WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 25, 2022 – 11:05 pm BST

PDB ID	:	8AIJ
Title	:	STRUCTURE OF THE LECB LECTIN FROM PSEUDOMONAS AERUG-INOSA STRAIN PAO1 IN COMPLEX WITH N-(alpha-L-Fucopyranosyl)benzamide (6)
Authors	:	Meiers, J.; Mala, P.; Varrot, A.; Siebs, E.; Imberty, A.; Titz, A.
Deposited on	:	2022-07-26
Resolution	:	1.50 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 ⓘ 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 ⓘ) were used in the production of this report:

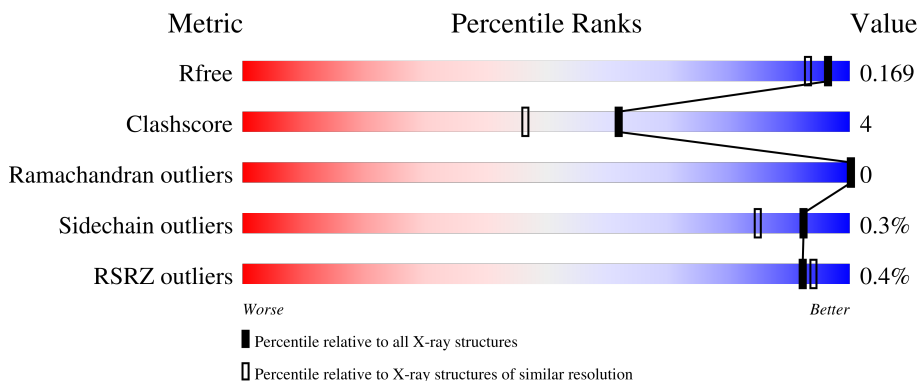|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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		0.169
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0.3%
RSRZ outliers		0.4%

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	114	94% 6%
1	BBB	114	93% 7%
1	CCC	114	96% •
1	DDD	114	90% 10%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 4065 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.
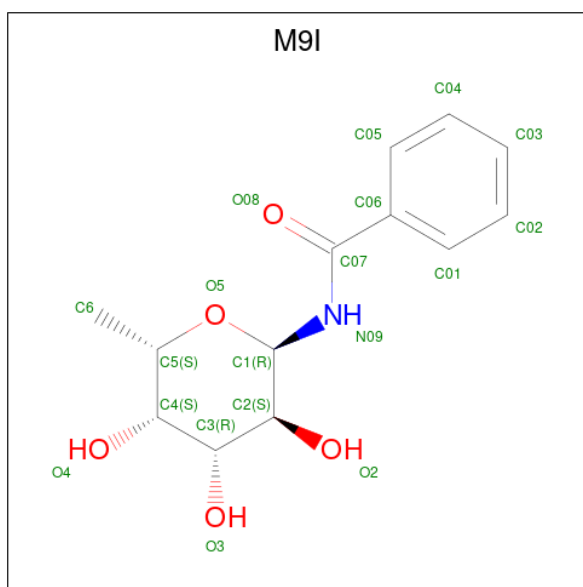
- Molecule 1 is a protein called Fucose-binding lectin PA-IIL.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	AAA	114	Total 846	C 520	N 147	O 179	0	3	0
1	BBB	114	Total 869	C 532	N 151	O 186	0	7	0
1	CCC	114	Total 843	C 520	N 144	O 179	0	4	0
1	DDD	114	Total 869	C 535	N 151	O 183	0	6	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	2	Total 2	Ca 2	0	0
2	BBB	2	Total 2	Ca 2	0	0
2	CCC	2	Total 2	Ca 2	0	0
2	DDD	2	Total 2	Ca 2	0	0

- Molecule 3 is N-(alpha-L-Fucopyranosyl)benzamide (three-letter code: M9I) (formula: $C_{13}H_{17}NO_5$) (labeled as "Ligand of Interest" by depositor).

M9I

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total 19	C 13	N 1	O 5	0	0
3	BBB	1	Total 19	C 13	N 1	O 5	0	0
3	CCC	1	Total 19	C 13	N 1	O 5	0	0
3	DDD	1	Total 19	C 13	N 1	O 5	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total 4	C 2	O 2	0	0
4	AAA	1	Total 4	C 2	O 2	0	0
4	AAA	1	Total 4	C 2	O 2	0	0
4	BBB	1	Total 4	C 2	O 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	121	Total 131	O 131	0	10
5	BBB	125	Total 132	O 132	0	7
5	CCC	126	Total 133	O 133	0	7
5	DDD	137	Total 142	O 142	0	5

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: Fucose-binding lectin PA-IIL
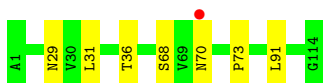
Chain AAA:

• Molecule 1: Fucose-binding lectin PA-IIL

Chain BBB:

• Molecule 1: Fucose-binding lectin PA-IIL

Chain CCC:

• Molecule 1: Fucose-binding lectin PA-IIL

Chain DDD:

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	45.38Å 51.59Å 52.60Å 101.81° 99.47° 115.81°	Depositor
Resolution (Å)	40.69 – 1.50 40.69 – 1.50	Depositor EDS
% Data completeness (in resolution range)	95.5 (40.69-1.50) 95.5 (40.69-1.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.32 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.138 , 0.169 0.139 , 0.169	Depositor DCC
R_{free} test set	7570 reflections (12.29%)	wwPDB-VP
Wilson B-factor (Å^2)	13.1	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]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4065	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.89% 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: M9I, EDO, CA

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	0.81	0/856	0.92	0/1172
1	BBB	0.80	0/879	0.93	0/1204
1	CCC	0.83	0/853	0.90	0/1170
1	DDD	0.78	0/879	0.87	0/1204
All	All	0.80	0/3467	0.90	0/4750

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	AAA	846	0	817	10	0
1	BBB	869	0	833	10	0
1	CCC	843	0	808	5	0
1	DDD	869	0	841	9	0
2	AAA	2	0	0	0	0
2	BBB	2	0	0	0	0
2	CCC	2	0	0	0	0
2	DDD	2	0	0	0	0
3	AAA	19	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	BBB	19	0	0	0	0
3	CCC	19	0	0	0	0
3	DDD	19	0	0	0	0
4	AAA	12	0	18	4	0
4	BBB	4	0	6	1	0
5	AAA	131	0	0	1	0
5	BBB	132	0	0	4	0
5	CCC	133	0	0	1	0
5	DDD	142	0	0	2	0
All	All	4065	0	3323	30	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 4.

The worst 5 of 30 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:29[B]:ASN:OD1	1:BBB:36:THR:HG23	1.60	1.02
1:BBB:86:GLU:HG3	5:BBB:301:HOH:O	1.72	0.88
1:AAA:31[B]:LEU:HD23	1:AAA:36:THR:HA	1.64	0.78
1:AAA:29[B]:ASN:OD1	1:AAA:36:THR:HG23	1.87	0.74
1:CCC:46[A]:ASN:OD1	1:DDD:13:ARG:NH2	2.21	0.73

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	AAA	115/114 (101%)	112 (97%)	3 (3%)	0	100	100
1	BBB	119/114 (104%)	115 (97%)	4 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	CCC	116/114 (102%)	114 (98%)	2 (2%)	0	100	100
1	DDD	118/114 (104%)	115 (98%)	3 (2%)	0	100	100
All	All	468/456 (103%)	456 (97%)	12 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	95/93 (102%)	95 (100%)	0	100	100
1	BBB	99/93 (106%)	99 (100%)	0	100	100
1	CCC	94/93 (101%)	94 (100%)	0	100	100
1	DDD	98/93 (105%)	97 (99%)	1 (1%)	76	57
All	All	386/372 (104%)	385 (100%)	1 (0%)	92	85

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	34	ASN

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 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	M9I	CCC	203	2	20,20,20	0.48	0	27,28,28	1.06	2 (7%)
4	EDO	AAA	205	-	3,3,3	0.26	0	2,2,2	0.54	0
3	M9I	BBB	203	2	20,20,20	0.41	0	27,28,28	0.85	2 (7%)
3	M9I	AAA	203	2	20,20,20	0.60	0	27,28,28	0.96	1 (3%)
4	EDO	AAA	206	-	3,3,3	0.23	0	2,2,2	0.69	0
4	EDO	BBB	204	-	3,3,3	0.25	0	2,2,2	0.85	0
4	EDO	AAA	204	-	3,3,3	0.34	0	2,2,2	0.65	0
3	M9I	DDD	203	2	20,20,20	0.62	0	27,28,28	0.98	1 (3%)

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	M9I	CCC	203	2	-	0/8/28/28	0/2/2/2
4	EDO	AAA	205	-	-	1/1/1/1	-
3	M9I	BBB	203	2	-	4/8/28/28	0/2/2/2
3	M9I	AAA	203	2	-	1/8/28/28	0/2/2/2
4	EDO	AAA	206	-	-	1/1/1/1	-
4	EDO	BBB	204	-	-	1/1/1/1	-
4	EDO	AAA	204	-	-	1/1/1/1	-
3	M9I	DDD	203	2	-	0/8/28/28	0/2/2/2

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	203	M9I	O5-C1-N09	2.80	113.05	108.01
3	CCC	203	M9I	O5-C1-N09	2.66	112.79	108.01
3	CCC	203	M9I	O4-C4-C3	-2.56	104.44	110.35
3	BBB	203	M9I	O5-C1-N09	2.27	112.09	108.01
3	DDD	203	M9I	O2-C2-C1	2.16	113.25	109.39

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

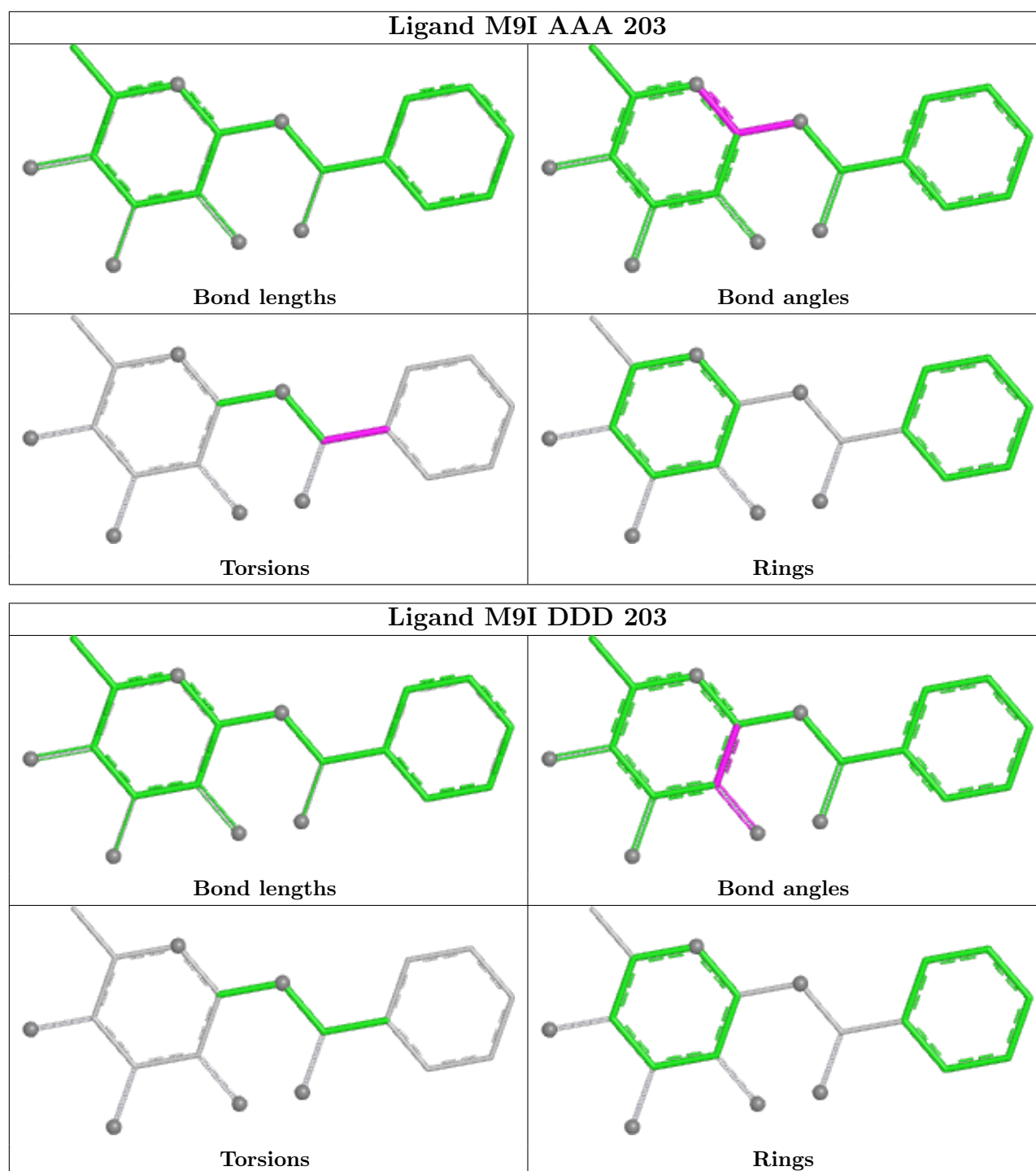| Mol	Chain	Res	Type	Atoms
3	BBB	203	M9I	C01-C06-C07-N09
3	BBB	203	M9I	C05-C06-C07-N09
3	BBB	203	M9I	C05-C06-C07-O08
3	BBB	203	M9I	C01-C06-C07-O08
4	AAA	206	EDO	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AAA	206	EDO	2	0
4	BBB	204	EDO	1	0
4	AAA	204	EDO	2	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 M9I CCC 203

Bond lengths

Bond angles

Torsions

Rings

Ligand M9I BBB 203

Bond lengths

Bond angles

Torsions

Rings

Ligand M9I AAA 203

Bond lengths

Bond angles

Torsions

Rings

Ligand M9I DDD 203

Bond lengths

Bond angles

Torsions

Rings
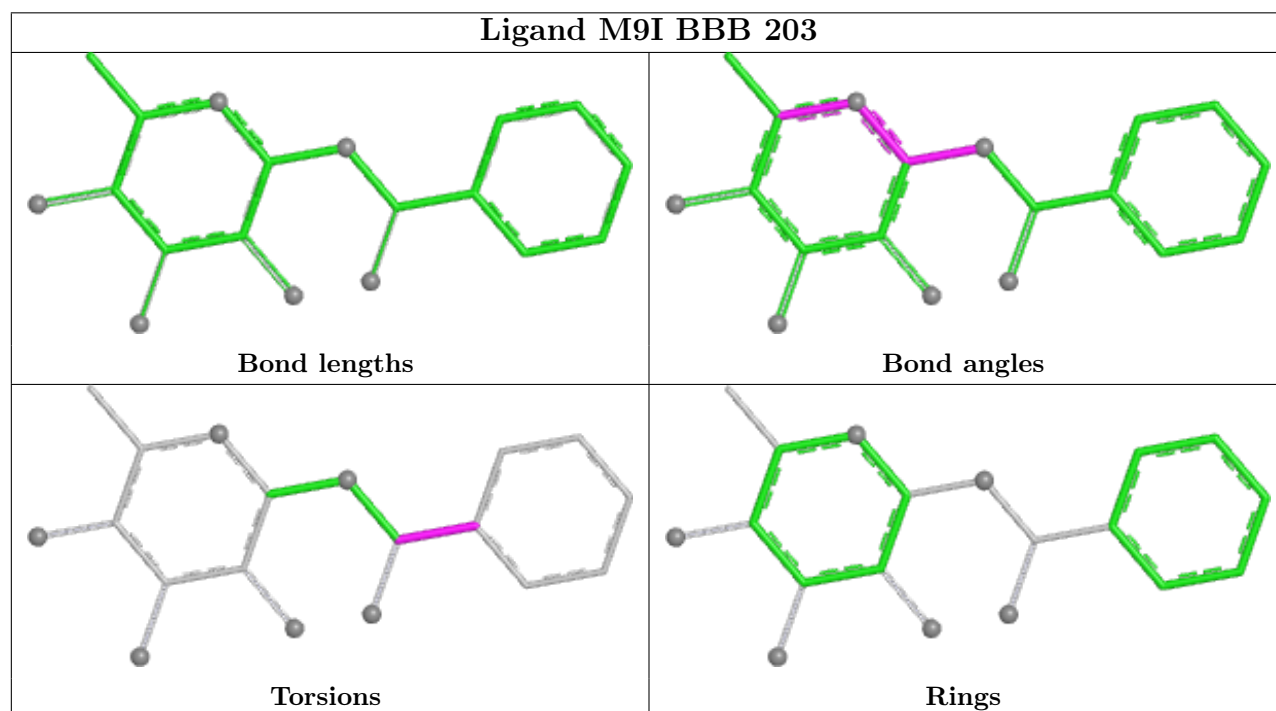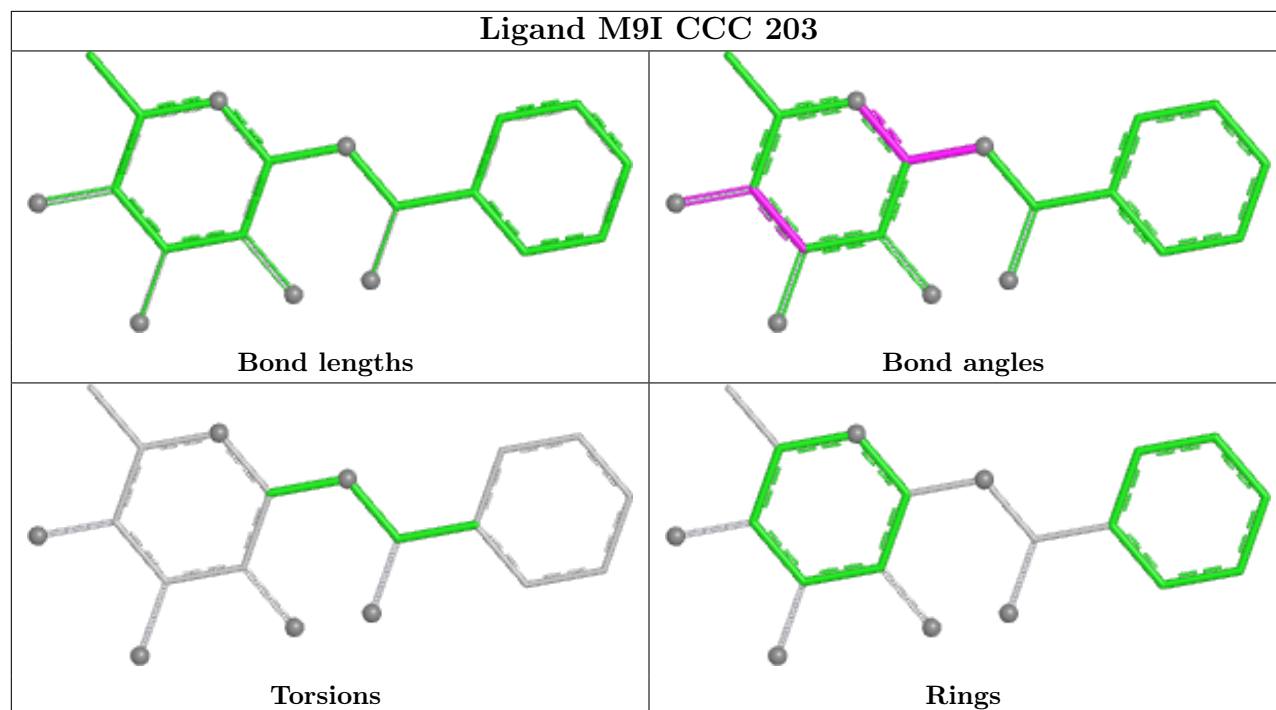
5.7 Other polymers (i)

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 ⓘ

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	114/114 (100%)	-0.23	1 (0%)	84	87	9, 13, 21, 30	1 (0%)
1	BBB	114/114 (100%)	-0.34	0	100	100	9, 13, 23, 34	1 (0%)
1	CCC	114/114 (100%)	-0.37	0	100	100	9, 13, 19, 32	0
1	DDD	114/114 (100%)	-0.30	1 (0%)	84	87	9, 12, 21, 31	1 (0%)
All	All	456/456 (100%)	-0.31	2 (0%)	92	94	9, 13, 22, 34	3 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	70	ASN	2.8
1	DDD	32[A]	VAL	2.0

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

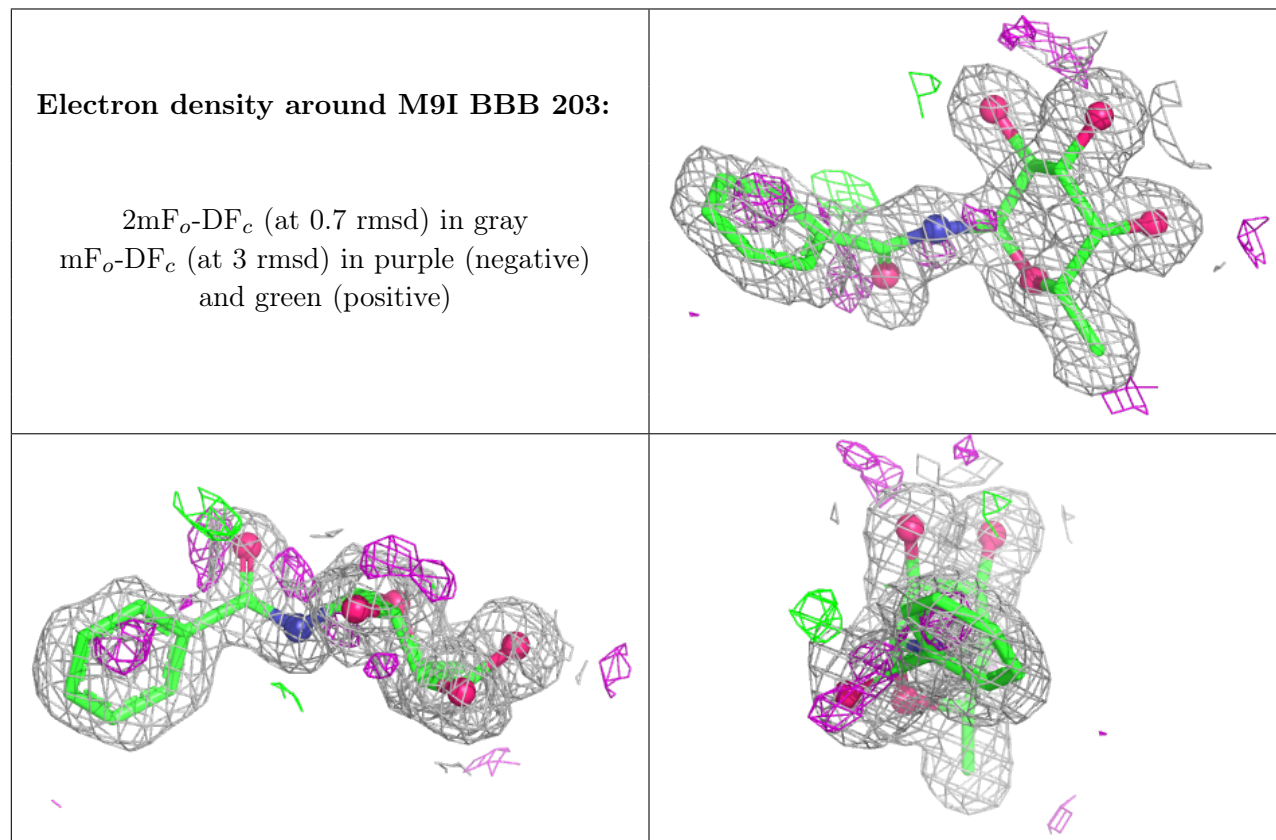There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	AAA	205	4/4	0.83	0.18	38,39,41,41	0
4	EDO	BBB	204	4/4	0.85	0.25	30,30,32,32	0
4	EDO	AAA	204	4/4	0.87	0.21	27,29,29,31	0
4	EDO	AAA	206	4/4	0.91	0.21	26,30,30,31	0
3	M9I	BBB	203	19/19	0.96	0.12	12,14,30,31	0
3	M9I	DDD	203	19/19	0.96	0.10	11,13,23,23	0
3	M9I	AAA	203	19/19	0.96	0.09	11,14,29,30	0
3	M9I	CCC	203	19/19	0.97	0.09	11,13,30,30	0
2	CA	AAA	201	1/1	1.00	0.06	10,10,10,10	0
2	CA	AAA	202	1/1	1.00	0.05	11,11,11,11	0
2	CA	BBB	201	1/1	1.00	0.04	11,11,11,11	0
2	CA	BBB	202	1/1	1.00	0.04	9,9,9,9	0
2	CA	CCC	201	1/1	1.00	0.06	10,10,10,10	0
2	CA	CCC	202	1/1	1.00	0.04	11,11,11,11	0
2	CA	DDD	201	1/1	1.00	0.06	9,9,9,9	0
2	CA	DDD	202	1/1	1.00	0.04	11,11,11,11	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.
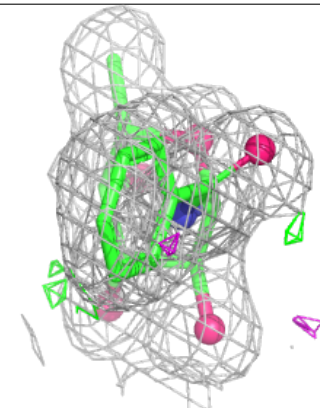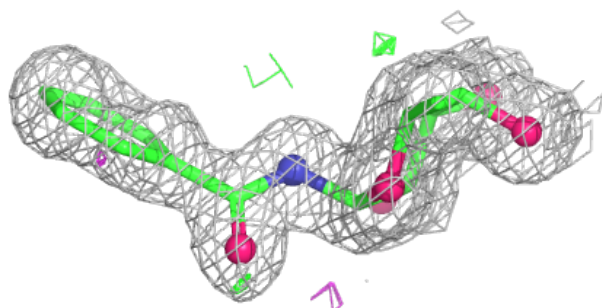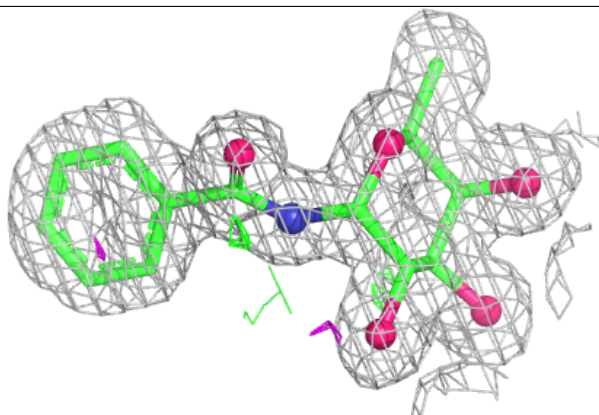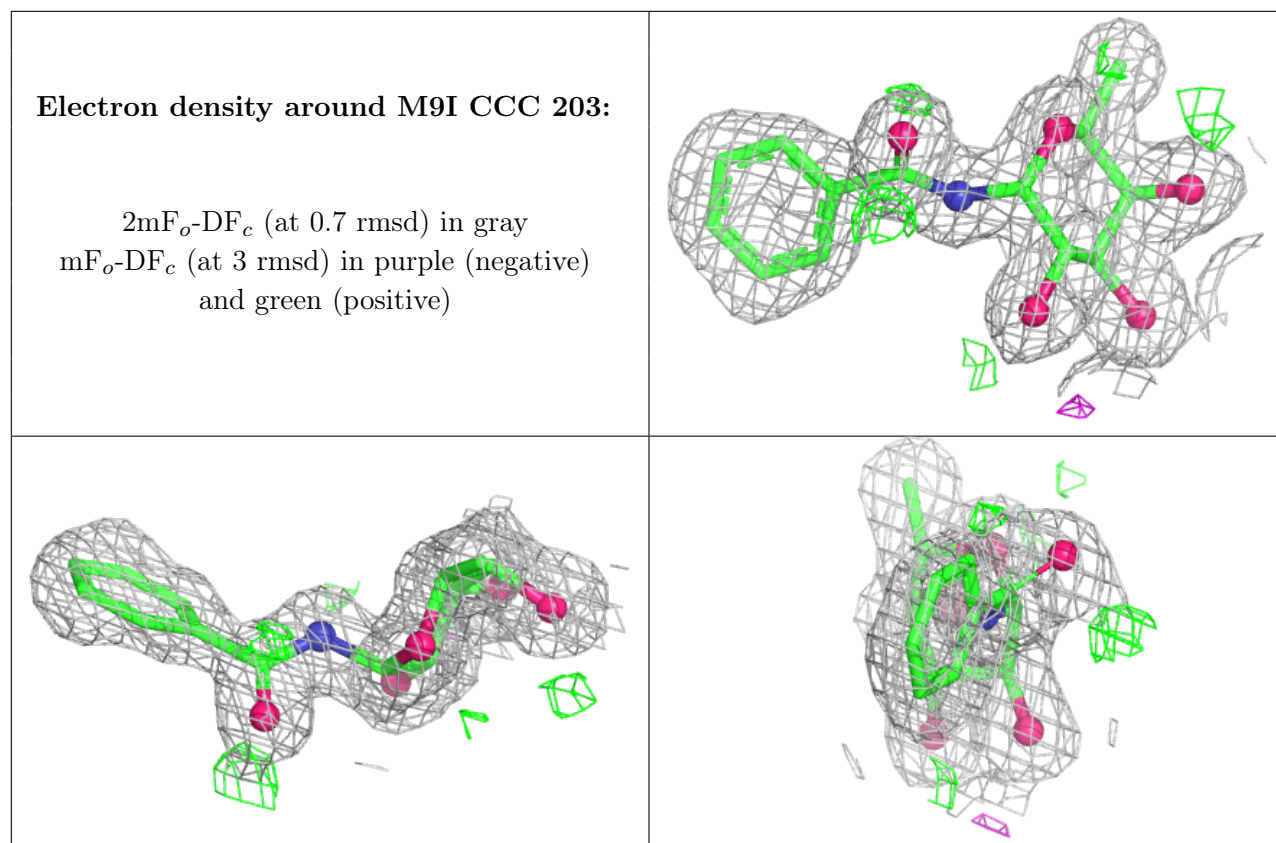

Electron density around M9I BBB 203:

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

Electron density around M9I DDD 203:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
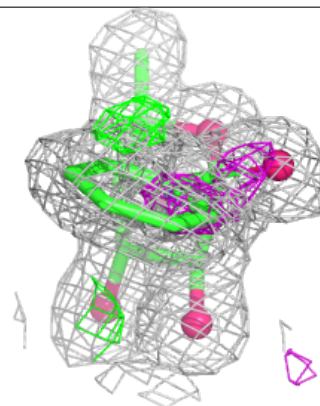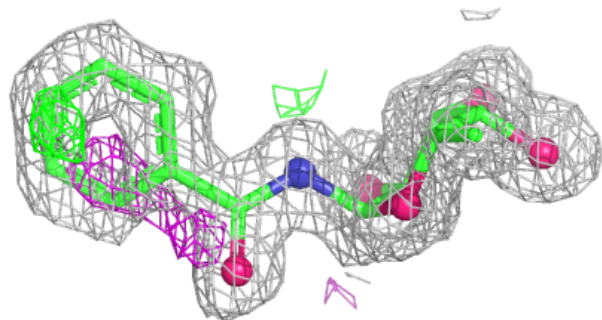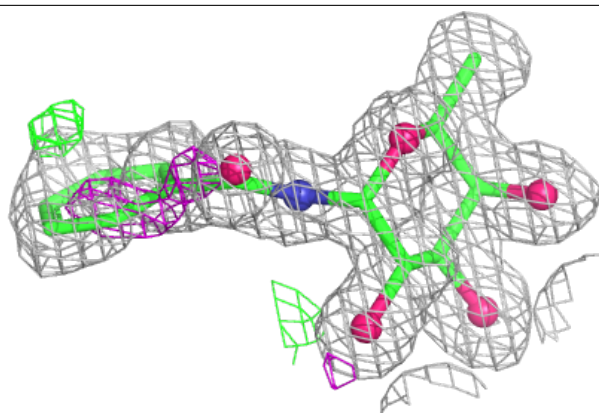
Electron density around M9I AAA 203:

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

Electron density around M9I CCC 203:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.