wwPDB X-ray Structure Validation Summary Report (i)

Jan 10, 2023 – 01:17 pm GMT

PDB ID	:	7Z63
Title	:	Structure of the LecA lectin from Pseudomonas aeruginosa in complex with a biaryl-thiogalactoside
Authors	:	Varrot, A.
Deposited on	:	2022-03-10
Resolution	:	1.75 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 (i)) were used in the production of this report:

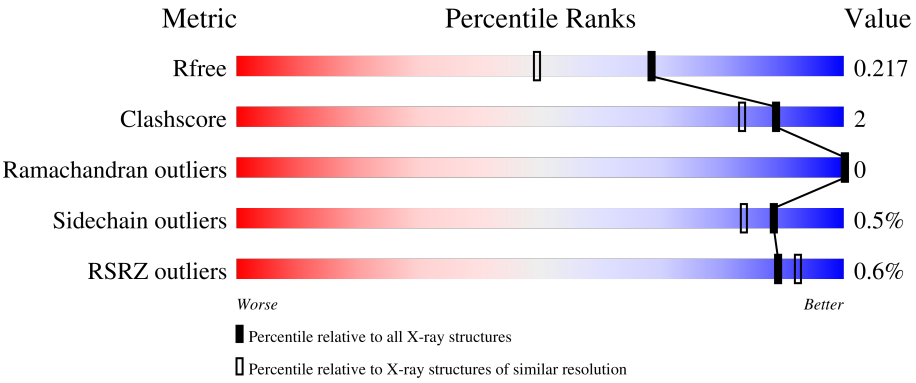
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	121	94% 6%
1	BBB	121	2% 96% .
1	CCC	121	96% .
1	DDD	121	95% 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4649 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.

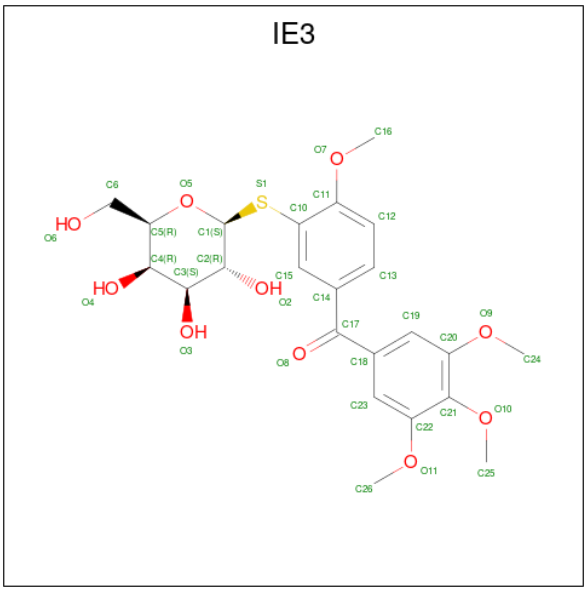
- Molecule 1 is a protein called PA-I galactophilic lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	121	Total	C	N	O	S	0	2	0
			914	575	158	178	3			
1	BBB	121	Total	C	N	O	S	0	1	0
			907	570	157	177	3			
1	CCC	121	Total	C	N	O	S	0	2	0
			918	576	160	179	3			
1	DDD	121	Total	C	N	O	S	0	2	0
			916	576	159	178	3			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

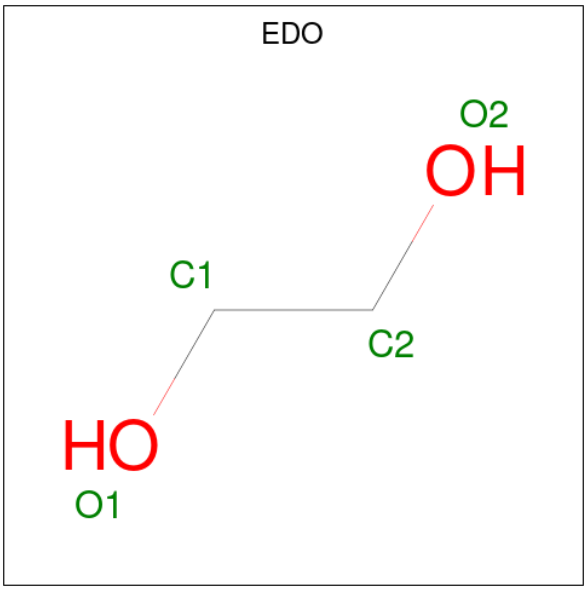
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	1	Total	Ca	0	0
			1	1		
2	BBB	1	Total	Ca	0	0
			1	1		
2	CCC	1	Total	Ca	0	0
			1	1		
2	DDD	1	Total	Ca	0	0
			1	1		

- Molecule 3 is [3-[(2 {S},3 {R},4 {S},5 {R},6 {R})-6-(hydroxymethyl)-3,4,5-tris(oxidanyl)oxan-2-yl]sulfanyl-4-methoxy-phenyl]-(3,4,5-trimethoxyphenyl)methanone (three-letter code: IE3) (formula: $C_{23}H_{28}O_{10}S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C O S 34 23 10 1	0	0
3	BBB	1	Total C O S 23 15 7 1	0	0
3	CCC	1	Total C O S 34 23 10 1	0	0
3	DDD	1	Total C O S 34 23 10 1	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	CCC	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

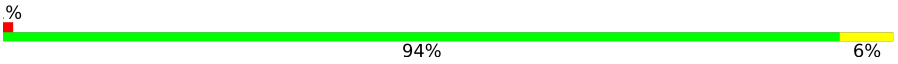
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	195	Total O 201 201	0	6
5	BBB	174	Total O 178 178	0	4
5	CCC	246	Total O 257 257	0	11
5	DDD	218	Total O 225 225	0	7

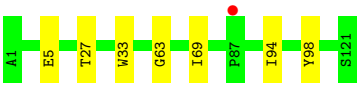
3 Residue-property plots (i)

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: PA-I galactophilic lectin

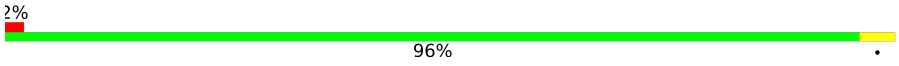
Chain AAA:

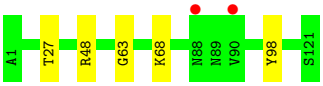




• Molecule 1: PA-I galactophilic lectin

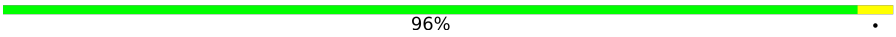
Chain BBB:

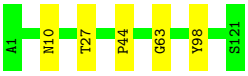




• Molecule 1: PA-I galactophilic lectin

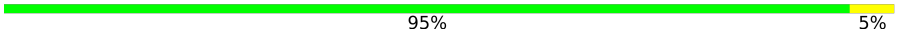
Chain CCC:





• Molecule 1: PA-I galactophilic lectin

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	42.26Å 61.93Å 180.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.13 – 1.75 43.09 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.9 (43.13-1.75) 99.9 (43.09-1.75)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.22 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.161 , 0.207 0.173 , 0.217	Depositor DCC
R_{free} test set	2506 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å^2)	10.3	Xtriage
Anisotropy	1.167	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 41.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4649	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.95% 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: IE3, 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	0.88	0/937	0.88	0/1279
1	BBB	0.90	0/930	0.90	0/1270
1	CCC	0.83	0/941	0.91	0/1285
1	DDD	0.87	0/939	0.94	0/1281
All	All	0.87	0/3747	0.91	0/5115

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	914	0	870	4	0
1	BBB	907	0	865	4	0
1	CCC	918	0	873	5	0
1	DDD	916	0	877	4	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
2	CCC	1	0	0	0	0
2	DDD	1	0	0	0	0
3	AAA	34	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	23	0	0	0	0
3	CCC	34	0	0	0	0
3	DDD	34	0	0	0	0
4	CCC	4	0	5	2	0
5	AAA	201	0	0	0	0
5	BBB	178	0	0	1	0
5	CCC	257	0	0	2	0
5	DDD	225	0	0	1	0
All	All	4649	0	3490	14	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.

The worst 5 of 14 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:CCC:10[B]:ASN:ND2	5:CCC:301:HOH:O	2.21	0.72
1:CCC:44:PRO:HD2	4:CCC:203:EDO:H22	1.73	0.69
4:CCC:203:EDO:H11	1:DDD:44:PRO:HD2	1.83	0.58
1:CCC:10[A]:ASN:ND2	5:CCC:304:HOH:O	2.43	0.52
1:CCC:63:GLY:HA2	1:CCC:98:TYR:CZ	2.47	0.49

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	121/121 (100%)	118 (98%)	3 (2%)	0	100	100
1	BBB	120/121 (99%)	115 (96%)	5 (4%)	0	100	100
1	CCC	121/121 (100%)	117 (97%)	4 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	DDD	121/121 (100%)	118 (98%)	3 (2%)	0	100 100
All	All	483/484 (100%)	468 (97%)	15 (3%)	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	AAA	95/94 (101%)	94 (99%)	1 (1%)	73 60
1	BBB	95/94 (101%)	95 (100%)	0	100 100
1	CCC	96/94 (102%)	96 (100%)	0	100 100
1	DDD	96/94 (102%)	95 (99%)	1 (1%)	76 63
All	All	382/376 (102%)	380 (100%)	2 (0%)	88 83

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	5	GLU
1	DDD	121	SER

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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	IE3	BBB	202	2	24,24,36	1.79	5 (20%)	34,34,51	1.65	3 (8%)
3	IE3	AAA	202	2	36,36,36	1.62	4 (11%)	51,51,51	1.41	6 (11%)
4	EDO	CCC	203	-	3,3,3	1.11	0	2,2,2	0.45	0
3	IE3	CCC	202	2	36,36,36	1.71	5 (13%)	51,51,51	1.64	10 (19%)
3	IE3	DDD	202	2	36,36,36	1.28	4 (11%)	51,51,51	1.95	13 (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
3	IE3	BBB	202	2	-	4/12/32/42	0/2/2/3
3	IE3	AAA	202	2	-	0/22/42/42	0/3/3/3
4	EDO	CCC	203	-	-	1/1/1/1	-
3	IE3	CCC	202	2	-	2/22/42/42	0/3/3/3
3	IE3	DDD	202	2	-	6/22/42/42	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	BBB	202	IE3	C11-C10	5.60	1.47	1.40
3	CCC	202	IE3	C11-C10	5.37	1.47	1.40
3	AAA	202	IE3	C11-C10	5.33	1.47	1.40
3	CCC	202	IE3	C10-S1	-5.21	1.69	1.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	202	IE3	C10-S1	-4.49	1.70	1.77

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	BBB	202	IE3	C16-O7-C11	6.50	127.33	117.53
3	DDD	202	IE3	O11-C22-C21	5.56	124.94	115.16
3	DDD	202	IE3	O9-C20-C21	5.26	124.40	115.16
3	CCC	202	IE3	C24-O9-C20	4.85	124.85	117.53
3	DDD	202	IE3	O7-C11-C10	4.78	121.05	115.87

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	BBB	202	IE3	O5-C1-S1-C10
3	BBB	202	IE3	C2-C1-S1-C10
3	DDD	202	IE3	O5-C1-S1-C10
3	DDD	202	IE3	C2-C1-S1-C10
3	DDD	202	IE3	C21-C22-O11-C26

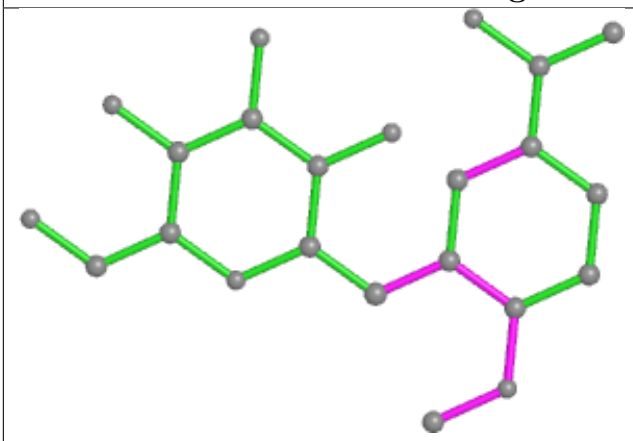
There are no ring outliers.

1 monomer is involved in 2 short contacts:

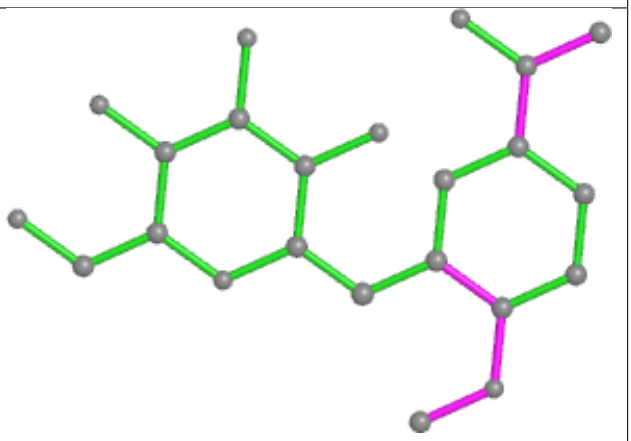
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	CCC	203	EDO	2	0

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.

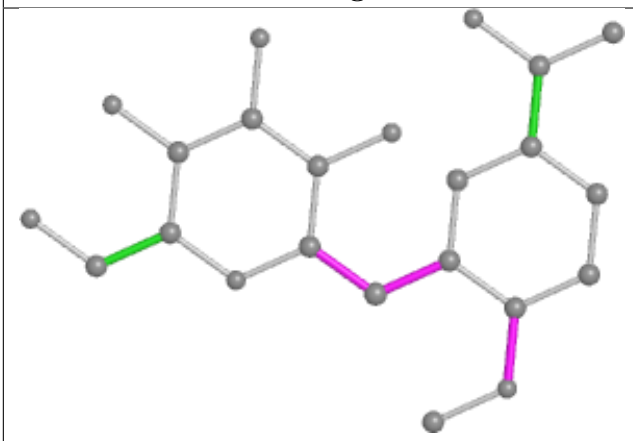
Ligand IE3 BBB 202



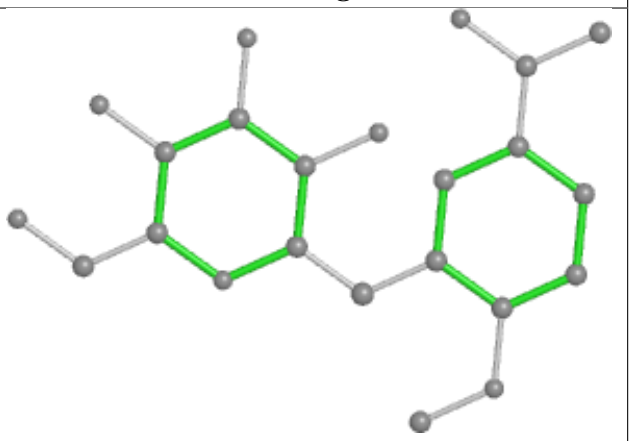
Bond lengths



Bond angles

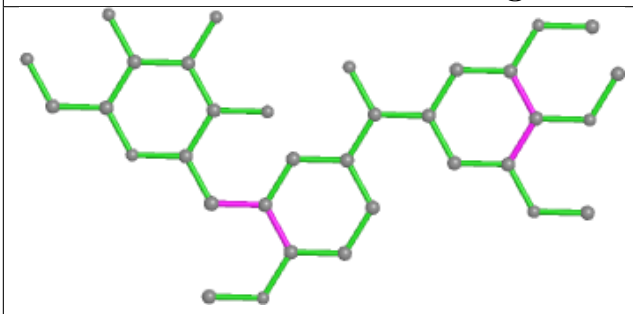


Torsions

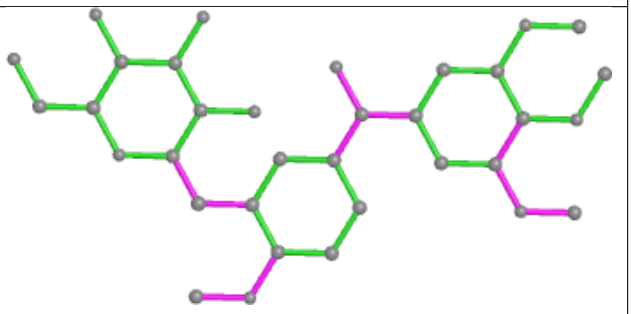


Rings

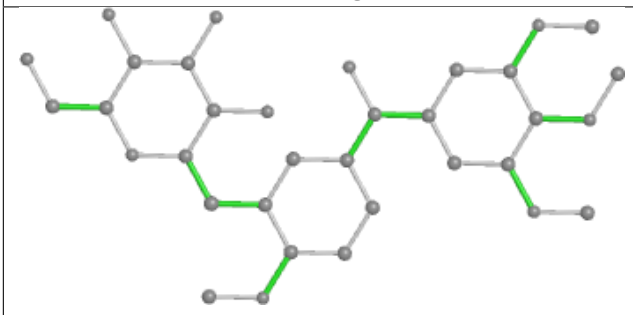
Ligand IE3 AAA 202



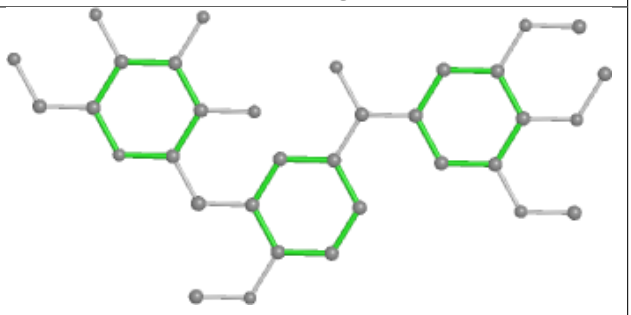
Bond lengths



Bond angles

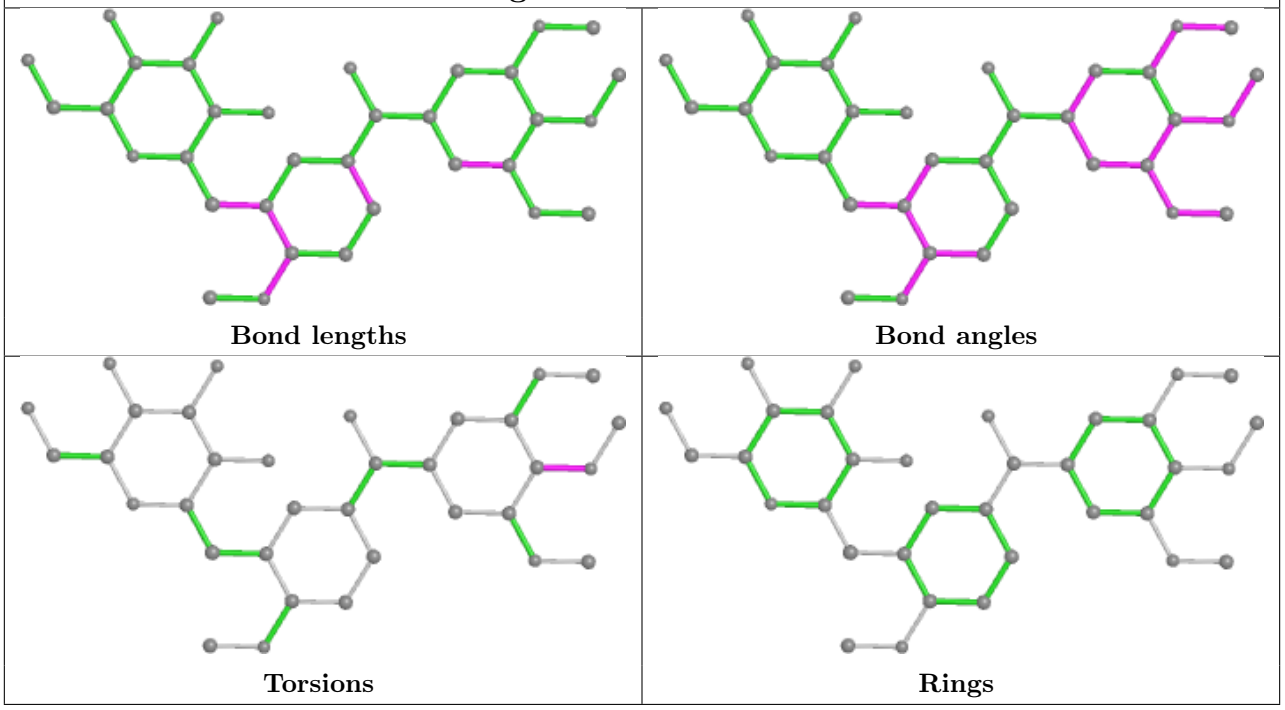


Torsions

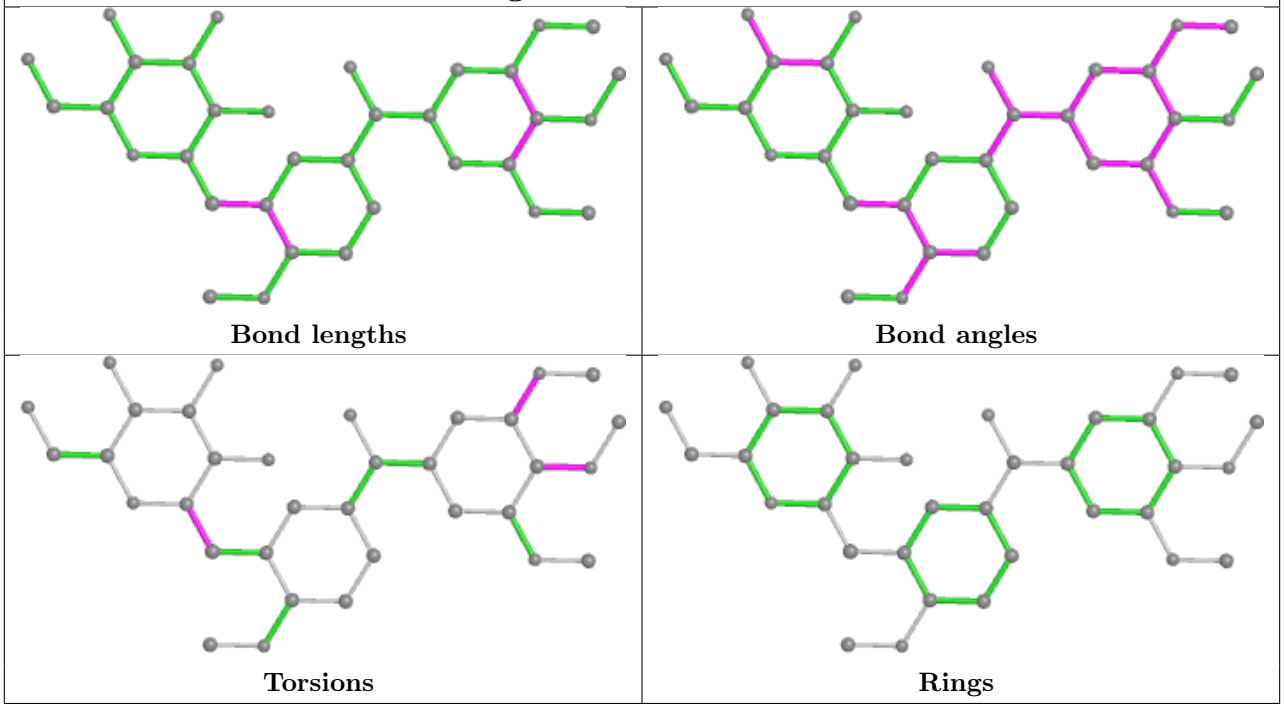


Rings

Ligand IE3 CCC 202



Ligand IE3 DDD 202



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, 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	AAA	121/121 (100%)	-0.09	1 (0%) 86 90	9, 15, 24, 41	0
1	BBB	121/121 (100%)	-0.29	2 (1%) 70 77	8, 12, 23, 39	0
1	CCC	121/121 (100%)	-0.43	0 100 100	5, 8, 16, 21	0
1	DDD	121/121 (100%)	-0.38	0 100 100	6, 10, 20, 33	0
All	All	484/484 (100%)	-0.30	3 (0%) 89 92	5, 11, 22, 41	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	88	ASN	3.1
1	BBB	90	VAL	2.6
1	AAA	87	PRO	2.5

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

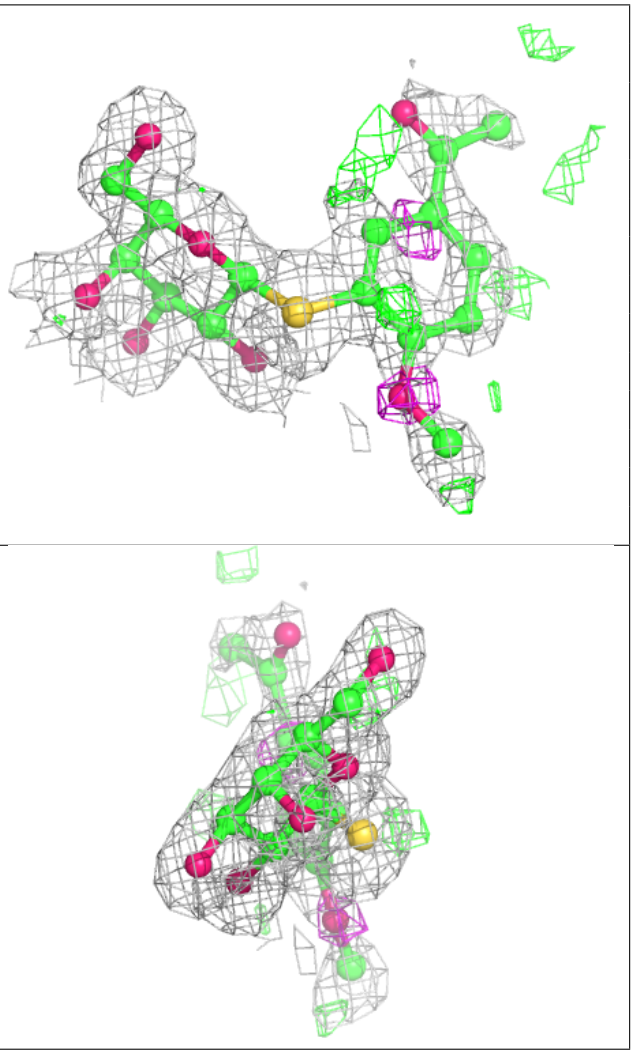
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	CCC	203	4/4	0.89	0.16	8,9,15,16	0
3	IE3	BBB	202	23/34	0.90	0.14	14,18,35,42	0
3	IE3	DDD	202	34/34	0.95	0.12	11,25,29,31	12
3	IE3	AAA	202	34/34	0.96	0.09	11,19,23,26	0
3	IE3	CCC	202	34/34	0.96	0.09	7,13,24,40	0
2	CA	DDD	201	1/1	0.99	0.04	9,9,9,9	0
2	CA	AAA	201	1/1	0.99	0.04	15,15,15,15	0
2	CA	BBB	201	1/1	0.99	0.04	11,11,11,11	0
2	CA	CCC	201	1/1	1.00	0.04	7,7,7,7	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.

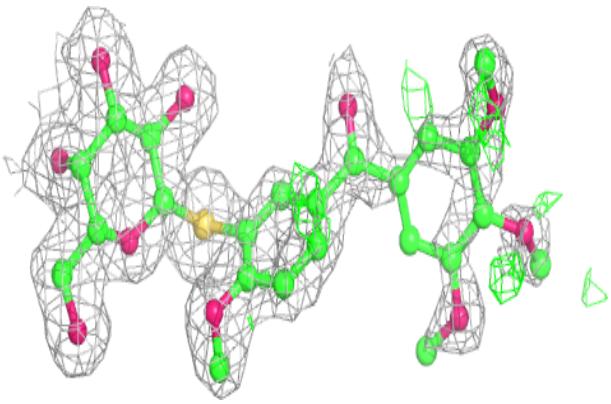
Electron density around IE3 BBB 202:

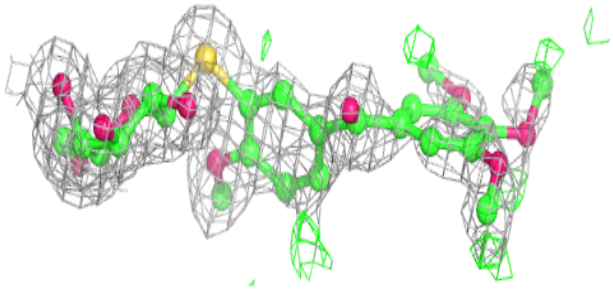
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

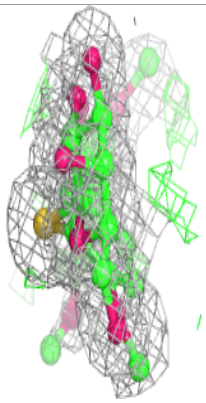


Electron density around IE3 DDD 202:

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

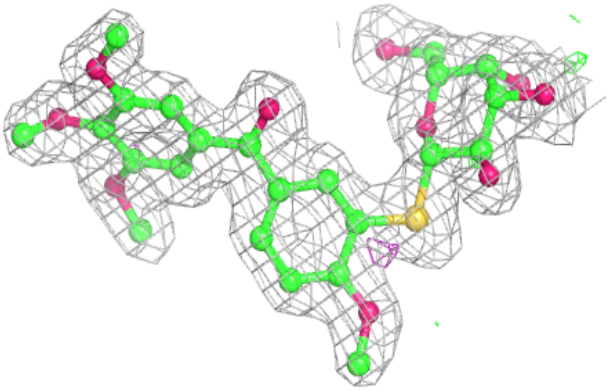


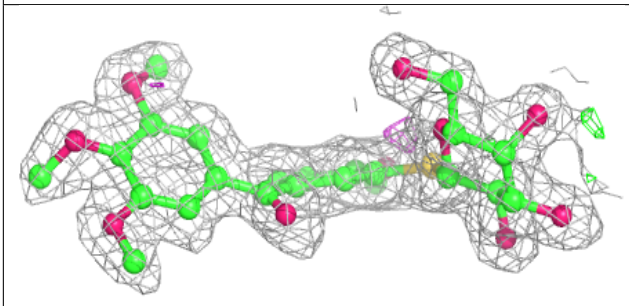


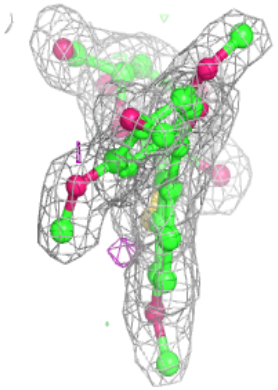


Electron density around IE3 AAA 202:

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

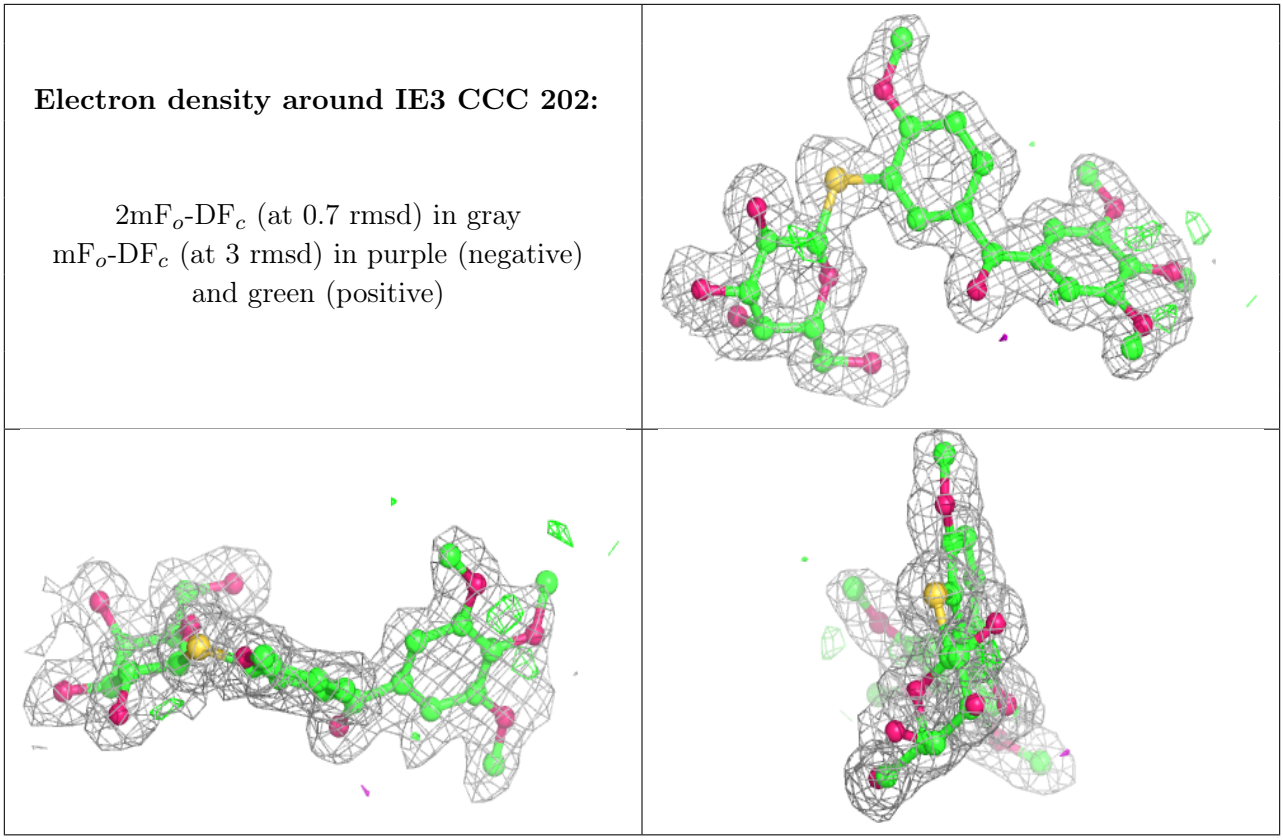






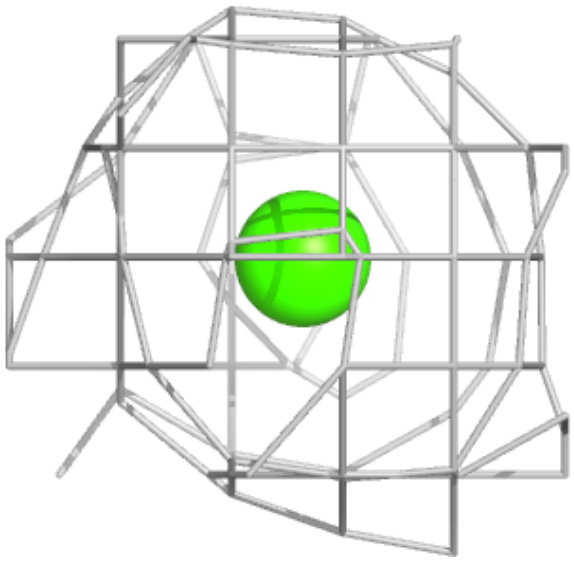
Electron density around IE3 CCC 202:

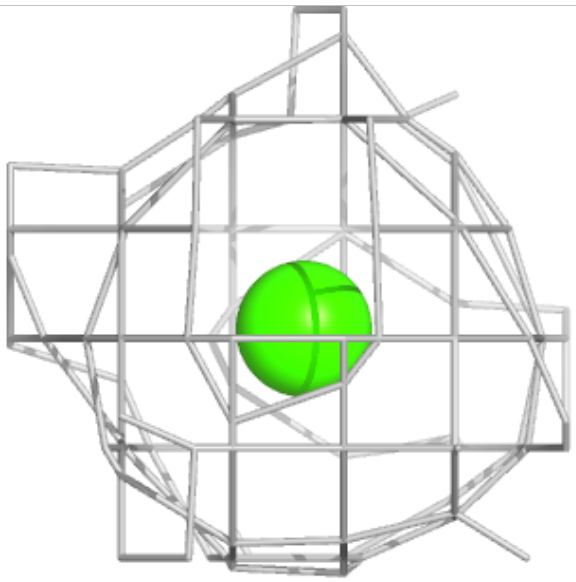
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

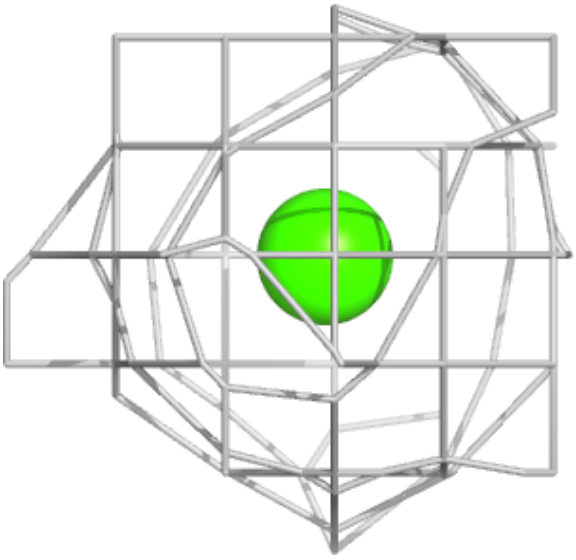


Electron density around CA DDD 201:

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

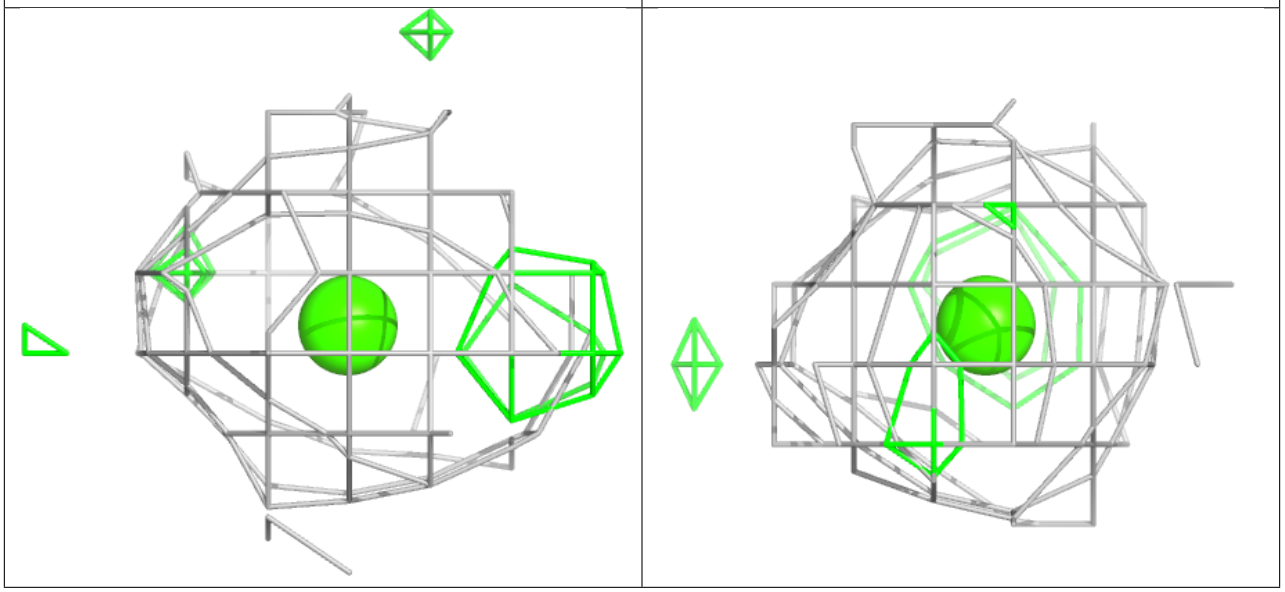






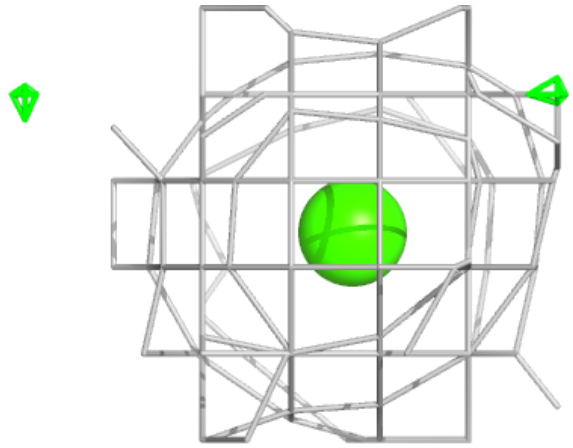
Electron density around CA AAA 201:

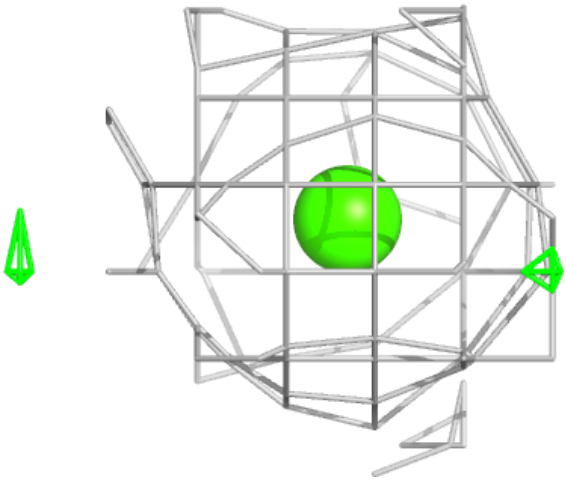
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

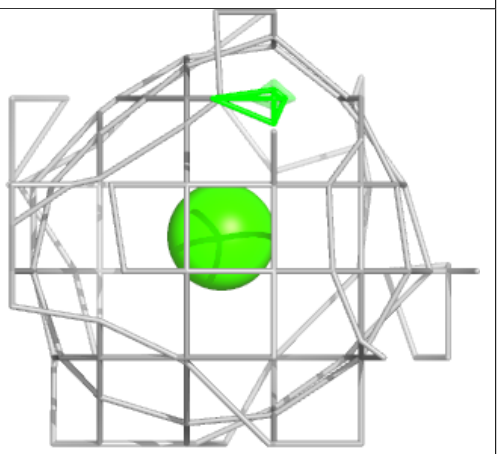


Electron density around CA BBB 201:

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

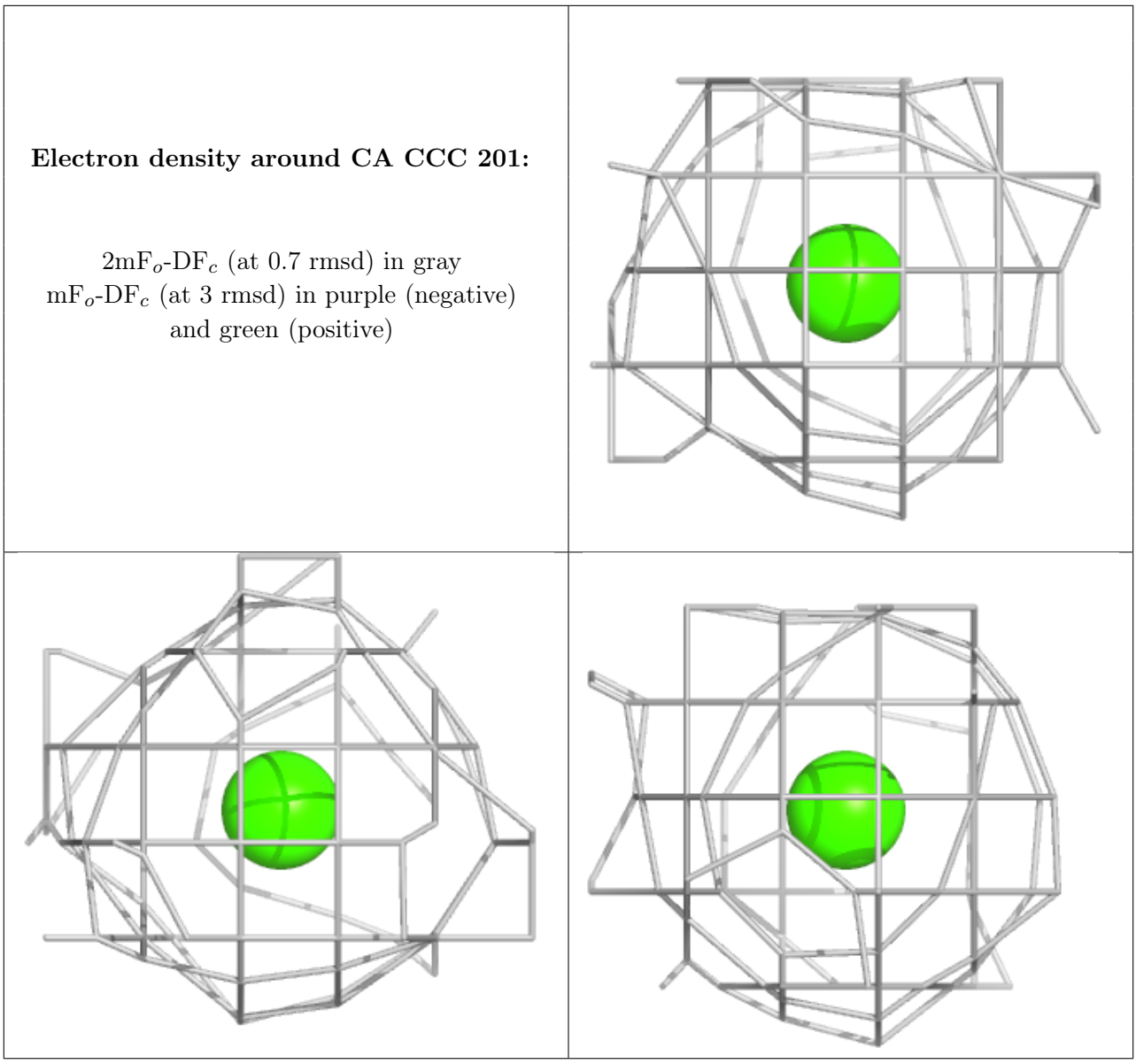






Electron density around CA CCC 201:

$2mF_o$-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 (i)

There are no such residues in this entry.