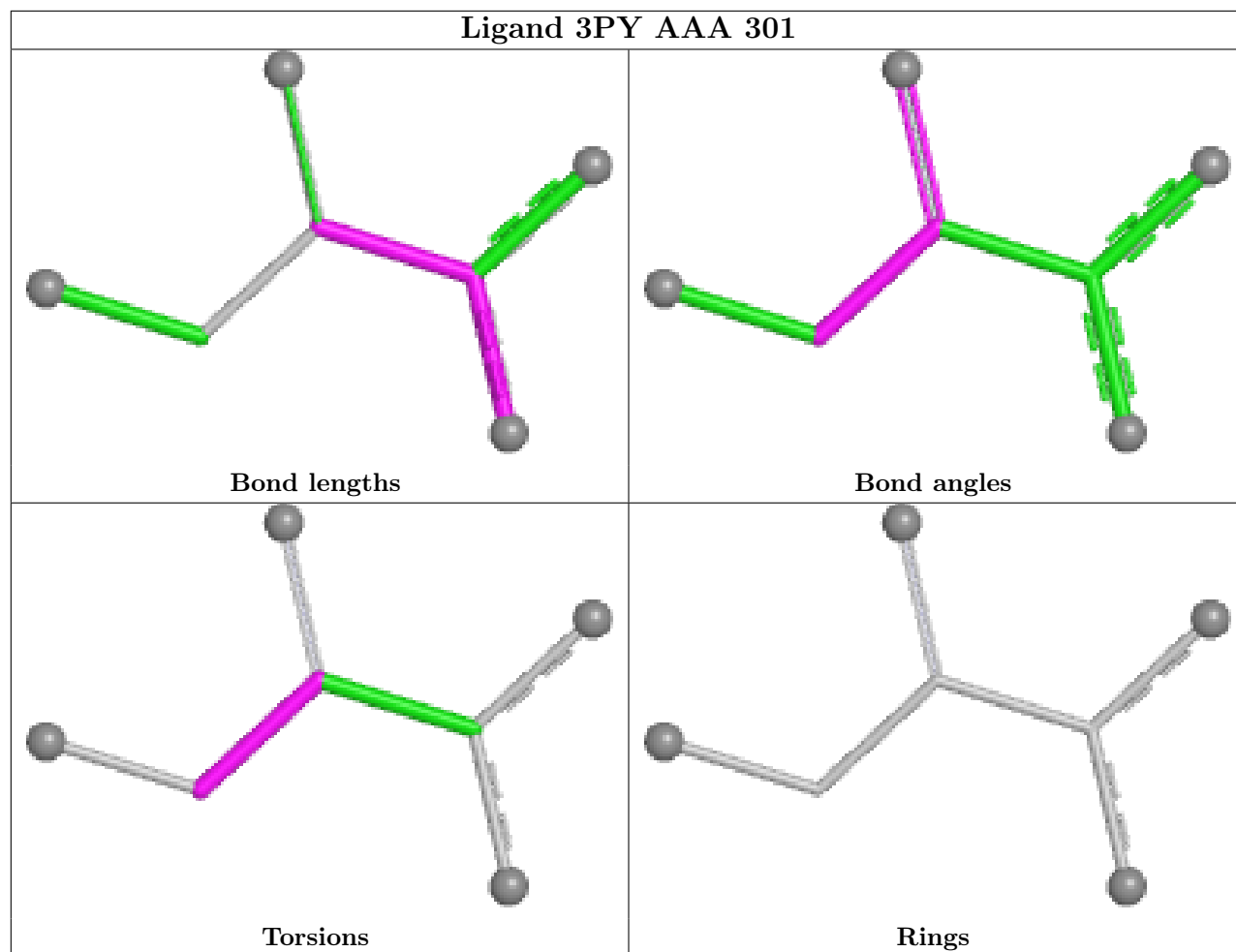
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	AAA	252/252 (100%)	0.13	5 (1%)	65	70	10, 15, 27, 46	0
1	BBB	250/252 (99%)	0.16	6 (2%)	59	63	11, 17, 29, 41	0
All	All	502/504 (99%)	0.15	11 (2%)	62	67	10, 16, 28, 46	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	210[A]	TRP	4.7
1	BBB	11	GLY	3.7
1	AAA	0	HIS	3.5
1	BBB	1	HIS	3.5
1	AAA	128	SER	3.1
1	BBB	160[A]	ASP	2.7
1	AAA	129	GLY	2.4
1	AAA	1	HIS	2.4
1	BBB	12	ARG	2.2
1	AAA	11	GLY	2.0
1	BBB	7	CYS	2.0

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.4 Ligands ⓘ

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
3	PEG	BBB	302	7/7	0.80	0.17	25,29,34,34	17
3	PEG	AAA	302	4/7	0.89	0.17	17,24,25,34	1
2	3PY	BBB	301	7/7	0.93	0.14	15,20,27,35	1
2	3PY	AAA	301	7/7	0.94	0.14	16,23,30,37	1
4	MG	AAA	303	1/1	0.99	0.05	14,14,14,14	0
4	MG	AAA	304	1/1	0.99	0.06	15,15,15,15	1
4	MG	BBB	303	1/1	0.99	0.06	15,15,15,15	0
5	K	AAA	305	1/1	0.99	0.06	14,14,14,14	0
5	K	BBB	304	1/1	0.99	0.06	27,27,27,27	1
5	K	BBB	305	1/1	0.99	0.07	16,16,16,16	0
6	BR	BBB	306	1/1	0.99	0.07	28,28,28,28	1
6	BR	AAA	306	1/1	1.00	0.07	29,29,29,29	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.
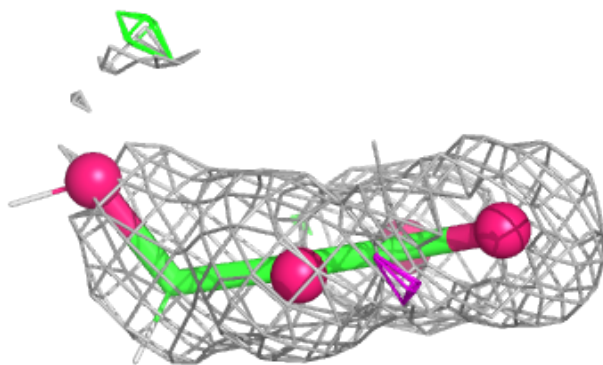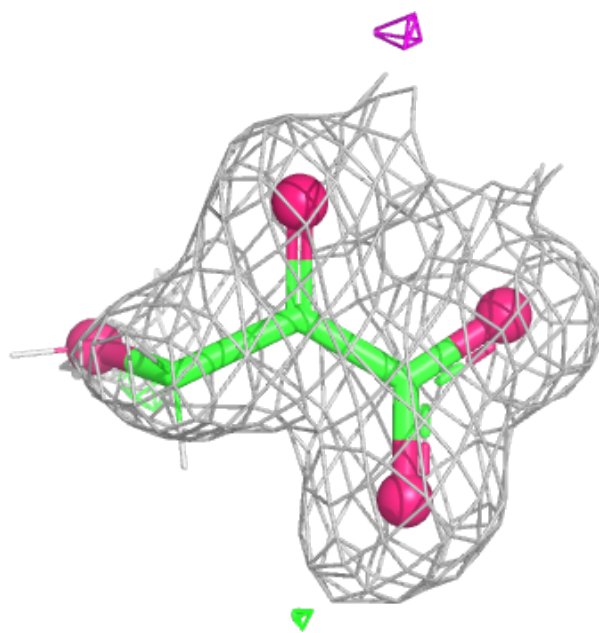
Electron density around 3PY BBB 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
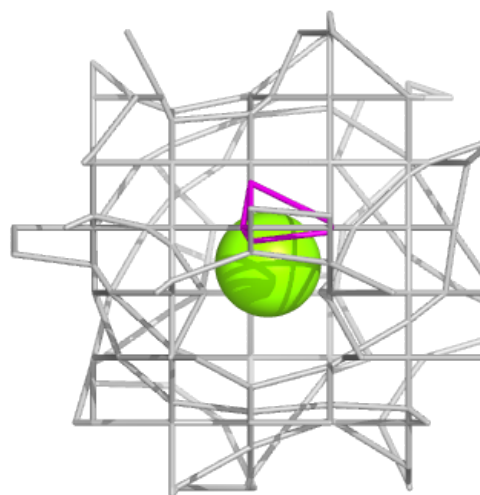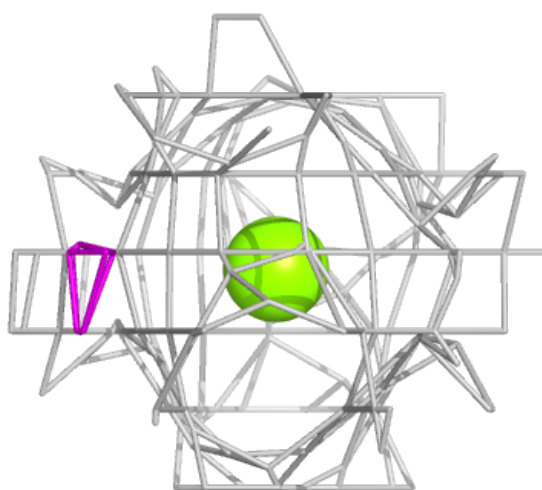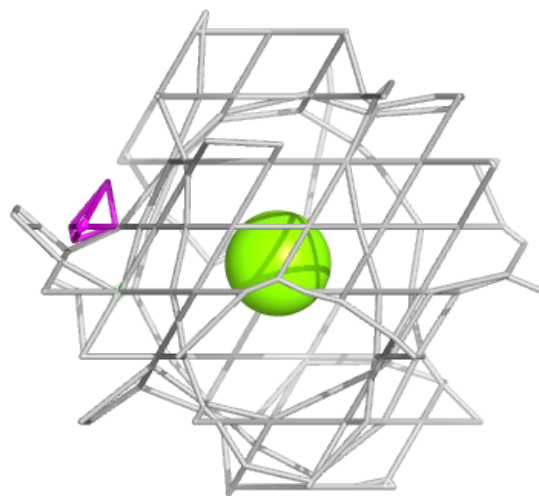mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
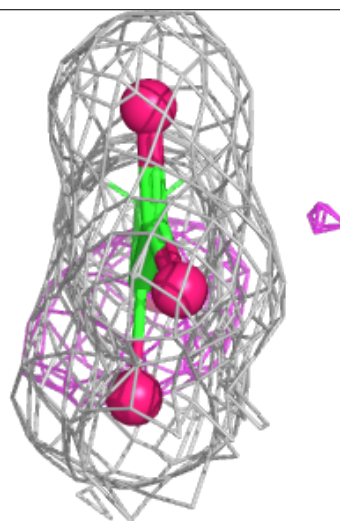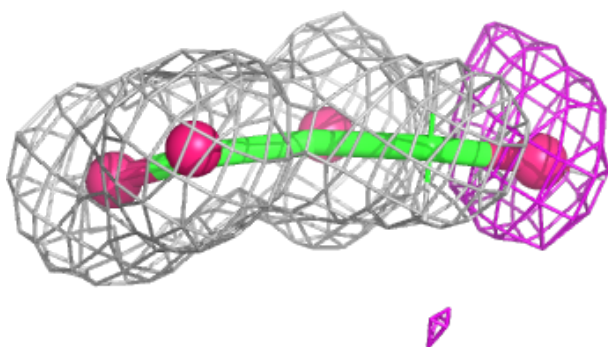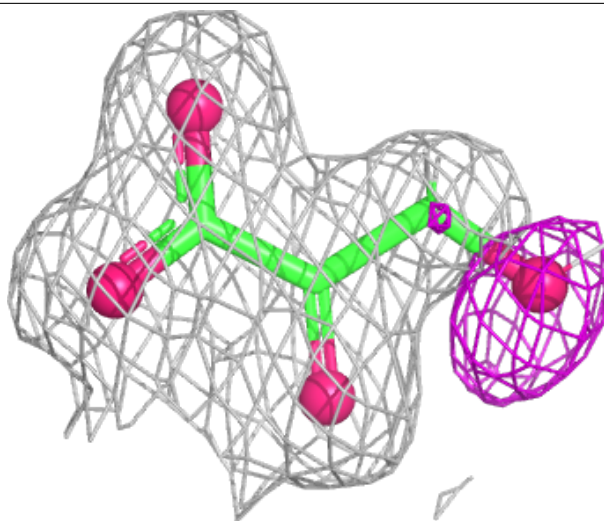
Electron density around 3PY AAA 301:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
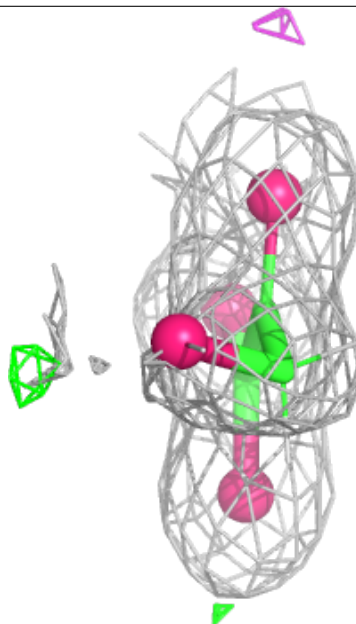
Electron density around MG AAA 303:

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

Electron density around MG AAA 304:

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

Electron density around MG BBB 303:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
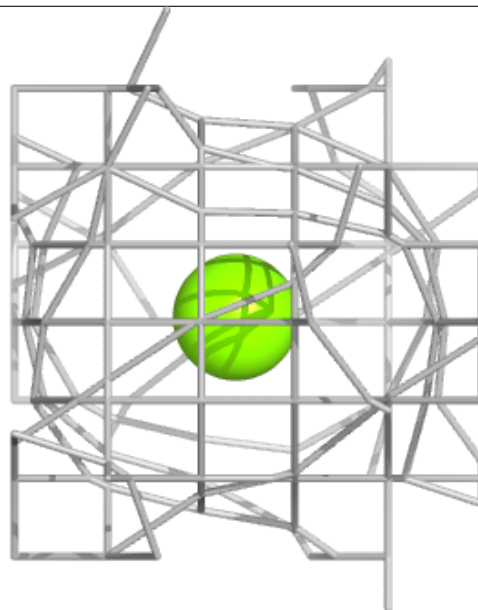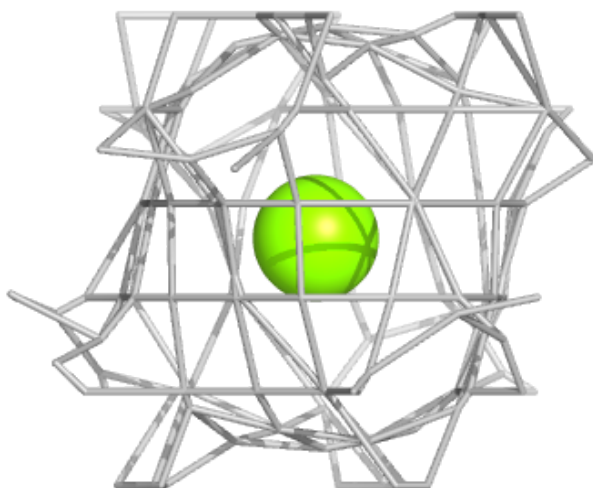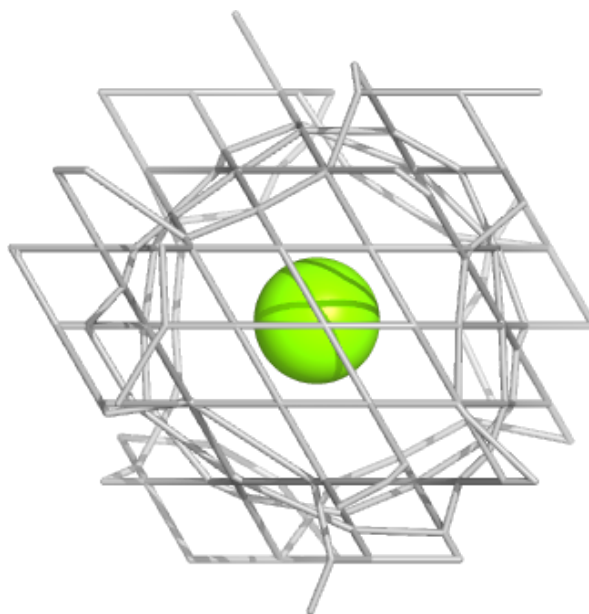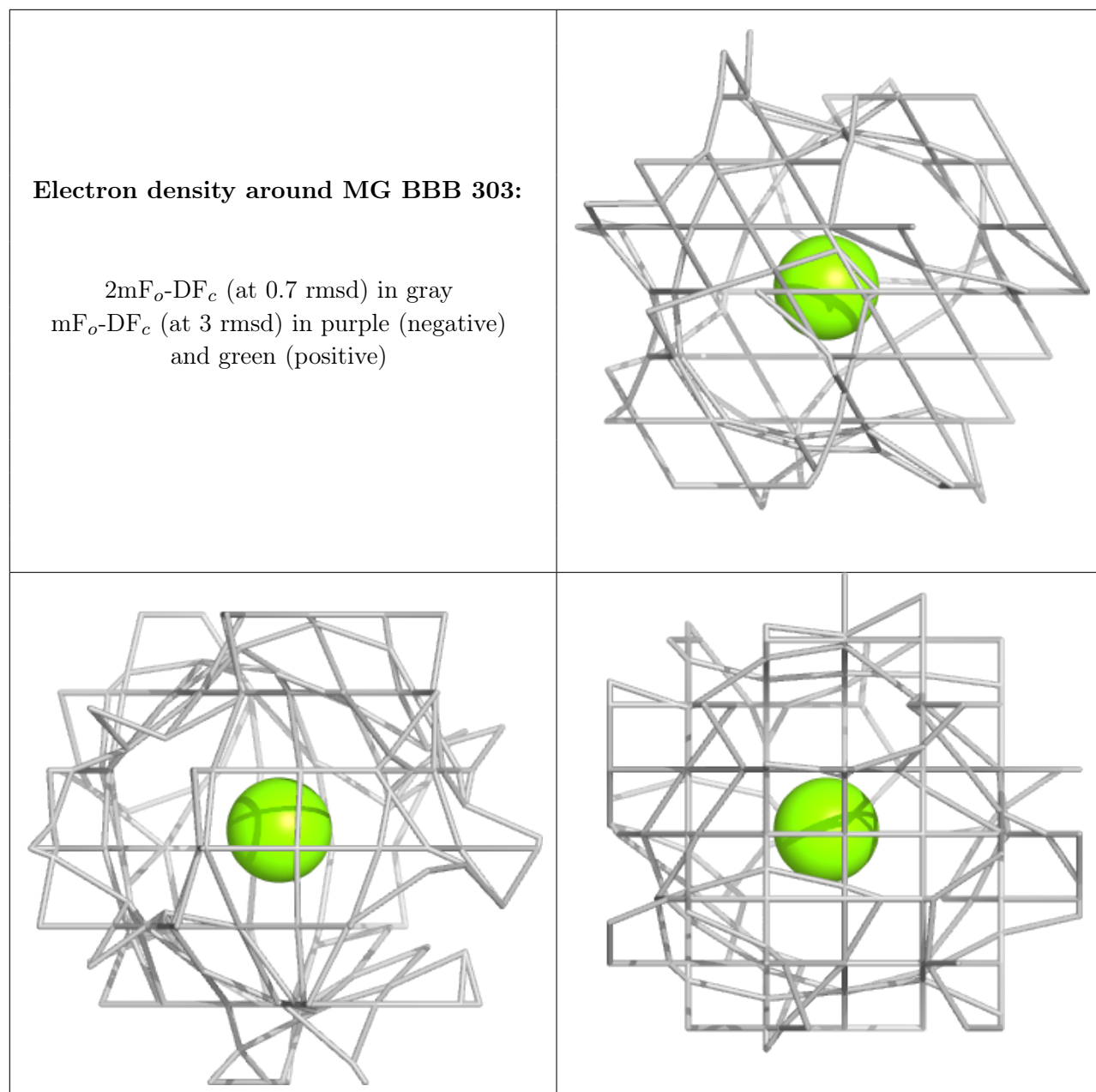
6.5 Other polymers (i)

There are no such residues in this entry.