W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 20, 2024 – 04:42 pm GMT

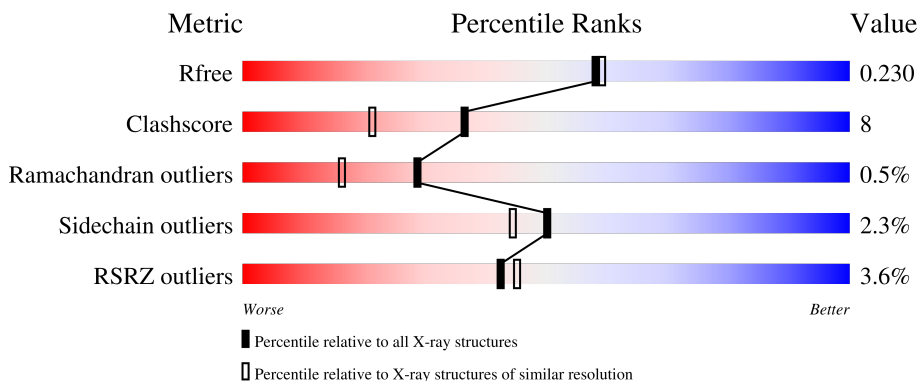|		
PDB ID	:	7NWN
Title	:	A carbohydrate binding module family 9 (CBM9) from Caldicellulsiruptor kristjanssonii
Authors	:	Krska, D.; Mazurkewich, S.; Navarro Poulsen, J.; Larsbrink, J.; Lo Leggio, L.
Deposited on	:	2021-03-17
Resolution	:	1.97 Å(reported)

---

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.97 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		0.230
Clashscore		8
Ramachandran outliers		0.5%
Sidechain outliers		2.3%
RSRZ outliers		3.6%

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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	215	3% / 82% / 8% / 9%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PEG	AAA	303	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	EDO	AAA	315	-	X	-	-

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 1853 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 Beta-xylanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	195	1616	1021	276	316	3	0	5	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP E4S6E9
AAA	2	GLY	-	expression tag	UNP E4S6E9
AAA	3	SER	-	expression tag	UNP E4S6E9
AAA	4	SER	-	expression tag	UNP E4S6E9
AAA	5	HIS	-	expression tag	UNP E4S6E9
AAA	6	HIS	-	expression tag	UNP E4S6E9
AAA	7	HIS	-	expression tag	UNP E4S6E9
AAA	8	HIS	-	expression tag	UNP E4S6E9
AAA	9	HIS	-	expression tag	UNP E4S6E9
AAA	10	HIS	-	expression tag	UNP E4S6E9
AAA	11	SER	-	expression tag	UNP E4S6E9
AAA	12	SER	-	expression tag	UNP E4S6E9
AAA	13	GLU	-	expression tag	UNP E4S6E9
AAA	14	ASN	-	expression tag	UNP E4S6E9
AAA	15	LEU	-	expression tag	UNP E4S6E9
AAA	16	TYR	-	expression tag	UNP E4S6E9
AAA	17	PHE	-	expression tag	UNP E4S6E9
AAA	18	GLN	-	expression tag	UNP E4S6E9
AAA	19	GLY	-	expression tag	UNP E4S6E9
AAA	20	HIS	-	expression tag	UNP E4S6E9

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total 7	C 4	O 3	0	0
2	AAA	1	Total 7	C 4	O 3	0	0
2	AAA	1	Total 7	C 4	O 3	0	0
2	AAA	1	Total 7	C 4	O 3	0	0
2	AAA	1	Total 7	C 4	O 3	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	3	Total 3	Ca 3	0	0

- Molecule 5 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$).
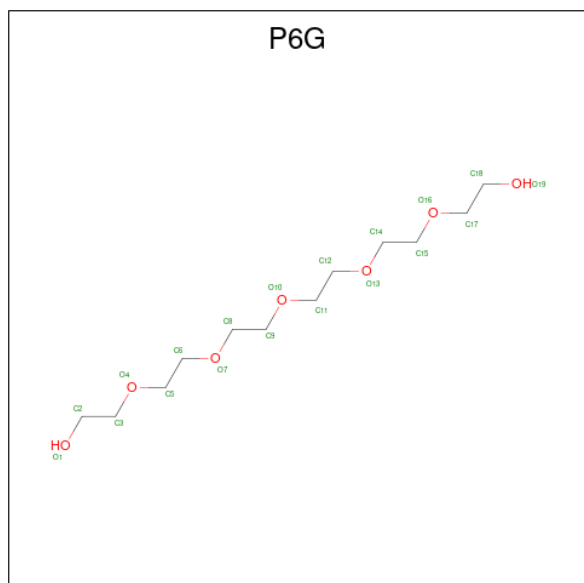
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total 19	C 12	O 7	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	140	Total 140	O 140	0	0

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: Beta-xylanase

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 3 2	Depositor
Cell constants a, b, c, α, β, γ	172.67Å 172.67Å 172.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.19 – 1.97	Depositor
	46.15 – 1.97	EDS
% Data completeness	100.0 (46.19-1.97)	Depositor
(in resolution range)	100.0 (46.15-1.97)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.180 , 0.230	Depositor
	0.180 , 0.230	DCC
R_{free} test set	1521 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	41.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 51.7	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	1853	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 3.77% 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, CA, EDO, PEG

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	0.54	0/1653	0.89	0/2247

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	1616	0	1551	19	0
2	AAA	35	0	50	5	0
3	AAA	40	0	60	6	0
4	AAA	3	0	0	0	0
5	AAA	19	0	26	2	0
6	AAA	140	0	0	10	0
All	All	1853	0	1687	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 8.

The worst 5 of 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:AAA:206:SER:HB2	2:AAA:303:PEG:H21	1.43	1.01
3:AAA:314:EDO:H21	6:AAA:513:HOH:O	1.82	0.79
1:AAA:52:SER:HA	2:AAA:304:PEG:H21	1.64	0.78
1:AAA:206:SER:CB	2:AAA:303:PEG:H21	2.18	0.73
1:AAA:60[A]:LYS:NZ	6:AAA:401:HOH:O	2.23	0.72

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	198/215 (92%)	191 (96%)	6 (3%)	1 (0%)	29	16

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	210	VAL

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	177/190 (93%)	173 (98%)	4 (2%)	50	44

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	37	ASP
1	AAA	176	ARG
1	AAA	208	ASN
1	AAA	210	VAL

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 19 ligands modelled in this entry, 3 are monoatomic - leaving 16 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	EDO	AAA	308	-	3,3,3	0.74	0	2,2,2	1.44	0
3	EDO	AAA	311	-	3,3,3	0.30	0	2,2,2	0.46	0
2	PEG	AAA	304	-	6,6,6	0.36	0	5,5,5	0.41	0
3	EDO	AAA	309	-	3,3,3	0.07	0	2,2,2	0.15	0
5	P6G	AAA	319	-	18,18,18	0.97	1 (5%)	17,17,17	0.50	0
3	EDO	AAA	306	-	3,3,3	0.14	0	2,2,2	0.21	0
3	EDO	AAA	310	-	3,3,3	0.41	0	2,2,2	0.44	0

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	PEG	AAA	303	-	6,6,6	0.25	0	5,5,5	0.17	0
3	EDO	AAA	315	-	3,3,3	3.01	3 (100%)	2,2,2	2.06	1 (50%)
3	EDO	AAA	312	-	3,3,3	0.10	0	2,2,2	0.10	0
3	EDO	AAA	313	-	3,3,3	0.32	0	2,2,2	0.42	0
2	PEG	AAA	301	-	6,6,6	0.16	0	5,5,5	0.11	0
3	EDO	AAA	307	-	3,3,3	0.16	0	2,2,2	0.33	0
3	EDO	AAA	314	-	3,3,3	0.21	0	2,2,2	0.46	0
2	PEG	AAA	305	-	6,6,6	0.28	0	5,5,5	0.15	0
2	PEG	AAA	302	-	6,6,6	0.11	0	5,5,5	0.20	0

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	EDO	AAA	308	-	-	0/1/1/1	-
3	EDO	AAA	311	-	-	1/1/1/1	-
2	PEG	AAA	304	-	-	4/4/4/4	-
3	EDO	AAA	309	-	-	1/1/1/1	-
5	P6G	AAA	319	-	-	12/16/16/16	-
3	EDO	AAA	306	-	-	0/1/1/1	-
3	EDO	AAA	310	-	-	1/1/1/1	-
2	PEG	AAA	303	-	-	4/4/4/4	-
3	EDO	AAA	315	-	-	1/1/1/1	-
3	EDO	AAA	312	-	-	1/1/1/1	-
3	EDO	AAA	313	-	-	1/1/1/1	-
2	PEG	AAA	301	-	-	3/4/4/4	-
3	EDO	AAA	307	-	-	1/1/1/1	-
3	EDO	AAA	314	-	-	0/1/1/1	-
2	PEG	AAA	305	-	-	2/4/4/4	-
2	PEG	AAA	302	-	-	3/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	315	EDO	O2-C2	3.59	1.60	1.42
3	AAA	315	EDO	O1-C1	3.05	1.57	1.42
5	AAA	319	P6G	O19-C18	2.48	1.54	1.42
3	AAA	315	EDO	C2-C1	2.26	1.63	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	315	EDO	O1-C1-C2	-2.13	96.61	111.91

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	AAA	319	P6G	C11-C12-O13-C14
2	AAA	304	PEG	O2-C3-C4-O4
5	AAA	319	P6G	O4-C5-C6-O7
5	AAA	319	P6G	C12-C11-O10-C9
2	AAA	303	PEG	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	304	PEG	1	0
5	AAA	319	P6G	2	0
3	AAA	306	EDO	1	0
3	AAA	310	EDO	2	0
2	AAA	303	PEG	4	0
3	AAA	307	EDO	1	0
3	AAA	314	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 P6G AAA 319	
Bond lengths
Bond angles
Torsions |
Rings |
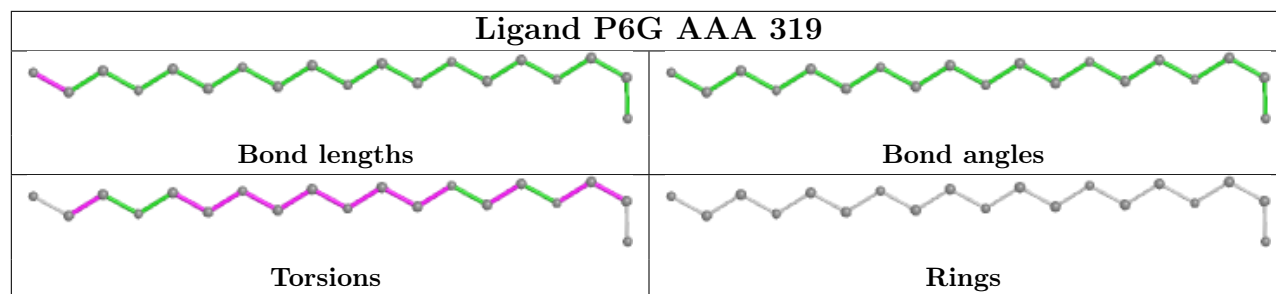
5.7 Other polymers ⓘ

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 (i)

6.1 Protein, DNA and RNA chains (i)

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	195/215 (90%)	0.14	7 (3%)	42	45	29, 40, 63, 163	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	211	VAL	9.7
1	AAA	210	VAL	6.9
1	AAA	213	ILE	6.9
1	AAA	209	ALA	6.8
1	AAA	212	PRO	6.1

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

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

6.3 Carbohydrates (i)

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, 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.
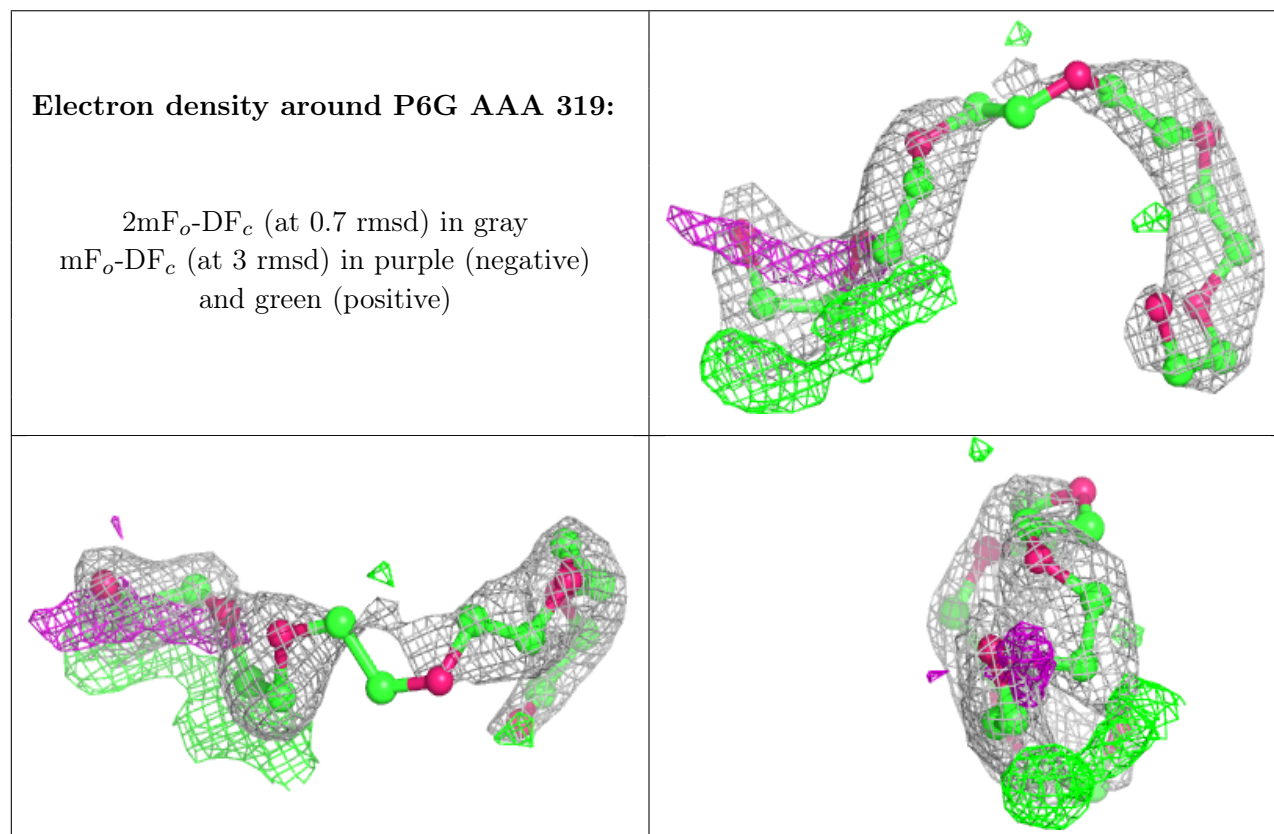
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	PEG	AAA	305	7/7	0.50	0.24	86,98,113,120	0
3	EDO	AAA	313	4/4	0.65	0.33	104,108,112,118	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EDO	AAA	309	4/4	0.72	0.17	81,83,83,91	0
2	PEG	AAA	304	7/7	0.74	0.28	69,75,98,100	0
2	PEG	AAA	301	7/7	0.74	0.21	74,81,93,94	0
3	EDO	AAA	311	4/4	0.78	0.21	78,79,79,80	0
2	PEG	AAA	303	7/7	0.79	0.29	58,63,83,87	0
3	EDO	AAA	308	4/4	0.80	0.17	60,73,82,85	0
3	EDO	AAA	314	4/4	0.80	0.31	58,72,73,87	0
5	P6G	AAA	319	19/19	0.81	0.22	39,65,86,87	0
3	EDO	AAA	306	4/4	0.82	0.25	68,68,75,75	0
3	EDO	AAA	312	4/4	0.83	0.15	80,87,87,95	0
3	EDO	AAA	307	4/4	0.85	0.16	71,72,72,81	0
2	PEG	AAA	302	7/7	0.85	0.18	63,77,80,84	0
3	EDO	AAA	315	4/4	0.90	0.26	31,33,36,37	0
3	EDO	AAA	310	4/4	0.94	0.28	51,51,57,59	0
4	CA	AAA	316	1/1	0.97	0.04	50,50,50,50	0
4	CA	AAA	318	1/1	0.98	0.20	51,51,51,51	1
4	CA	AAA	317	1/1	0.99	0.09	29,29,29,29	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 319:

$2mF_o\text{-}DF_c$ (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.