WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2023 – 09:54 PM EDT

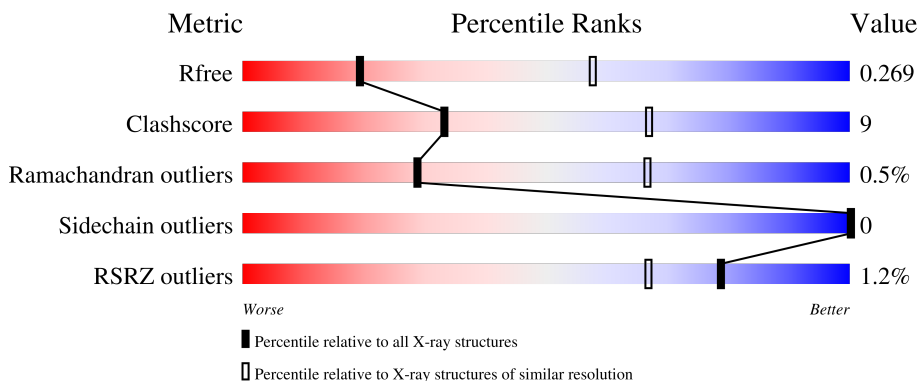|		
PDB ID	:	6PCP
Title	:	Mechanism for regulation of DNA binding of Bordetella bronchiseptica BpsR by 6-hydroxynicotinic acid
Authors	:	Booth, W.T.; Davis, R.R.; Deora, R.; Hollis, T.
Deposited on	:	2019-06-17
Resolution	:	3.20 Å(reported)

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 3.20 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		0.269
Clashscore		9
Ramachandran outliers		0.5%
Sidechain outliers		0
RSRZ outliers		1.2%

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	161	75% 12% 13%
1	B	161	71% 13% 16%
1	C	161	68% 18% · 13%
1	D	161	2% 68% 19% 14%
1	E	161	2% 79% 8% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	161	2% ▮ 75% 11% · 13%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6157 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.
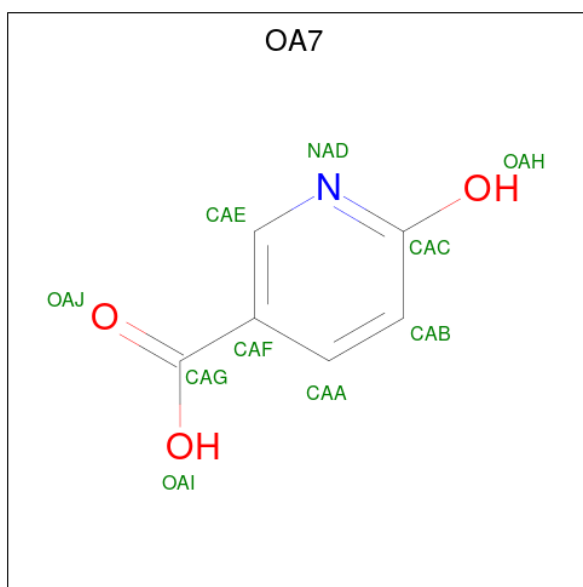
- Molecule 1 is a protein called MarR family transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	140	Total 1033	C 656	N 187	O 186	S 4	0	0	0
1	B	135	Total 1006	C 639	N 182	O 181	S 4	0	0	0
1	C	140	Total 1013	C 644	N 180	O 185	S 4	0	0	0
1	D	139	Total 1011	C 642	N 180	O 185	S 4	0	0	0
1	E	140	Total 1018	C 646	N 181	O 187	S 4	0	0	0
1	F	140	Total 1008	C 642	N 177	O 185	S 4	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	LEU	PRO	conflict	UNP A0A0T7MD48
A	160	ASP	GLU	conflict	UNP A0A0T7MD48
B	8	LEU	PRO	conflict	UNP A0A0T7MD48
B	160	ASP	GLU	conflict	UNP A0A0T7MD48
C	8	LEU	PRO	conflict	UNP A0A0T7MD48
C	160	ASP	GLU	conflict	UNP A0A0T7MD48
D	8	LEU	PRO	conflict	UNP A0A0T7MD48
D	160	ASP	GLU	conflict	UNP A0A0T7MD48
E	8	LEU	PRO	conflict	UNP A0A0T7MD48
E	160	ASP	GLU	conflict	UNP A0A0T7MD48
F	8	LEU	PRO	conflict	UNP A0A0T7MD48
F	160	ASP	GLU	conflict	UNP A0A0T7MD48

- Molecule 2 is 6-hydroxypyridine-3-carboxylic acid (three-letter code: OA7) (formula: $C_6H_5NO_3$) (labeled as "Ligand of Interest" by depositor).

OA7

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			10	6	1	3		
2	A	1	Total	C	N	O	0	0
			10	6	1	3		
2	B	1	Total	C	N	O	0	0
			10	6	1	3		
2	C	1	Total	C	N	O	0	0
			10	6	1	3		
2	E	1	Total	C	N	O	0	0
			10	6	1	3		
2	F	1	Total	C	N	O	0	0
			10	6	1	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	O	0	0
			4	4		
3	B	1	Total	O	0	0
			1	1		
3	C	1	Total	O	0	0
			1	1		
3	D	2	Total	O	0	0
			2	2		

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: MarR family transcriptional regulator

Chain A:
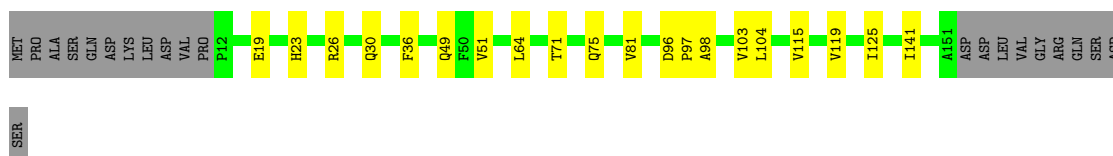
- Molecule 1: MarR family transcriptional regulator
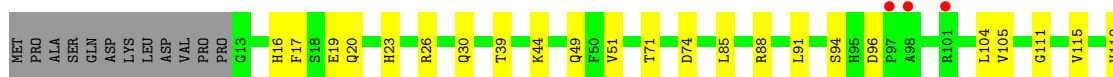
Chain B:

- Molecule 1: MarR family transcriptional regulator
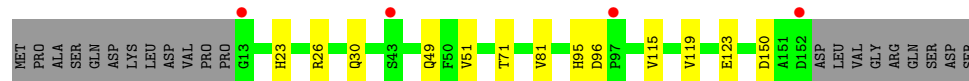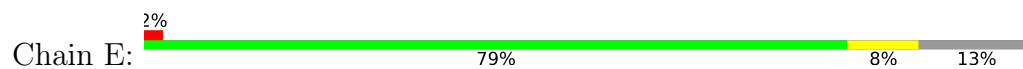
Chain C:

- Molecule 1: MarR family transcriptional regulator

Chain D:

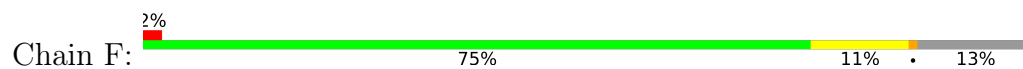I125 P135 R138 V139 A140 I141 L144 L145 R146 A151 ASP ASP LEU VAL GLY ARG GLN SER ASP SER

• Molecule 1: MarR family transcriptional regulator

Chain E:

2%

79%　　　8%　　13%

MET PRO ALA SER GLN ASP LYS LEU ASP VAL PRO PRO G13 H23 R26 Q30 S43 Q49 F50 V51 T71 V81 H95 D96 P97 V115 V119 E123 D150 A151 D152 ASP LEU VAL GLY ARG GLN SER ASP SER

• Molecule 1: MarR family transcriptional regulator

Chain F:

2%

75%　　11%　•　13%

MET PRO ALA SER GLN ASP LYS LEU ASP VAL PRO P12 F17 H23 R26 R27 Q30 S43 K44 L45 V51 V67 T71 D74 D96 D99 R100 R101 K102 V103 V115 V119 P135 R138 V139 A140 I141 A151 ASP ASP LEU VAL GLY ARG

GLN SER ASP SER

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	77.54Å 110.57Å 273.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.09 – 3.20 30.09 – 3.20	Depositor EDS
% Data completeness (in resolution range)	90.9 (30.09-3.20) 90.7 (30.09-3.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.47 (at 3.18Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.216 , 0.269 0.217 , 0.269	Depositor DCC
R_{free} test set	1816 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å^2)	45.5	Xtriage
Anisotropy	1.010	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 32.6	EDS
L-test for twinning[2]	$< |L| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	6157	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: OA7

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	0.25	0/1050	0.45	0/1432
1	B	0.24	0/1021	0.44	0/1389
1	C	0.25	0/1029	0.43	0/1408
1	D	0.25	0/1027	0.45	0/1403
1	E	0.25	0/1035	0.47	0/1416
1	F	0.24	0/1025	0.45	0/1402
All	All	0.25	0/6187	0.45	0/8450

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 ⓘ

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	1033	0	1028	24	0
1	B	1006	0	1006	15	0
1	C	1013	0	984	22	0
1	D	1011	0	989	26	0
1	E	1018	0	988	10	0
1	F	1008	0	980	14	0
2	A	20	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	10	0	0	1	0
2	C	10	0	0	2	0
2	E	10	0	0	1	0
2	F	10	0	0	1	0
3	A	4	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	2	0	0	0	0
All	All	6157	0	5975	104	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 9.

The worst 5 of 104 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:C:39:THR:HB	1:C:125:ILE:HD11	1.43	1.00
1:A:64:LEU:HD21	1:A:75:GLN:HG3	1.44	1.00
1:B:67:VAL:O	1:B:71:THR:HG22	1.61	0.98
1:C:115:VAL:O	1:C:119:VAL:HG23	1.63	0.97
1:C:121:PHE:O	1:C:125:ILE:HG12	1.71	0.91

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	138/161 (86%)	133 (96%)	5 (4%)	0	100	100
1	B	131/161 (81%)	128 (98%)	2 (2%)	1 (1%)	19	58
1	C	138/161 (86%)	133 (96%)	4 (3%)	1 (1%)	22	61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	137/161 (85%)	137 (100%)	0	0	100	100
1	E	138/161 (86%)	135 (98%)	3 (2%)	0	100	100
1	F	138/161 (86%)	134 (97%)	2 (1%)	2 (1%)	11	46
All	All	820/966 (85%)	800 (98%)	16 (2%)	4 (0%)	29	67

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	130	PHE
1	F	96	ASP
1	F	103	VAL
1	B	100	ARG

5.3.2   Protein sidechains ⓘ

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	103/135 (76%)	103 (100%)	0	100	100
1	B	102/135 (76%)	102 (100%)	0	100	100
1	C	98/135 (73%)	98 (100%)	0	100	100
1	D	99/135 (73%)	99 (100%)	0	100	100
1	E	100/135 (74%)	100 (100%)	0	100	100
1	F	98/135 (73%)	98 (100%)	0	100	100
All	All	600/810 (74%)	600 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	GLN
1	D	23	HIS

Continued on next page...

WORLDWIDE wwPDB PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	F	23	HIS

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   ⓘ

6 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	OA7	F	201	-	10,10,10	1.08	0	12,13,13	1.35	2 (16%)
2	OA7	A	201	-	10,10,10	1.08	0	12,13,13	1.46	3 (25%)
2	OA7	B	201	-	10,10,10	1.07	0	12,13,13	1.52	3 (25%)
2	OA7	C	201	-	10,10,10	1.09	0	12,13,13	1.41	3 (25%)
2	OA7	A	202	-	10,10,10	1.11	0	12,13,13	1.44	3 (25%)
2	OA7	E	201	-	10,10,10	1.03	0	12,13,13	1.61	3 (25%)

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
2	OA7	F	201	-	-	4/4/4/4	0/1/1/1
2	OA7	A	201	-	-	0/4/4/4	0/1/1/1
2	OA7	B	201	-	-	0/4/4/4	0/1/1/1
2	OA7	C	201	-	-	0/4/4/4	0/1/1/1
2	OA7	A	202	-	-	0/4/4/4	0/1/1/1
2	OA7	E	201	-	-	0/4/4/4	0/1/1/1

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	201	OA7	CAE-NAD-CAC	3.23	123.59	117.14
2	A	202	OA7	CAE-NAD-CAC	3.23	123.58	117.14
2	A	201	OA7	CAE-NAD-CAC	3.11	123.33	117.14
2	F	201	OA7	CAE-NAD-CAC	3.08	123.28	117.14
2	C	201	OA7	CAE-NAD-CAC	3.08	123.28	117.14

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	201	OA7	CAA-CAF-CAG-OAI
2	F	201	OA7	CAA-CAF-CAG-OAJ
2	F	201	OA7	CAE-CAF-CAG-OAI
2	F	201	OA7	CAE-CAF-CAG-OAJ

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	201	OA7	1	0
2	A	201	OA7	1	0
2	B	201	OA7	1	0
2	C	201	OA7	2	0
2	E	201	OA7	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be

highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand OA7 F 201

Bond lengths

Bond angles

Torsions

Rings

Ligand OA7 A 201

Bond lengths

Bond angles

Torsions

Rings

Ligand OA7 B 201

Bond lengths

Bond angles

Torsions

Rings

Ligand OA7 C 201

Bond lengths

Bond angles

Torsions

Rings

Ligand OA7 A 202

Bond lengths

Bond angles

Torsions

Rings

Ligand OA7 E 201

Bond lengths

Bond angles

Torsions

Rings
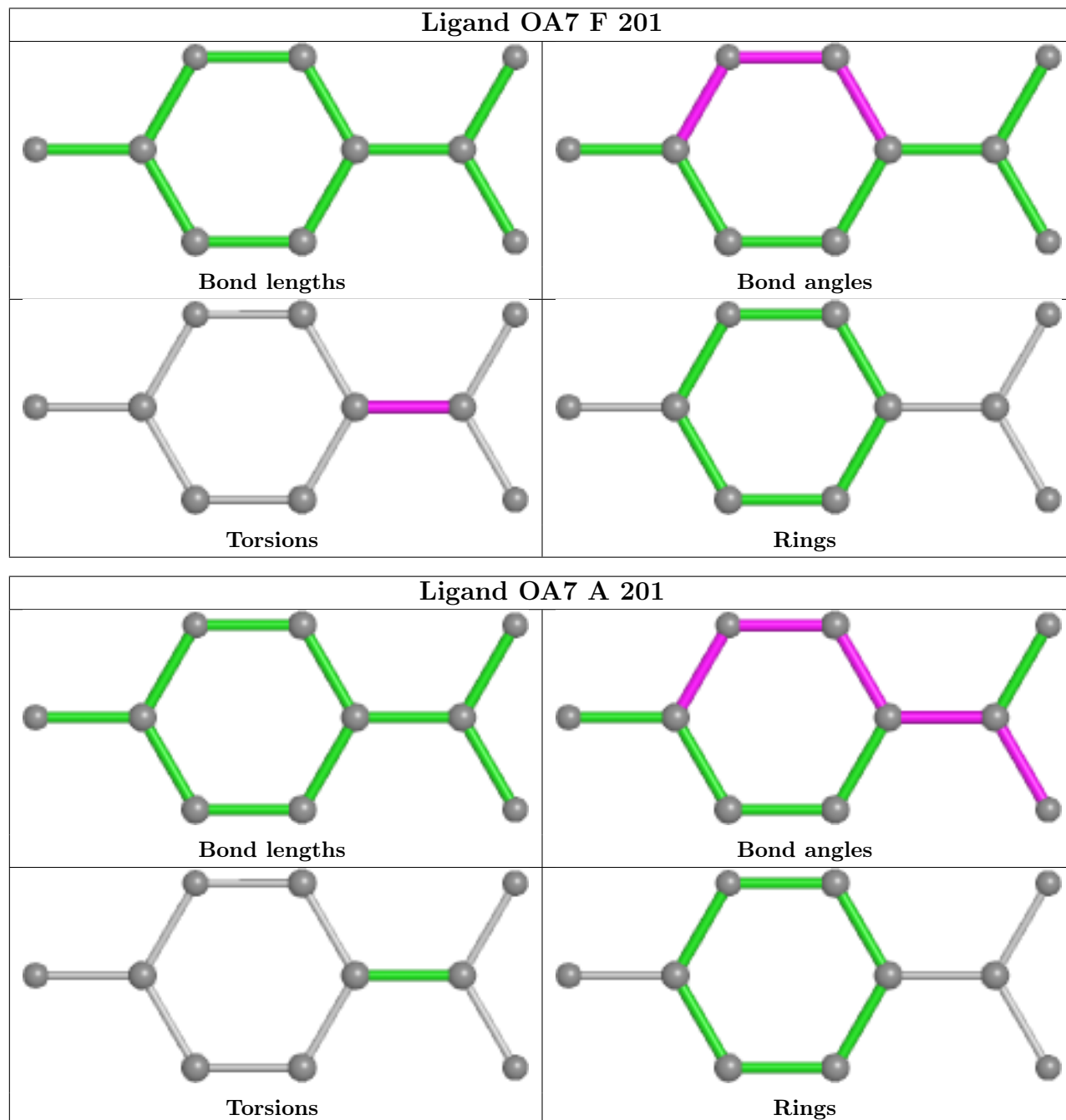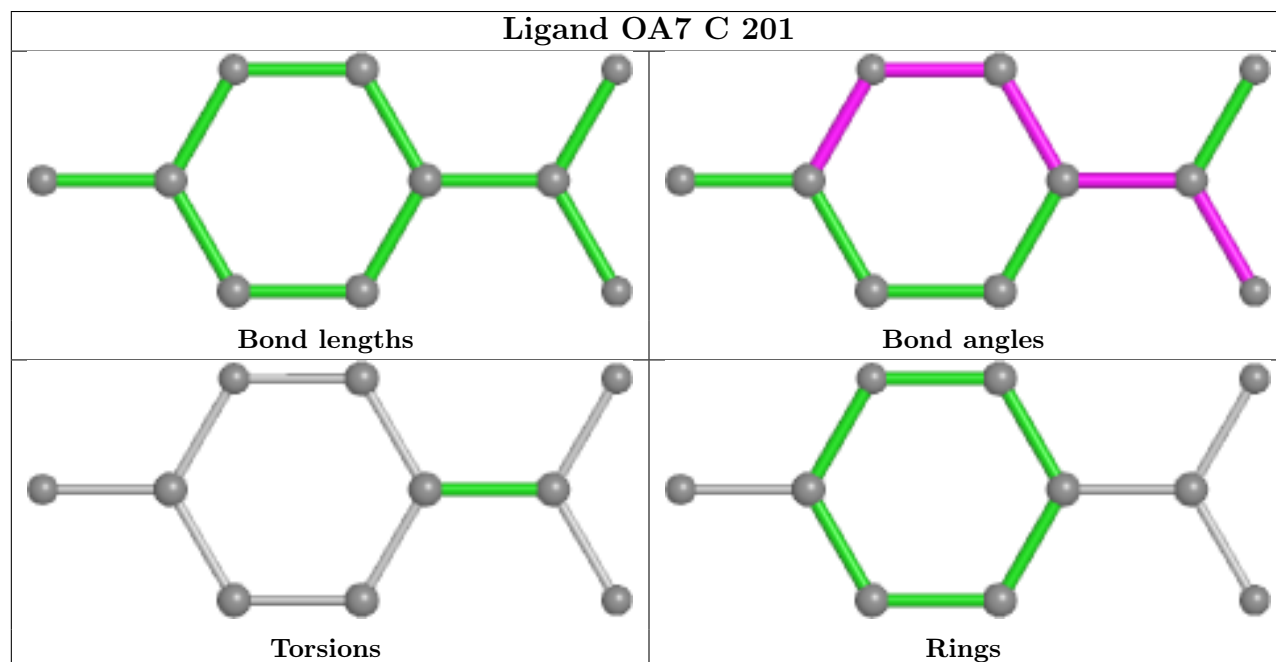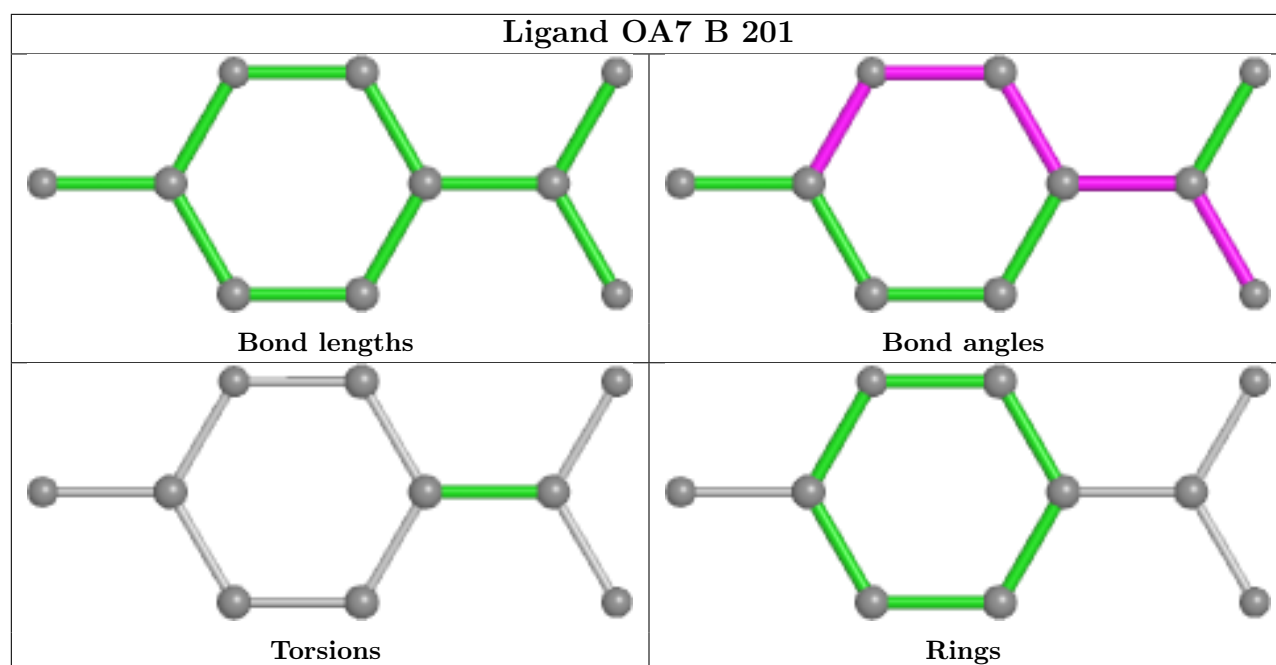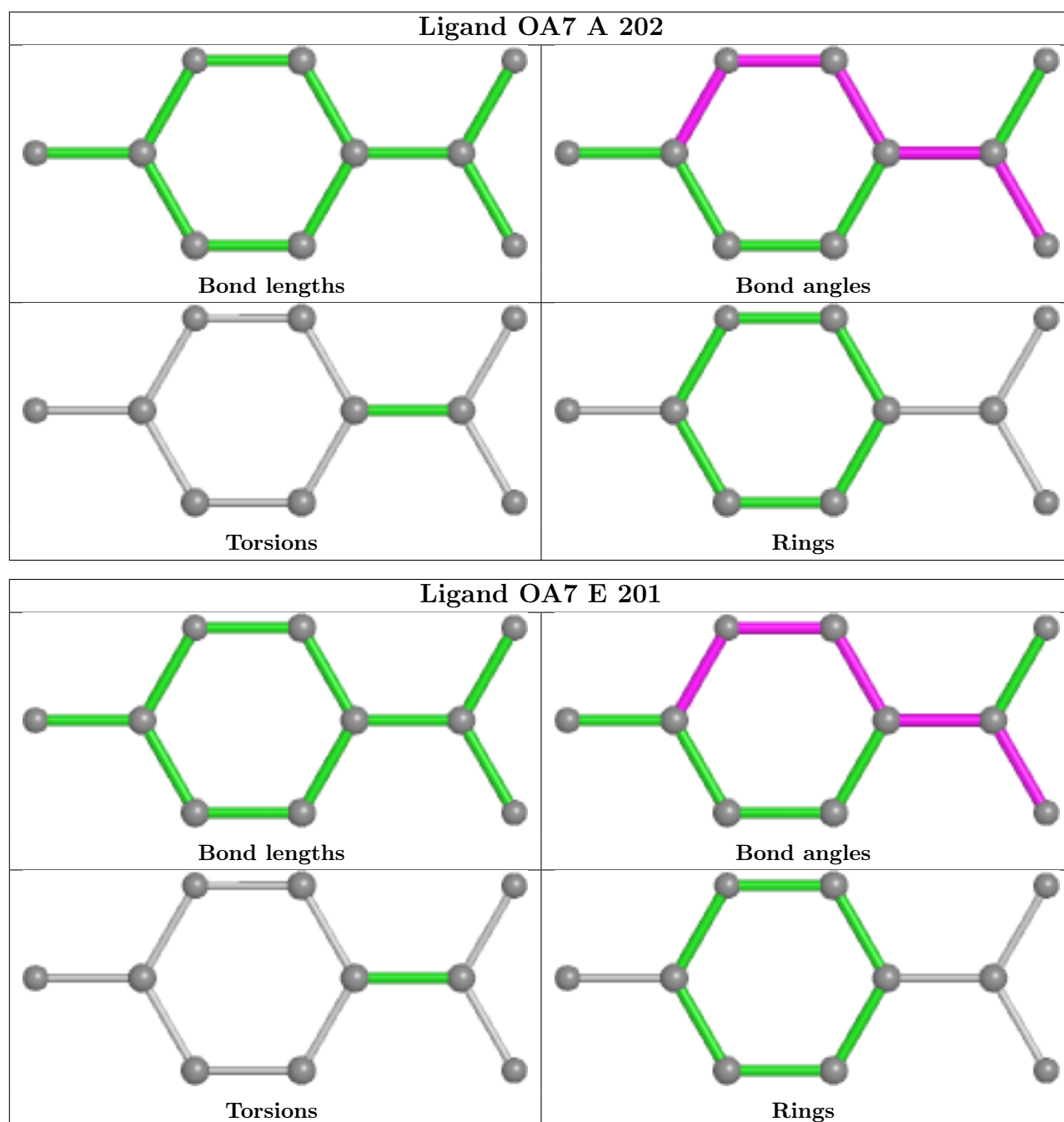
5.7 Other polymers ⓘ

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	A	140/161 (86%)	-0.40	0	100	100	17, 39, 80, 105	0
1	B	135/161 (83%)	-0.51	0	100	100	16, 37, 70, 106	0
1	C	140/161 (86%)	-0.49	0	100	100	22, 37, 65, 82	0
1	D	139/161 (86%)	-0.18	3 (2%)	62	48	19, 44, 96, 112	0
1	E	140/161 (86%)	-0.19	4 (2%)	51	36	20, 43, 88, 141	0
1	F	140/161 (86%)	-0.24	3 (2%)	63	49	24, 47, 88, 129	0
All	All	834/966 (86%)	-0.33	10 (1%)	79	67	16, 41, 82, 141	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	101	ARG	4.6
1	F	101	ARG	3.3
1	F	99	ASP	3.1
1	D	97	PRO	2.8
1	D	98	ALA	2.7

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
2	OA7	E	201	10/10	0.86	0.36	39,51,62,63	0
2	OA7	B	201	10/10	0.92	0.19	22,29,35,38	0
2	OA7	C	201	10/10	0.93	0.28	25,32,34,39	0
2	OA7	A	201	10/10	0.93	0.23	38,42,44,49	0
2	OA7	A	202	10/10	0.94	0.16	22,27,34,36	0
2	OA7	F	201	10/10	0.94	0.18	18,22,29,36	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.
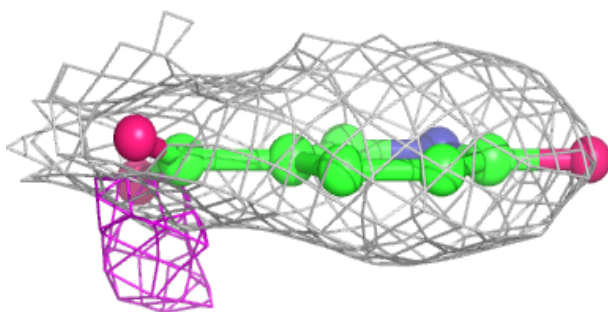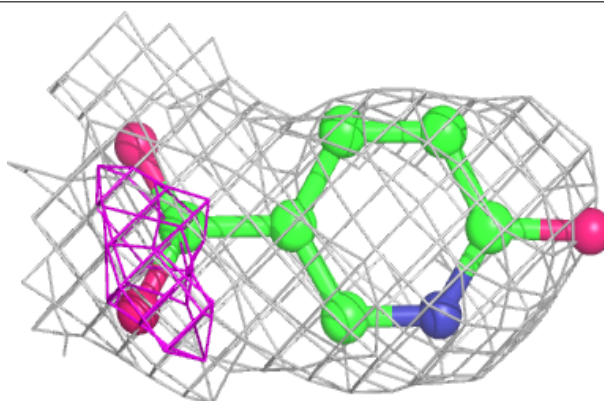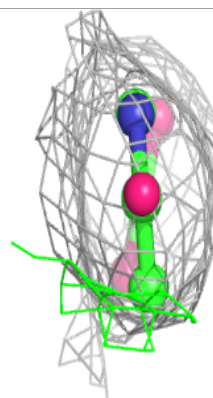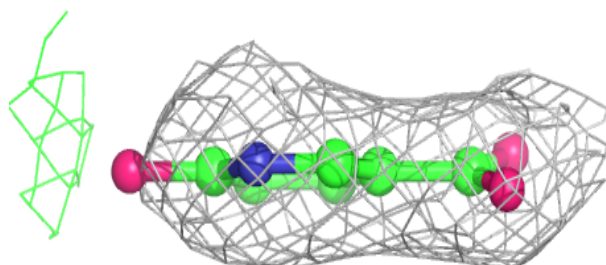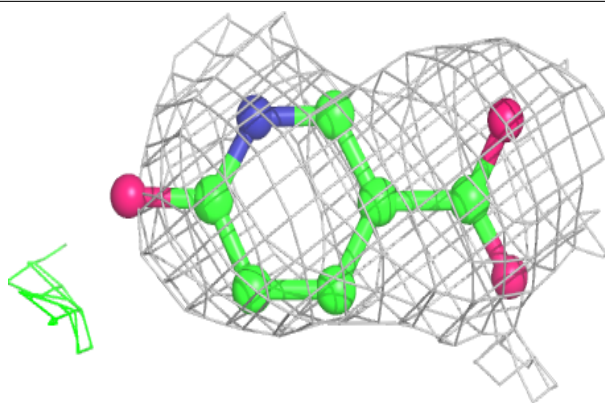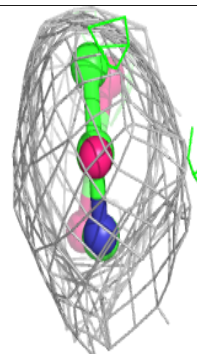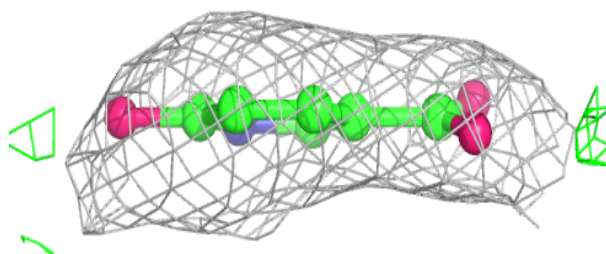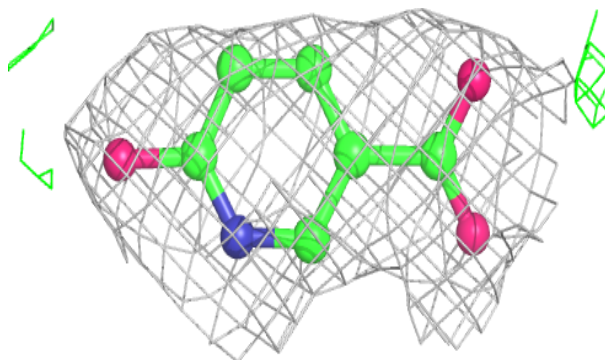
Electron density around OA7 E 201:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
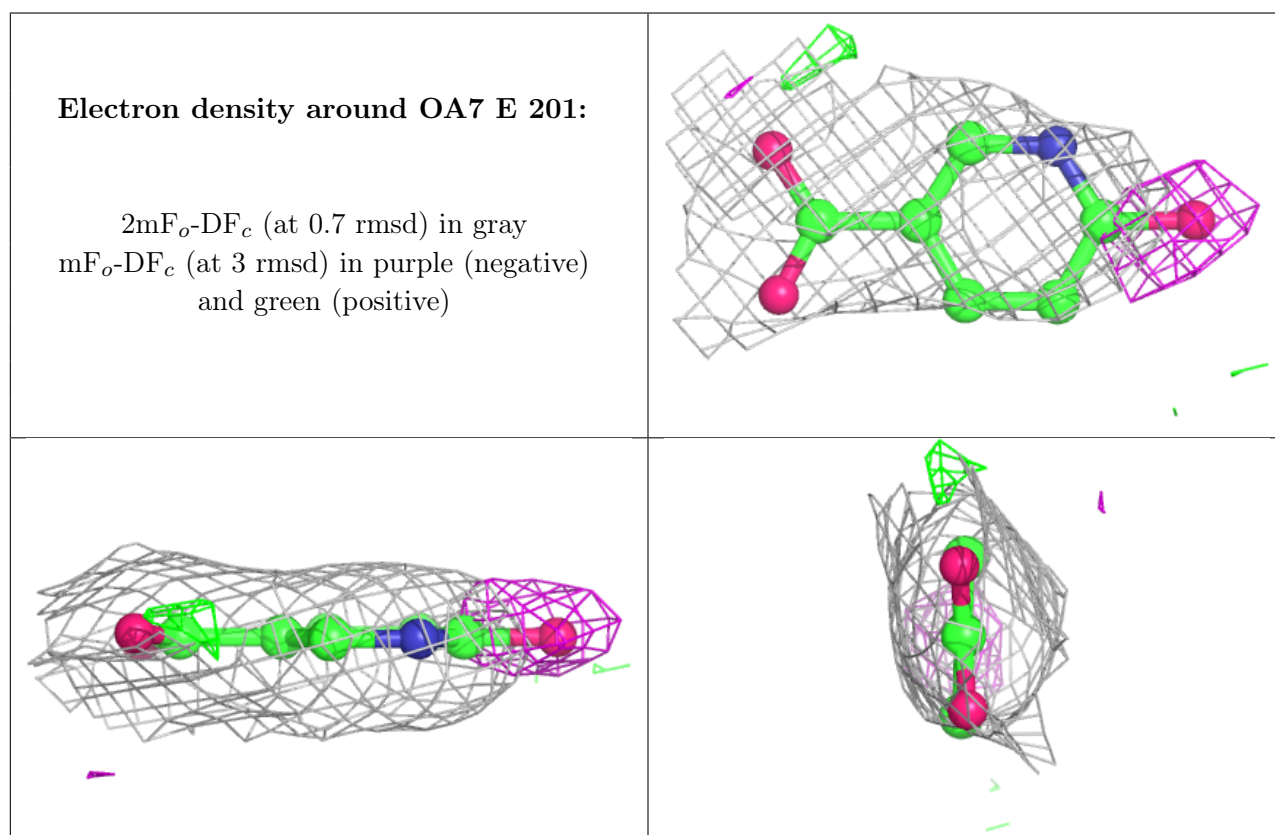
Electron density around OA7 B 201:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
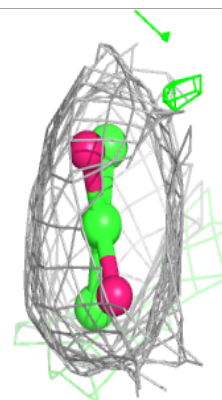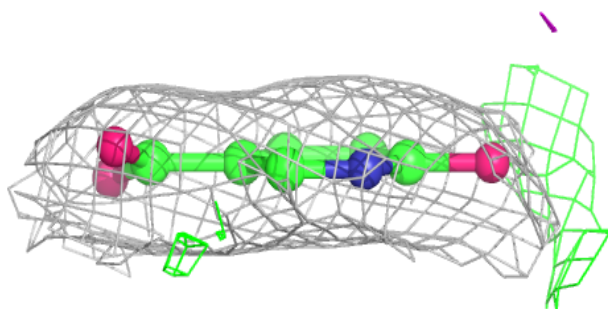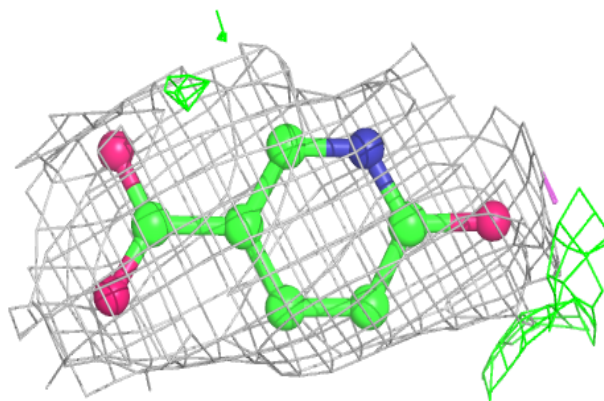
Electron density around OA7 C 201:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
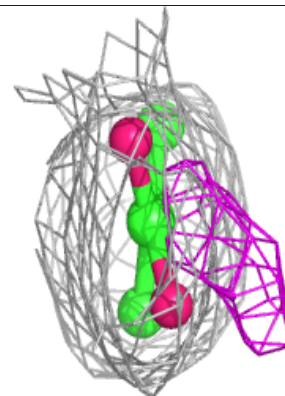
Electron density around OA7 A 201:

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

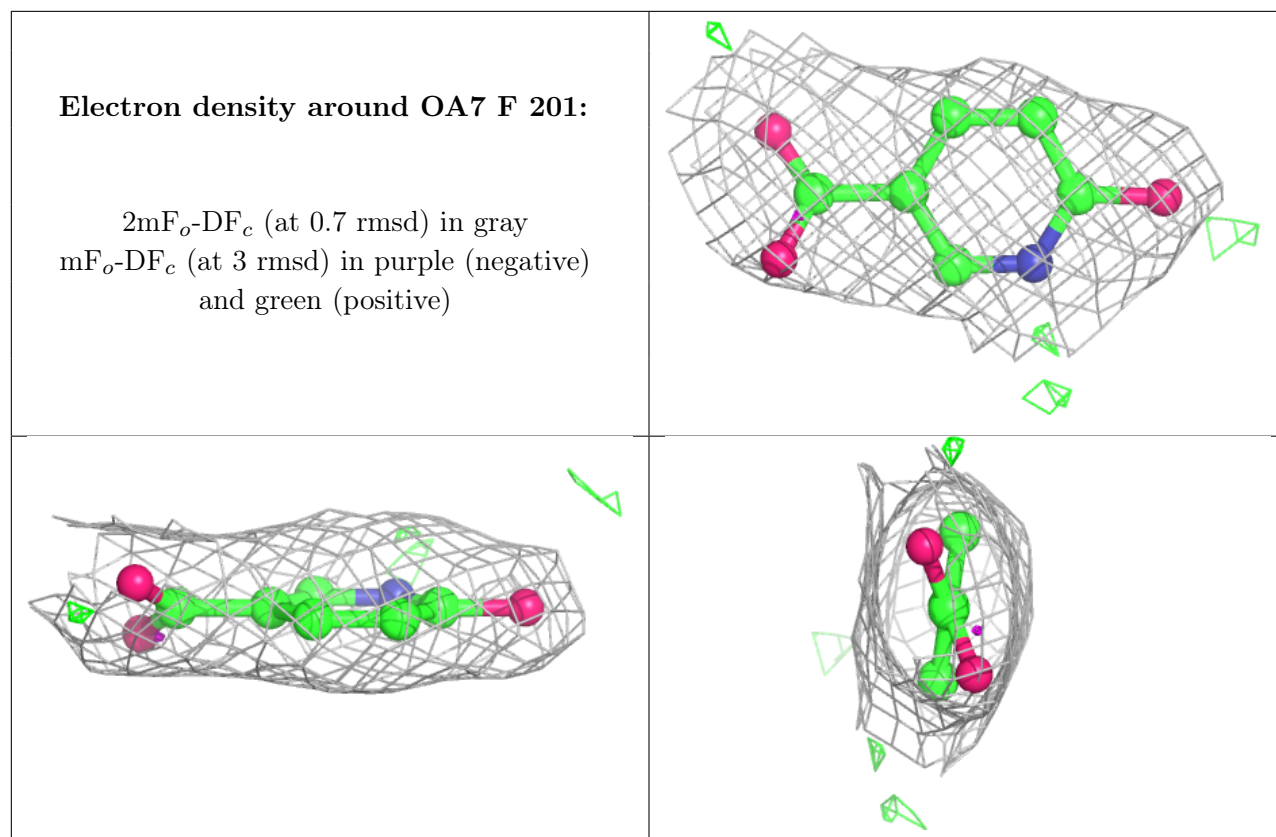**Electron density around OA7 A 202:**

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

Electron density around OA7 F 201:

$2mF_o$-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.