Full wwPDB X-ray Structure Validation Report ⓘ

Jun 8, 2023 – 01:10 pm BST

PDB ID	:	7Q2U
Title	:	The crystal structure of the HINT1 Q62A mutant.
Authors	:	Dolot, R.M.; Strom, A.M.; Wagner, C.R.
Deposited on	:	2021-10-26
Resolution	:	2.27 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

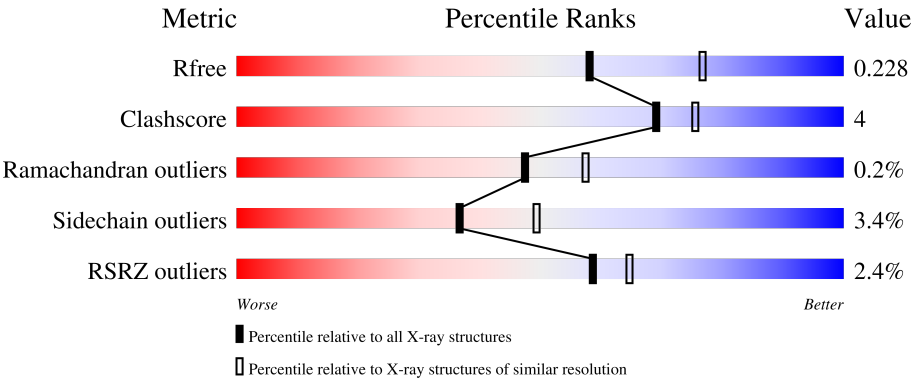
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.33
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 2.27 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	126	2% 83% 9% • 8%
1	BBB	126	2% 79% 12% • 8%
1	CCC	126	4% 87% 6% 8%
1	DDD	126	% 79% 13% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3917 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.

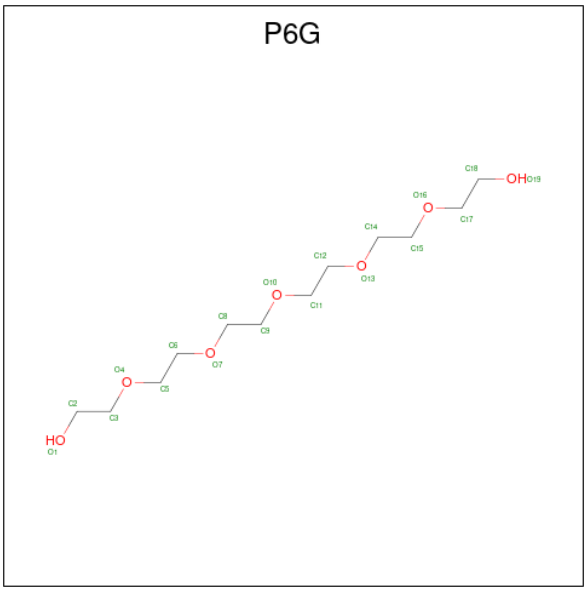
- Molecule 1 is a protein called Histidine triad nucleotide-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	116	Total	C	N	O	S	0	2	0
			907	574	163	165	5			
1	BBB	116	Total	C	N	O	S	0	2	0
			912	578	166	162	6			
1	CCC	116	Total	C	N	O	S	0	0	0
			893	567	161	160	5			
1	DDD	116	Total	C	N	O	S	0	2	0
			907	574	163	165	5			

There are 4 discrepancies between the modelled and reference sequences:

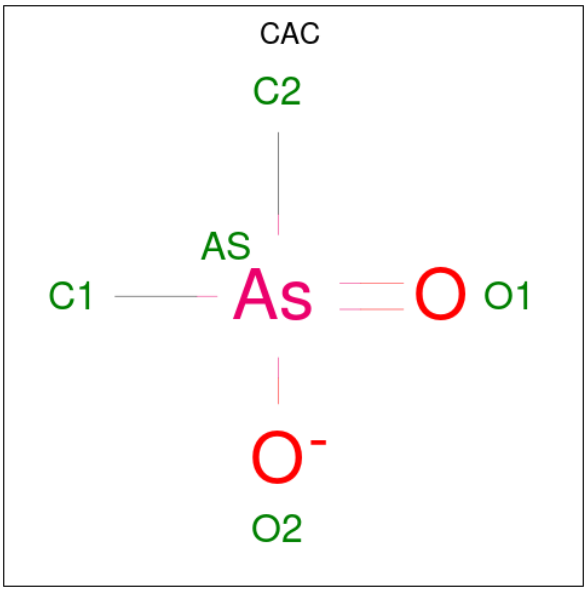
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	62	ALA	GLN	engineered mutation	UNP P49773
BBB	62	ALA	GLN	engineered mutation	UNP P49773
CCC	62	ALA	GLN	engineered mutation	UNP P49773
DDD	62	ALA	GLN	engineered mutation	UNP P49773

- Molecule 2 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$) (labeled as "Ligand of Interest" by depositor).



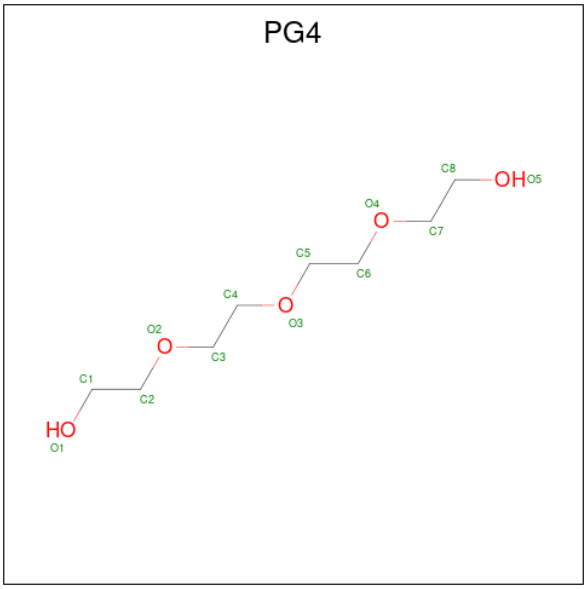
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C O 19 12 7	0	0

- Molecule 3 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	BBB	1	Total As C O 5 1 2 2	0	0
3	CCC	1	Total As C O 5 1 2 2	0	0

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	DDD	1	Total C O 13 8 5	0	0

- Molecule 5 is water.

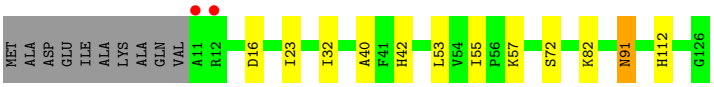
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	69	Total O 69 69	0	0
5	BBB	63	Total O 63 63	0	0
5	CCC	64	Total O 64 64	0	0
5	DDD	60	Total O 60 60	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

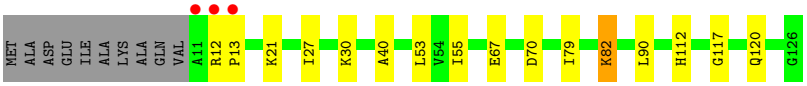
• Molecule 1: Histidine triad nucleotide-binding protein 1

Chain AAA:



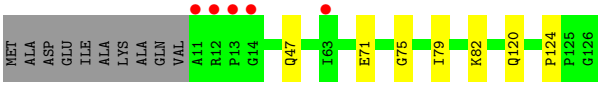
• Molecule 1: Histidine triad nucleotide-binding protein 1

Chain BBB:



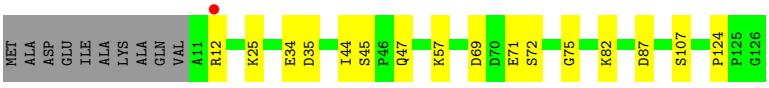
• Molecule 1: Histidine triad nucleotide-binding protein 1

Chain CCC:



• Molecule 1: Histidine triad nucleotide-binding protein 1

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	112.89Å 112.89Å 43.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.73 – 2.27 40.70 – 2.27	Depositor EDS
% Data completeness (in resolution range)	100.0 (40.73-2.27) 100.0 (40.70-2.27)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.164 , 0.220 0.172 , 0.228	Depositor DCC
R_{free} test set	1305 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	38.4	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 33.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3917	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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: P6G, CAC, PG4

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	0.79	0/929	0.88	0/1253
1	BBB	0.80	0/934	0.92	0/1259
1	CCC	0.81	0/915	0.90	0/1234
1	DDD	0.83	0/929	0.91	0/1253
All	All	0.81	0/3707	0.91	0/4999

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	907	0	898	6	0
1	BBB	912	0	912	12	0
1	CCC	893	0	891	6	0
1	DDD	907	0	898	8	0
2	AAA	19	0	26	0	0
3	BBB	5	0	0	0	0
3	CCC	5	0	0	0	0
4	DDD	13	0	18	0	0
5	AAA	69	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	BBB	63	0	0	1	0
5	CCC	64	0	0	1	0
5	DDD	60	0	0	0	0
All	All	3917	0	3643	27	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.

All (27) 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:12[A]:ARG:HG3	1:BBB:13:PRO:HD2	1.67	0.76
1:AAA:91:ASN:HD22	1:AAA:91:ASN:H	1.41	0.69
1:BBB:12[A]:ARG:CG	1:BBB:13:PRO:HD2	2.28	0.63
1:CCC:79:ILE:HG13	1:DDD:71:GLU:HG3	1.87	0.57
1:BBB:40:ALA:HA	1:BBB:53:LEU:O	2.10	0.51
1:AAA:91:ASN:H	1:AAA:91:ASN:ND2	2.09	0.50
1:DDD:25:LYS:NZ	1:DDD:34:GLU:OE1	2.45	0.50
1:BBB:12[B]:ARG:N	1:BBB:12[B]:ARG:HD3	2.27	0.49
1:BBB:12[B]:ARG:NH1	1:BBB:12[B]:ARG:HG3	2.28	0.49
1:BBB:12[B]:ARG:HG3	1:BBB:12[B]:ARG:HH11	1.77	0.48
1:CCC:120:GLN:NE2	5:CCC:302:HOH:O	2.43	0.48
1:CCC:71:GLU:O	1:DDD:75:GLY:HA3	2.15	0.46
1:AAA:55:ILE:HG22	1:AAA:112:HIS:HB3	1.98	0.46
1:AAA:23:ILE:HD13	1:AAA:57:LYS:HA	1.99	0.45
1:BBB:67:GLU:O	1:BBB:70:ASP:HB2	2.17	0.45
1:CCC:75:GLY:HA3	1:DDD:71:GLU:O	2.16	0.45
1:DDD:35:ASP:O	1:DDD:57:LYS:NZ	2.47	0.44
1:AAA:32:ILE:HD11	1:AAA:42:HIS:CE1	2.53	0.44
1:BBB:21:LYS:HB2	1:BBB:27:ILE:HD12	2.00	0.44
1:DDD:69[A]:ASP:OD1	1:DDD:69[A]:ASP:C	2.56	0.43
1:BBB:55:ILE:HG22	1:BBB:112:HIS:HB3	2.00	0.43
1:CCC:47:GLN:OE1	1:DDD:124:PRO:HB3	2.18	0.43
1:BBB:90:LEU:HD13	1:BBB:117:GLY:HA3	2.01	0.42
1:AAA:40:ALA:HA	1:AAA:53:LEU:O	2.19	0.42
1:BBB:79:ILE:O	1:BBB:82:LYS:HG3	2.20	0.41
1:BBB:120:GLN:NE2	5:BBB:302:HOH:O	2.52	0.41
1:CCC:124:PRO:HB3	1:DDD:47:GLN:HB3	2.03	0.40

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	116/126 (92%)	113 (97%)	3 (3%)	0	100	100
1	BBB	116/126 (92%)	114 (98%)	2 (2%)	0	100	100
1	CCC	114/126 (90%)	111 (97%)	3 (3%)	0	100	100
1	DDD	116/126 (92%)	109 (94%)	6 (5%)	1 (1%)	17	18
All	All	462/504 (92%)	447 (97%)	14 (3%)	1 (0%)	47	57

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	DDD	12	ARG

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	97/102 (95%)	93 (96%)	4 (4%)	30	41
1	BBB	97/102 (95%)	95 (98%)	2 (2%)	53	68
1	CCC	95/102 (93%)	94 (99%)	1 (1%)	73	84
1	DDD	97/102 (95%)	90 (93%)	7 (7%)	14	17
All	All	386/408 (95%)	372 (96%)	14 (4%)	37	47

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	16	ASP
1	AAA	72	SER
1	AAA	82	LYS
1	AAA	91	ASN
1	BBB	30	LYS
1	BBB	82	LYS
1	CCC	82	LYS
1	DDD	44	ILE
1	DDD	45	SER
1	DDD	72	SER
1	DDD	82	LYS
1	DDD	87	ASP
1	DDD	107[A]	SER
1	DDD	107[B]	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)

4 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	P6G	AAA	201	-	18,18,18	0.45	0	17,17,17	0.30	0
4	PG4	DDD	201	-	12,12,12	0.34	0	11,11,11	0.16	0
3	CAC	CCC	201	-	0,4,4	-	-	0,6,6	-	-
3	CAC	BBB	201	-	0,4,4	-	-	0,6,6	-	-

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	P6G	AAA	201	-	-	12/16/16/16	-
4	PG4	DDD	201	-	-	7/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

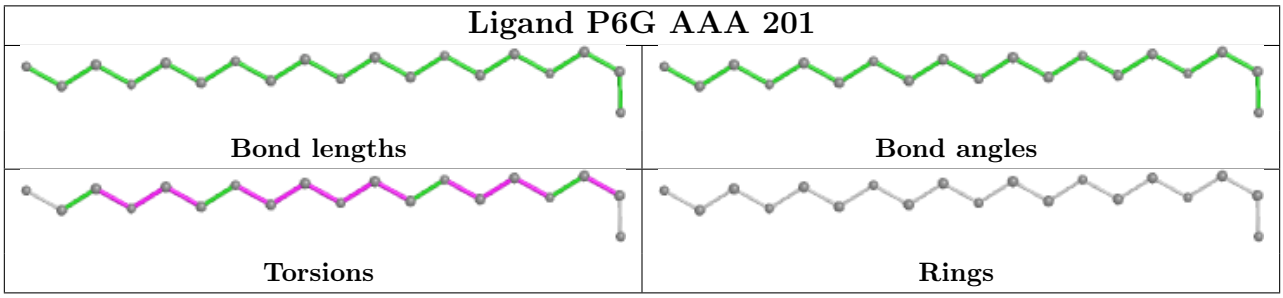
All (19) torsion outliers are listed below:

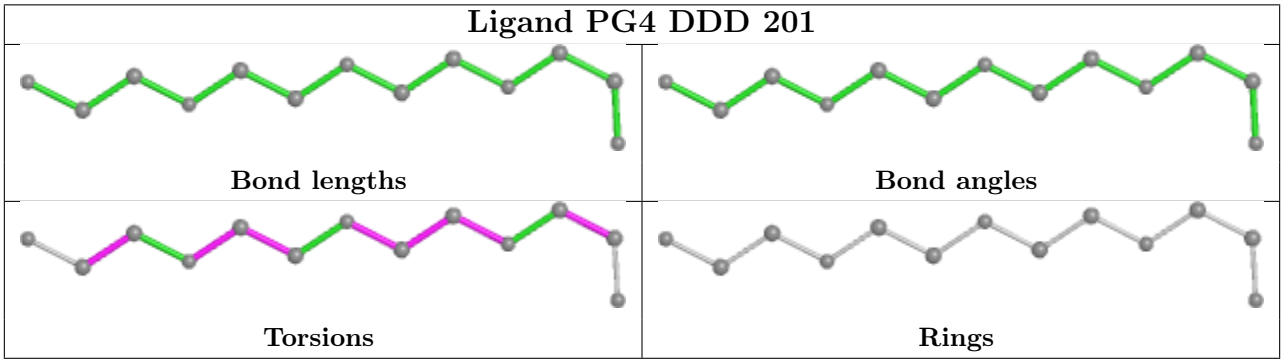
Mol	Chain	Res	Type	Atoms
4	DDD	201	PG4	O2-C3-C4-O3
4	DDD	201	PG4	O1-C1-C2-O2
4	DDD	201	PG4	O4-C7-C8-O5
2	AAA	201	P6G	O10-C11-C12-O13
2	AAA	201	P6G	C5-C6-O7-C8
4	DDD	201	PG4	C5-C6-O4-C7
2	AAA	201	P6G	C8-C9-O10-C11
2	AAA	201	P6G	O7-C8-C9-O10
2	AAA	201	P6G	C11-C12-O13-C14
2	AAA	201	P6G	C18-C17-O16-C15
4	DDD	201	PG4	O3-C5-C6-O4
2	AAA	201	P6G	O1-C2-C3-O4
4	DDD	201	PG4	C4-C3-O2-C2
4	DDD	201	PG4	C3-C4-O3-C5
2	AAA	201	P6G	C14-C15-O16-C17
2	AAA	201	P6G	C6-C5-O4-C3
2	AAA	201	P6G	O13-C14-C15-O16
2	AAA	201	P6G	C12-C11-O10-C9
2	AAA	201	P6G	O4-C5-C6-O7

There are no ring outliers.

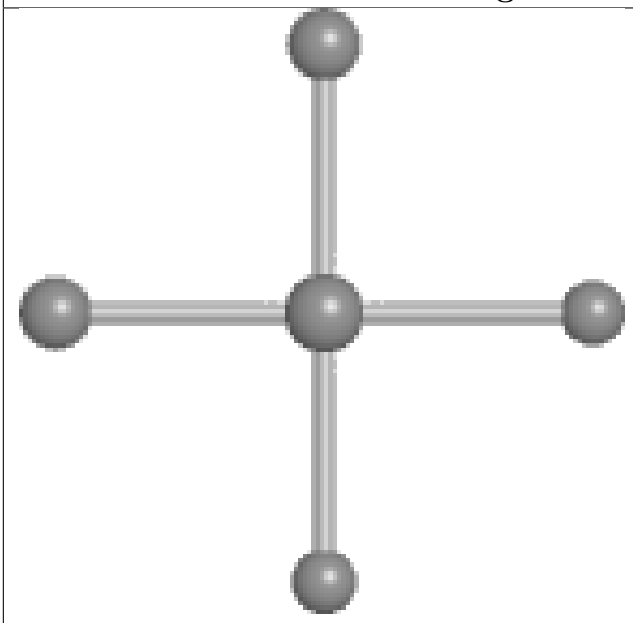
No monomer is involved in short contacts.

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.

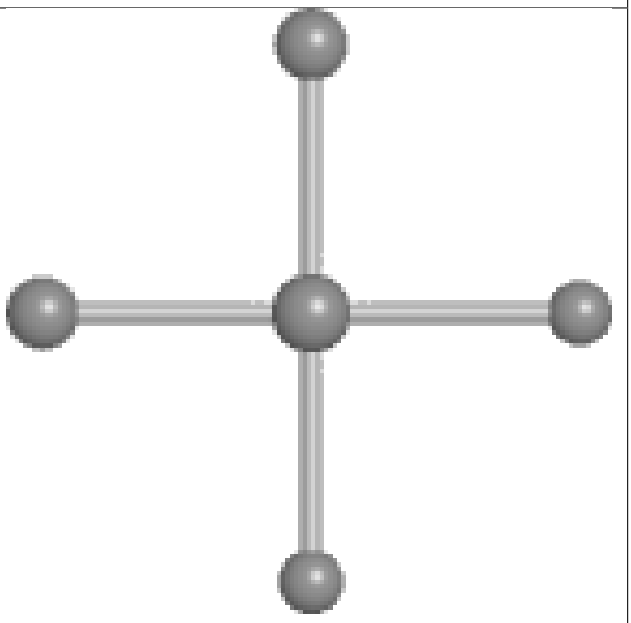




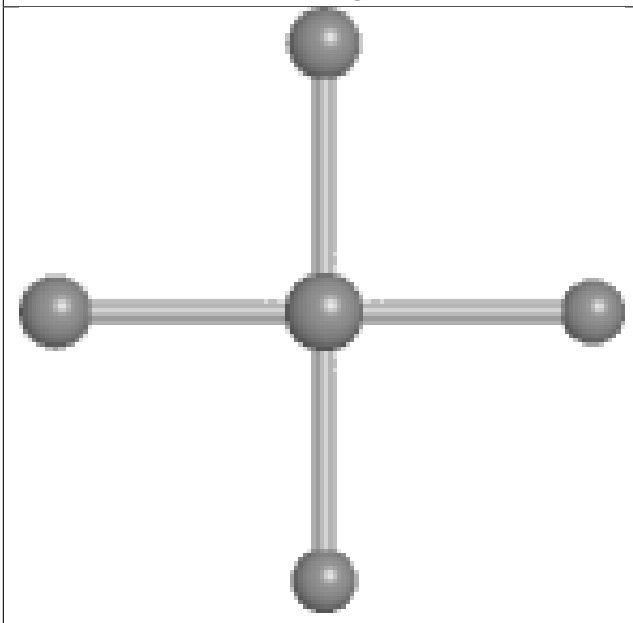
Ligand CAC CCC 201



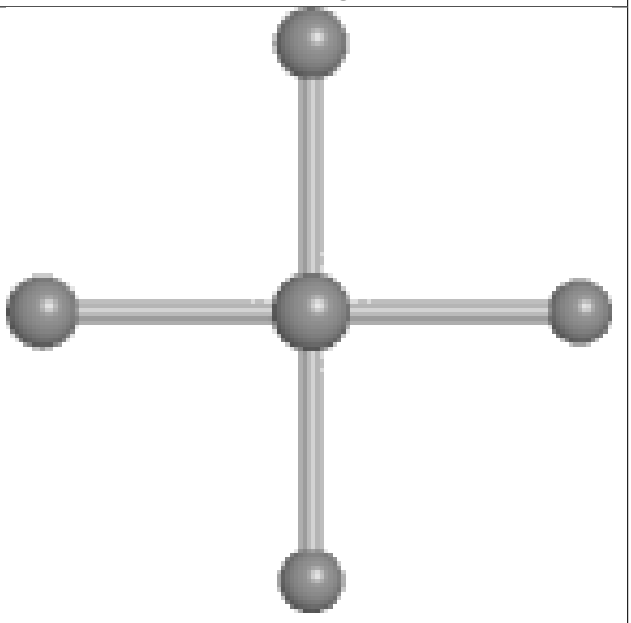
Bond lengths



Bond angles

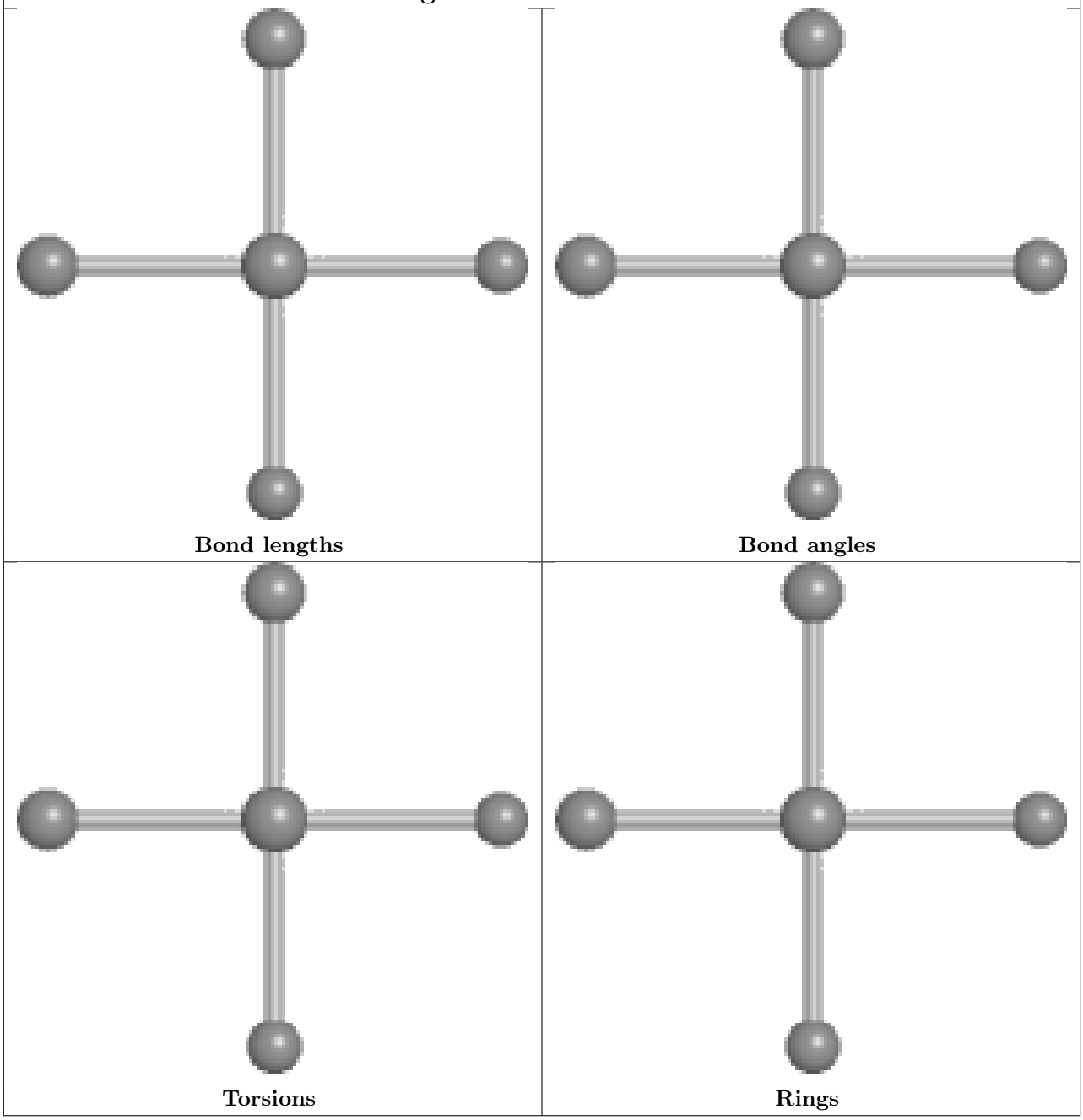


Torsions



Rings

Ligand CAC BBB 201



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	116/126 (92%)	-0.38	2 (1%)	70	75	30, 40, 63, 130	0
1	BBB	116/126 (92%)	-0.50	3 (2%)	56	62	29, 41, 70, 131	0
1	CCC	116/126 (92%)	-0.29	5 (4%)	35	40	29, 42, 67, 160	0
1	DDD	116/126 (92%)	-0.51	1 (0%)	84	87	32, 45, 76, 123	0
All	All	464/504 (92%)	-0.42	11 (2%)	59	65	29, 42, 71, 160	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	11	ALA	14.3
1	BBB	11	ALA	9.7
1	AAA	11	ALA	9.2
1	CCC	13	PRO	3.9
1	AAA	12	ARG	3.9
1	BBB	12[A]	ARG	3.9
1	CCC	12	ARG	3.4
1	CCC	14	GLY	3.2
1	DDD	12	ARG	2.7
1	BBB	13	PRO	2.4
1	CCC	63	ILE	2.1

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 (i)

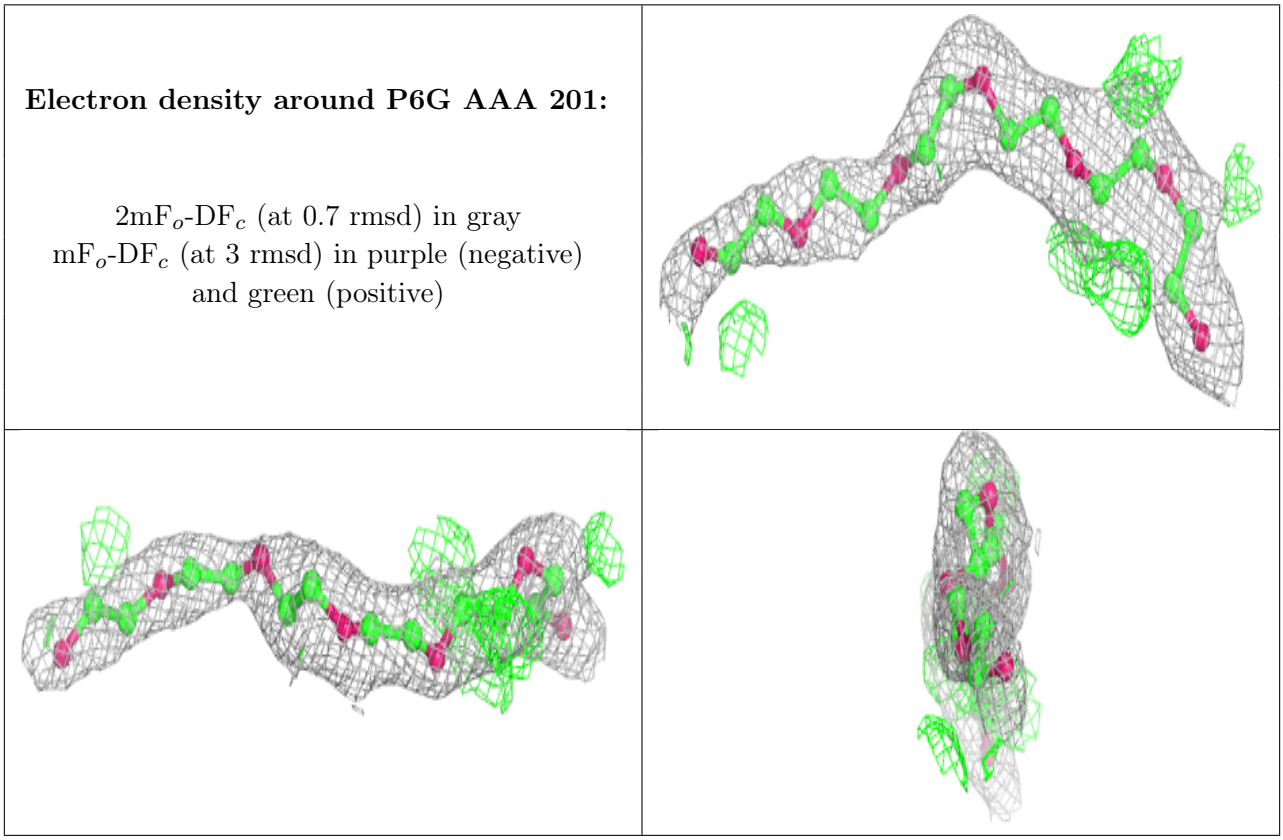
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	P6G	AAA	201	19/19	0.86	0.16	59,70,80,81	0
4	PG4	DDD	201	13/13	0.87	0.22	72,80,102,105	0
3	CAC	CCC	201	5/5	0.98	0.11	46,59,61,63	0
3	CAC	BBB	201	5/5	0.99	0.11	58,59,69,70	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.

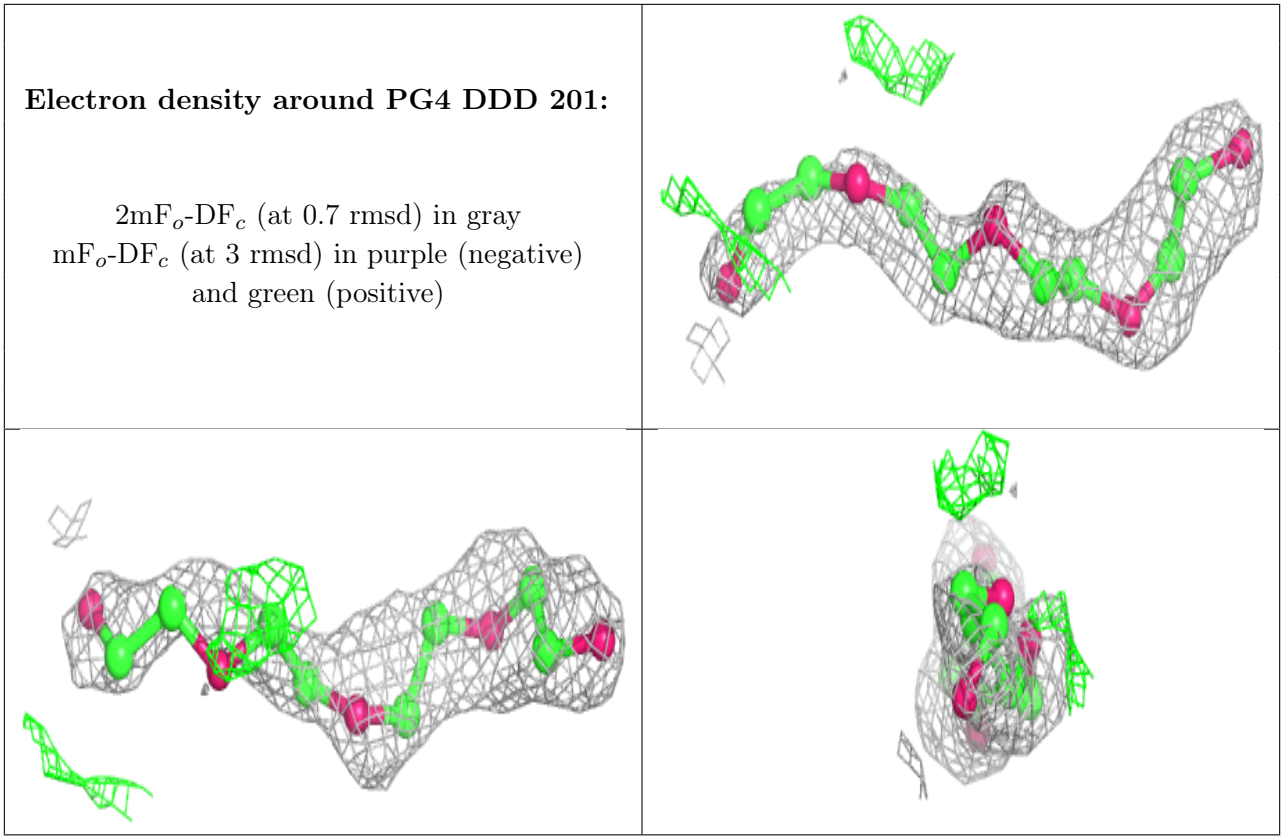
Electron density around P6G AAA 201:

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



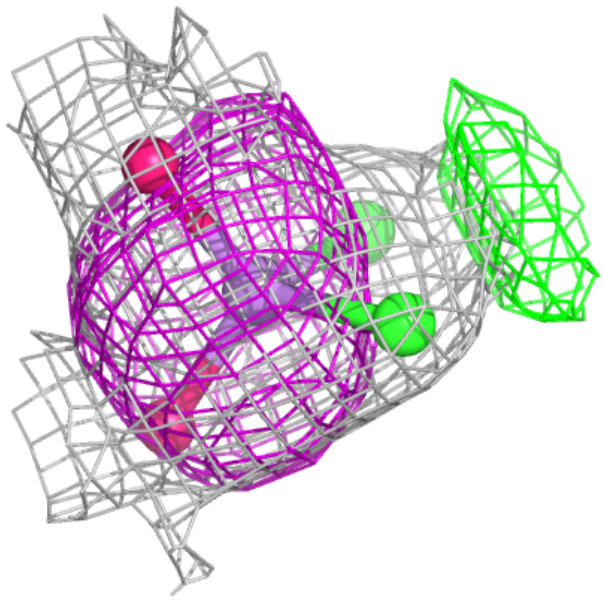
Electron density around PG4 DDD 201:

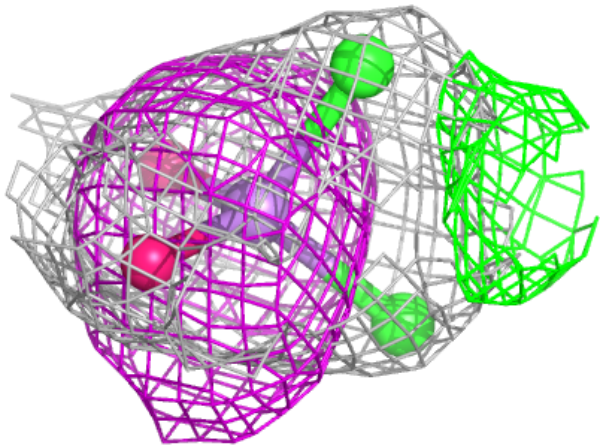
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

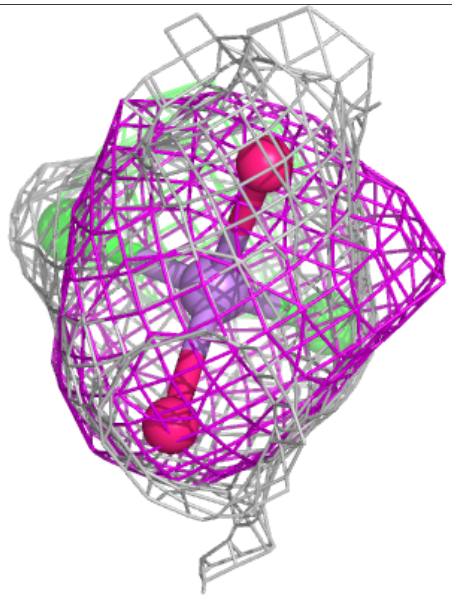


Electron density around CAC CCC 201:

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

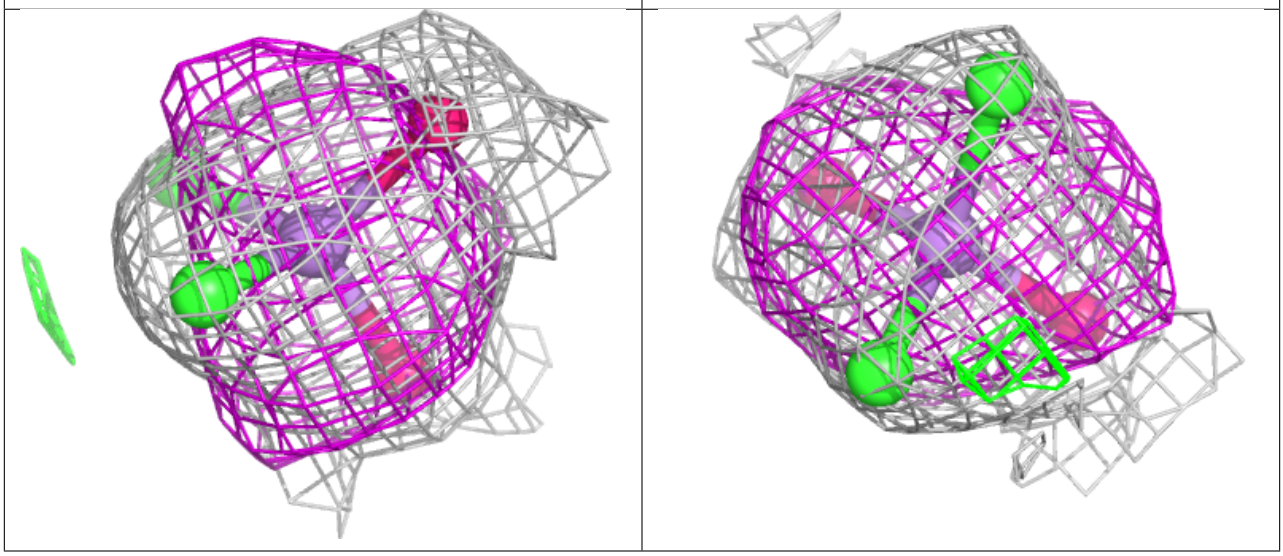






Electron density around CAC BBB 201:

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



6.5 Other polymers

There are no such residues in this entry.