



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2023 – 04:25 pm BST

PDB ID : 8C34
Title : Aerobic light exposed 1.8 Angstrom crystal structure of cobalamin binding domain belonging to a light-dependent transcription regulator TtCarH
Authors : Poddar, H.; Leys, D.
Deposited on : 2022-12-23
Resolution : 1.82 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 : 2.35
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) : 2.35

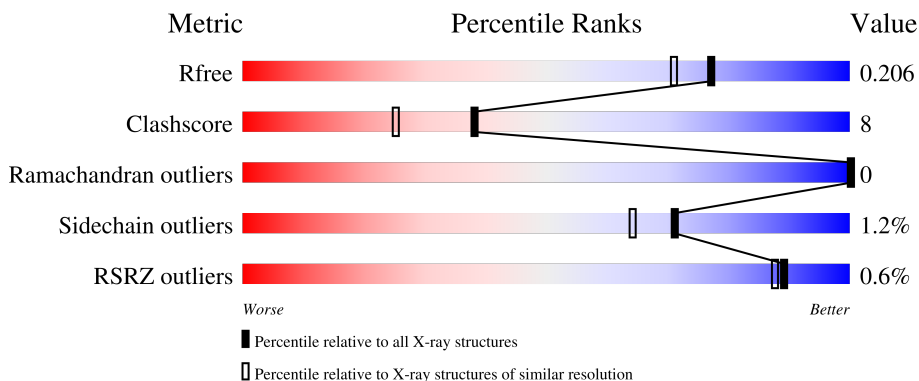
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.82 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	215	 80% 11% 8%
1	B	215	 86% 8% 7%
1	C	215	 83% 7% 8%
1	D	215	 83% 8% 7%

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	B12	C	301	X	-	-	-
2	B12	D	301	X	-	-	-

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 7119 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 Probable transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	197	1516	971	279	264	2	0	4	0
1	B	201	1539	988	281	268	2	0	2	0
1	C	197	1491	953	271	265	2	0	1	0
1	D	200	1533	983	281	267	2	0	2	0

There are 28 discrepancies between the modelled and reference sequences:

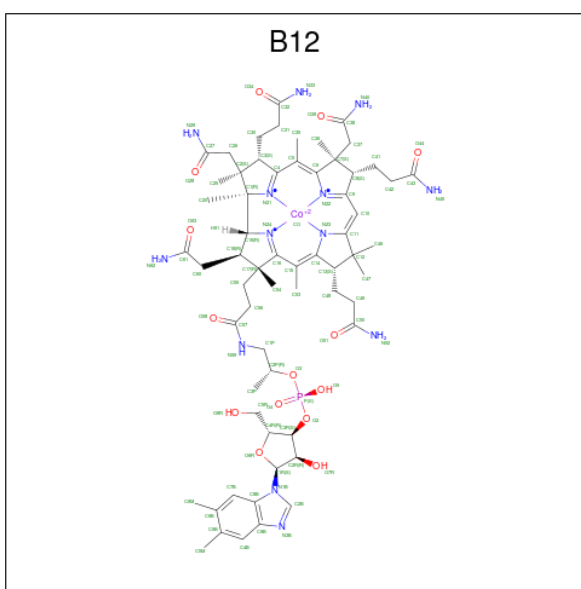
Chain	Residue	Modelled	Actual	Comment	Reference
A	286	GLY	-	expression tag	UNP Q746J7
A	287	HIS	-	expression tag	UNP Q746J7
A	288	HIS	-	expression tag	UNP Q746J7
A	289	HIS	-	expression tag	UNP Q746J7
A	290	HIS	-	expression tag	UNP Q746J7
A	291	HIS	-	expression tag	UNP Q746J7
A	292	HIS	-	expression tag	UNP Q746J7
B	286	GLY	-	expression tag	UNP Q746J7
B	287	HIS	-	expression tag	UNP Q746J7
B	288	HIS	-	expression tag	UNP Q746J7
B	289	HIS	-	expression tag	UNP Q746J7
B	290	HIS	-	expression tag	UNP Q746J7
B	291	HIS	-	expression tag	UNP Q746J7
B	292	HIS	-	expression tag	UNP Q746J7
C	286	GLY	-	expression tag	UNP Q746J7
C	287	HIS	-	expression tag	UNP Q746J7
C	288	HIS	-	expression tag	UNP Q746J7
C	289	HIS	-	expression tag	UNP Q746J7
C	290	HIS	-	expression tag	UNP Q746J7
C	291	HIS	-	expression tag	UNP Q746J7
C	292	HIS	-	expression tag	UNP Q746J7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	286	GLY	-	expression tag	UNP Q746J7
D	287	HIS	-	expression tag	UNP Q746J7
D	288	HIS	-	expression tag	UNP Q746J7
D	289	HIS	-	expression tag	UNP Q746J7
D	290	HIS	-	expression tag	UNP Q746J7
D	291	HIS	-	expression tag	UNP Q746J7
D	292	HIS	-	expression tag	UNP Q746J7

- Molecule 2 is COBALAMIN (three-letter code: B12) (formula: $C_{62}H_{89}CoN_{13}O_{14}P$) (labeled as "Ligand of Interest" by depositor).



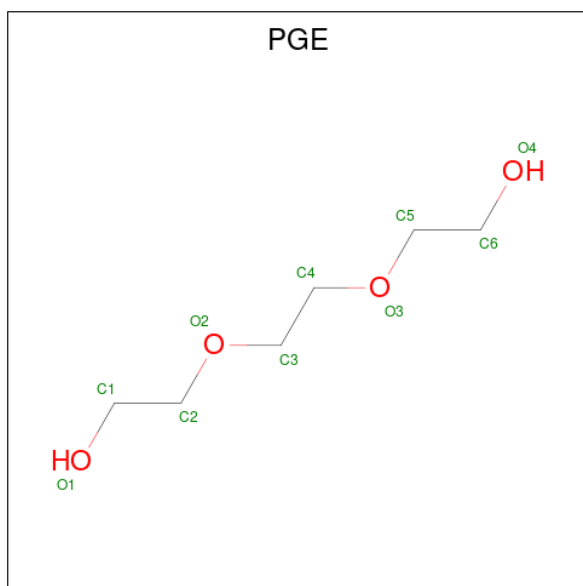
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	Co	N	O			P
2	A	1	Total 91	C 62	Co 1	N 13	O 14	P 1	0	0
2	B	1	Total 91	C 62	Co 1	N 13	O 14	P 1	0	0
2	C	1	Total 91	C 62	Co 1	N 13	O 14	P 1	0	0
2	D	1	Total 91	C 62	Co 1	N 13	O 14	P 1	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C₆H₁₄O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	174	Total 174	O 174	0	0
5	B	195	Total 195	O 195	0	0
5	C	135	Total 135	O 135	0	0
5	D	155	Total 155	O 155	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	60.87Å 72.69Å 201.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	52.10 – 1.82 52.10 – 1.82	Depositor EDS
% Data completeness (in resolution range)	100.0 (52.10-1.82) 99.9 (52.10-1.82)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.46 (at 1.82Å)	Xtrriage
Refinement program	REFMAC REFMAC 5.8.0352	Depositor
R, R_{free}	0.166 , 0.198 0.175 , 0.206	Depositor DCC
R_{free} test set	4038 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	33.7	Xtrriage
Anisotropy	0.186	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 42.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	7119	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.21% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: B12, PEG, PGE

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.53	0/1563	0.76	0/2122
1	B	0.52	0/1582	0.76	0/2152
1	C	0.51	0/1527	0.71	0/2075
1	D	0.52	0/1575	0.81	0/2141
All	All	0.52	0/6247	0.76	0/8490

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
1	C	0	5
1	D	0	6
All	All	0	14

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	93	ARG	Sidechain
1	B	105	ARG	Sidechain
1	B	212	ARG	Sidechain
1	C	79	ARG	Sidechain
1	C	93	ARG	Sidechain

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1516	0	1565	22	0
1	B	1539	0	1582	19	0
1	C	1491	0	1530	13	0
1	D	1533	0	1579	16	0
2	A	91	0	87	11	0
2	B	91	0	87	11	0
2	C	91	0	87	19	0
2	D	91	0	87	15	0
3	A	7	0	10	2	0
4	B	10	0	14	0	0
5	A	174	0	0	4	0
5	B	195	0	0	2	0
5	C	135	0	0	3	0
5	D	155	0	0	2	0
All	All	7119	0	6628	100	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 100 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)
2:D:301:B12:H362	2:D:301:B12:H351	1.35	1.08
2:A:301:B12:H362	2:A:301:B12:H351	1.45	0.97
2:B:301:B12:H362	2:B:301:B12:H351	1.49	0.92
2:A:301:B12:H531	2:A:301:B12:H552	1.57	0.87
2:C:301:B12:H521	2:C:301:B12:H471	1.48	0.77

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	A	199/215 (93%)	196 (98%)	3 (2%)	0	100	100
1	B	201/215 (94%)	197 (98%)	4 (2%)	0	100	100
1	C	196/215 (91%)	192 (98%)	4 (2%)	0	100	100
1	D	200/215 (93%)	196 (98%)	4 (2%)	0	100	100
All	All	796/860 (93%)	781 (98%)	15 (2%)	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	A	150/161 (93%)	149 (99%)	1 (1%)	84	80
1	B	152/161 (94%)	151 (99%)	1 (1%)	84	80
1	C	147/161 (91%)	144 (98%)	3 (2%)	55	43
1	D	151/161 (94%)	149 (99%)	2 (1%)	69	61
All	All	600/644 (93%)	593 (99%)	7 (1%)	71	64

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	142	HIS
1	C	196	LEU
1	D	264	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	176	ARG
1	C	79	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	249	GLN
1	C	132	HIS
1	D	249	GLN
1	D	142	HIS
1	B	142	HIS

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](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	PEG	A	302	-	6,6,6	0.23	0	5,5,5	0.26	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	B12	D	301	1	90,101,101	1.36	10 (11%)	137,166,166	1.85	23 (16%)
2	B12	B	301	-	90,101,101	1.32	10 (11%)	137,166,166	1.81	18 (13%)
2	B12	C	301	-	90,101,101	1.43	11 (12%)	137,166,166	2.04	26 (18%)
4	PGE	B	302	-	9,9,9	0.12	0	8,8,8	0.13	0
2	B12	A	301	-	90,101,101	1.29	8 (8%)	137,166,166	1.73	19 (13%)

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	PEG	A	302	-	-	1/4/4/4	-
2	B12	B	301	-	-	11/52/223/223	0/3/11/11
2	B12	D	301	1	1/1/36/38	11/52/223/223	0/3/11/11
2	B12	C	301	-	1/1/36/38	17/52/223/223	0/3/11/11
4	PGE	B	302	-	-	5/7/7/7	-
2	B12	A	301	-	-	10/52/223/223	0/3/11/11

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	B12	C14-N23	-6.50	1.26	1.35
2	D	301	B12	C19-N24	-4.80	1.39	1.48
2	A	301	B12	C8B-C9B	4.59	1.49	1.40
2	A	301	B12	C14-N23	-4.42	1.29	1.35
2	D	301	B12	C14-N23	-4.39	1.29	1.35

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	301	B12	C1-C19-N24	9.79	117.25	106.24
2	B	301	B12	C20-C1-C19	-9.67	100.03	109.36
2	A	301	B12	C1-C19-N24	8.71	116.03	106.24
2	C	301	B12	C20-C1-C19	-8.47	101.19	109.36
2	D	301	B12	C1-C19-N24	8.16	115.42	106.24

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	301	B12	C19
2	D	301	B12	C19

5 of 55 torsion outliers are listed below:

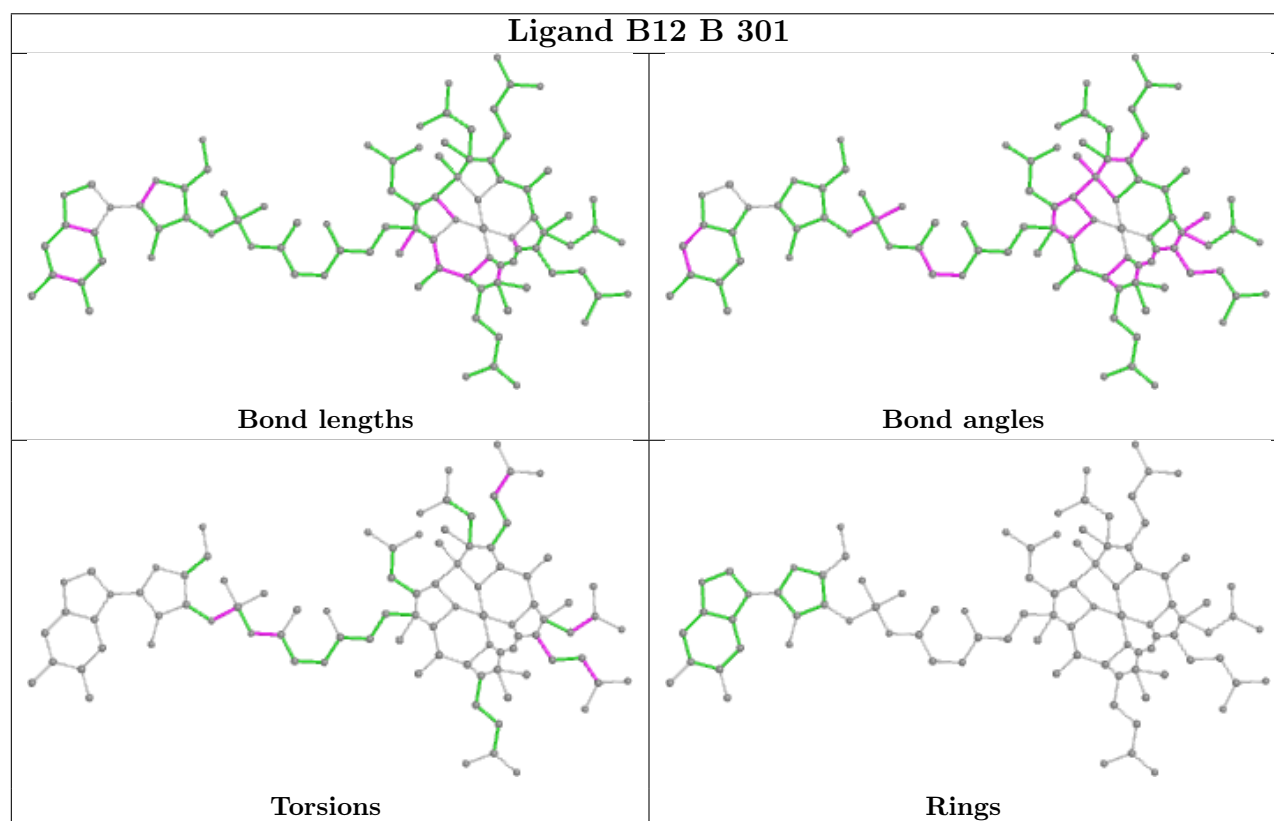
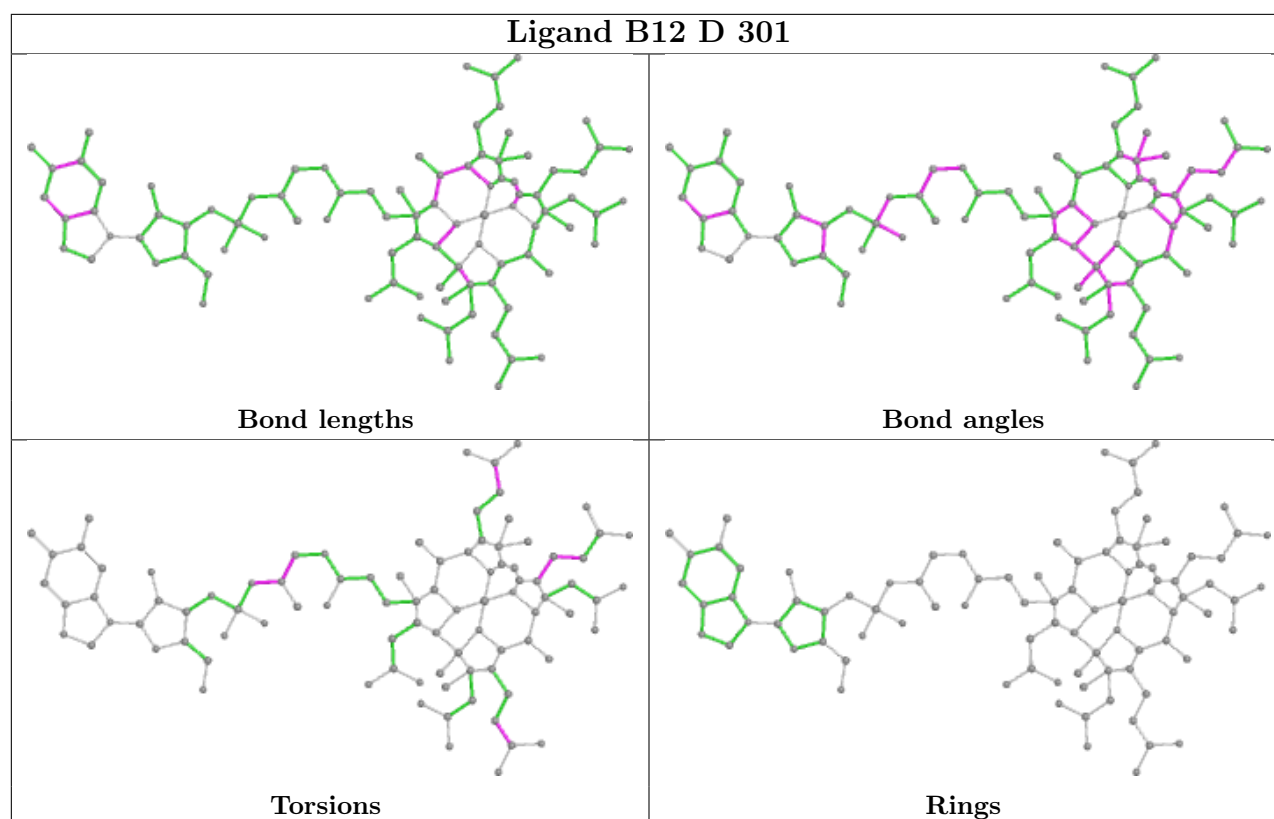
Mol	Chain	Res	Type	Atoms
2	A	301	B12	C42-C41-C8-C7
2	A	301	B12	C42-C41-C8-C9
2	C	301	B12	C38-C37-C7-C6
2	C	301	B12	C7-C37-C38-N40
2	C	301	B12	C42-C41-C8-C9

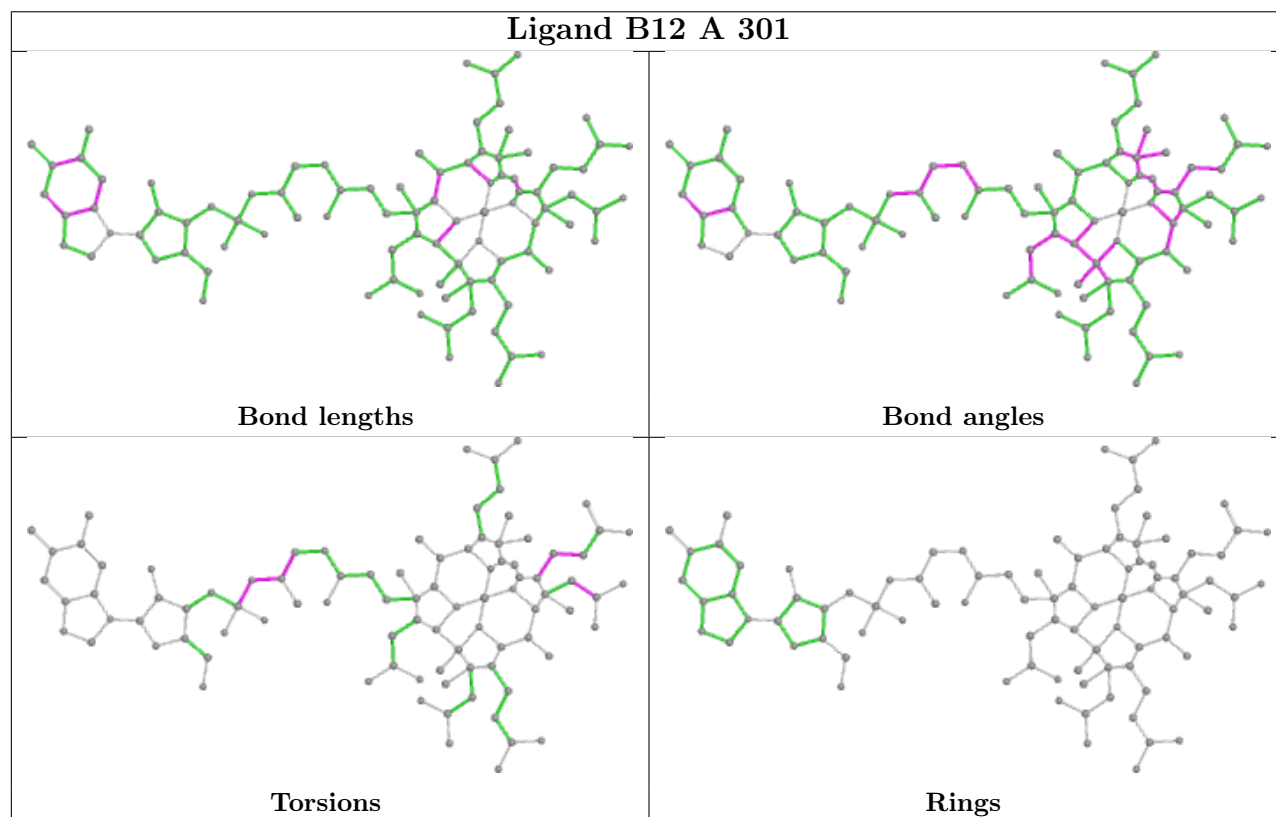
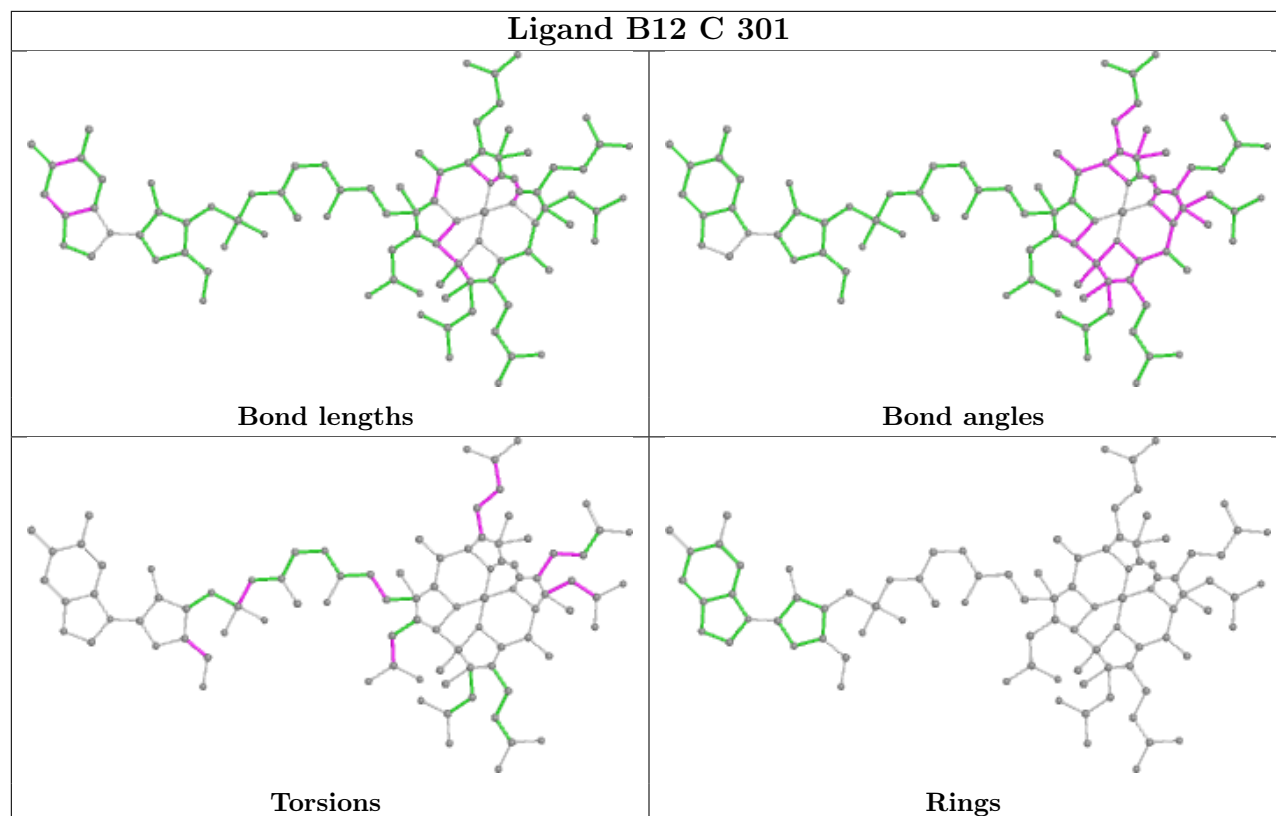
There are no ring outliers.

5 monomers are involved in 58 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	302	PEG	2	0
2	D	301	B12	15	0
2	B	301	B12	11	0
2	C	301	B12	19	0
2	A	301	B12	11	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 than 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.





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 [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	A	197/215 (91%)	-0.17	3 (1%) 73 70	25, 32, 59, 86	0
1	B	201/215 (93%)	-0.29	0 100 100	23, 33, 50, 60	0
1	C	197/215 (91%)	-0.02	2 (1%) 82 80	27, 37, 61, 80	0
1	D	200/215 (93%)	-0.21	0 100 100	25, 35, 53, 63	0
All	All	795/860 (92%)	-0.17	5 (0%) 89 88	23, 34, 56, 86	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	78	VAL	2.9
1	C	239	ASP	2.6
1	A	108[A]	ARG	2.1
1	A	274	LEU	2.1
1	C	268	LYS	2.1

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.

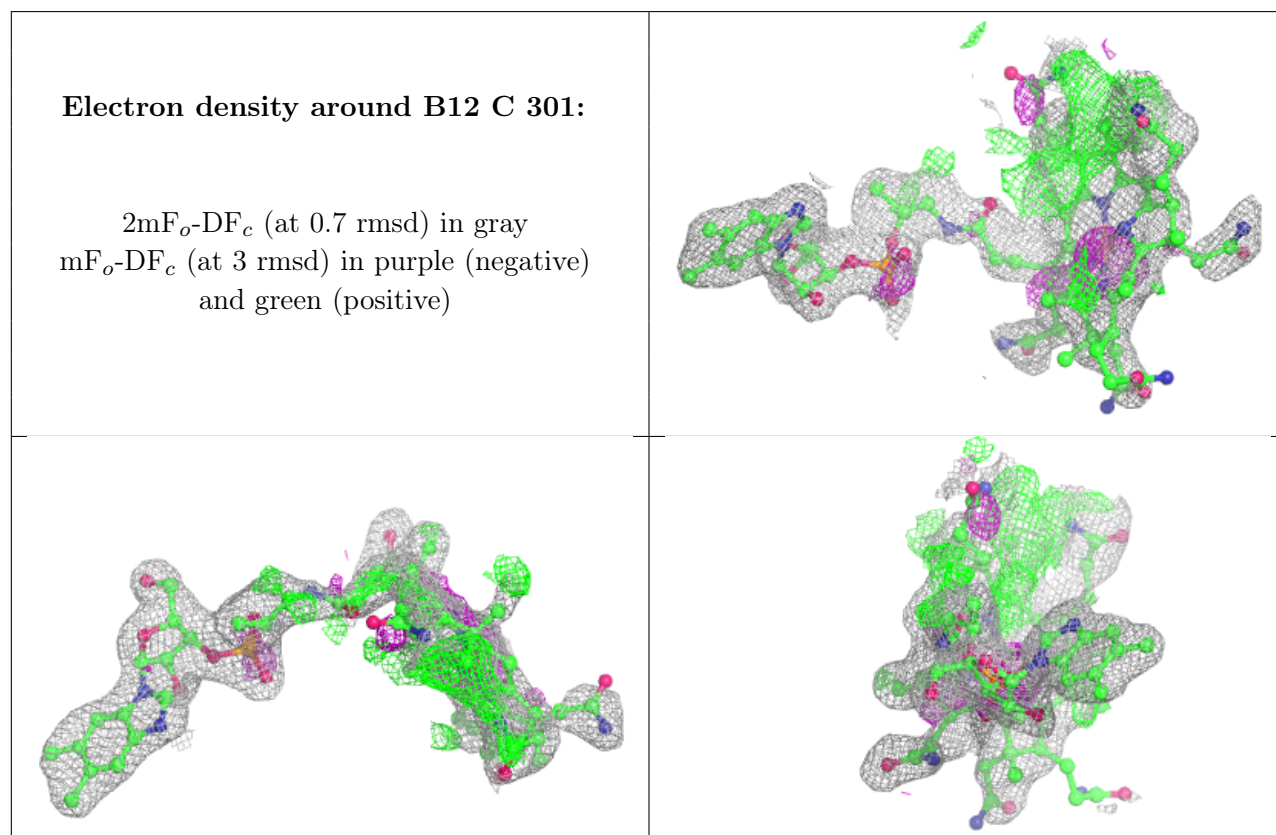
6.4 Ligands [i](#)

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.

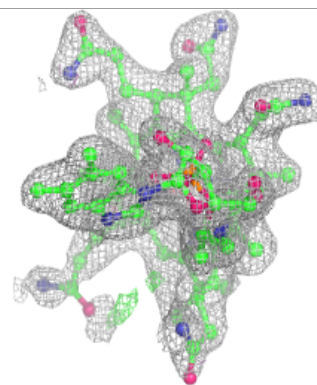
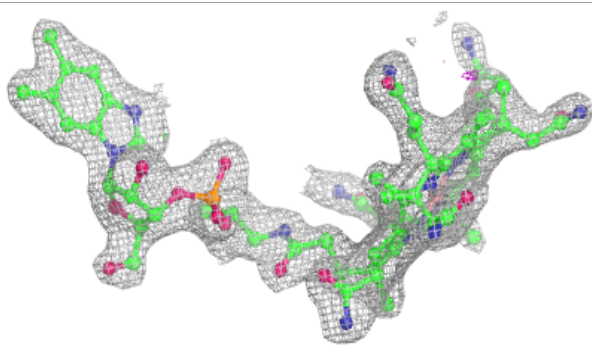
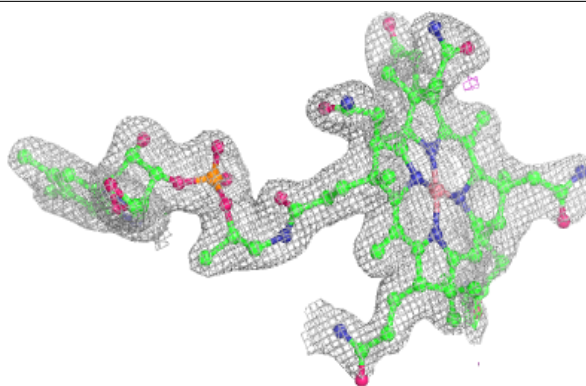
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	B12	C	301	91/91	0.79	0.22	24,61,92,115	0
4	PGE	B	302	10/10	0.83	0.19	54,62,67,68	0
3	PEG	A	302	7/7	0.87	0.21	58,67,69,71	0
2	B12	D	301	91/91	0.95	0.12	30,45,66,89	0
2	B12	A	301	91/91	0.96	0.10	29,35,57,113	0
2	B12	B	301	91/91	0.97	0.08	27,31,47,64	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.

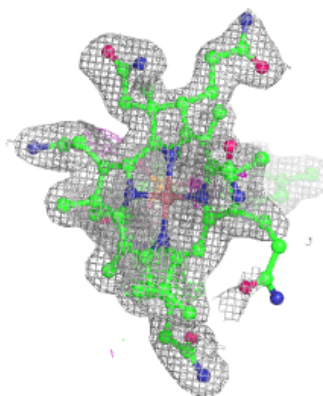
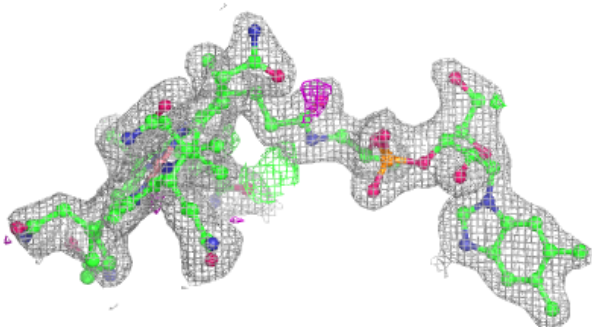
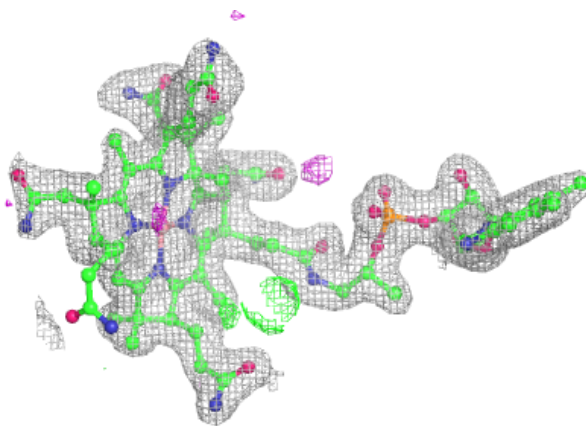


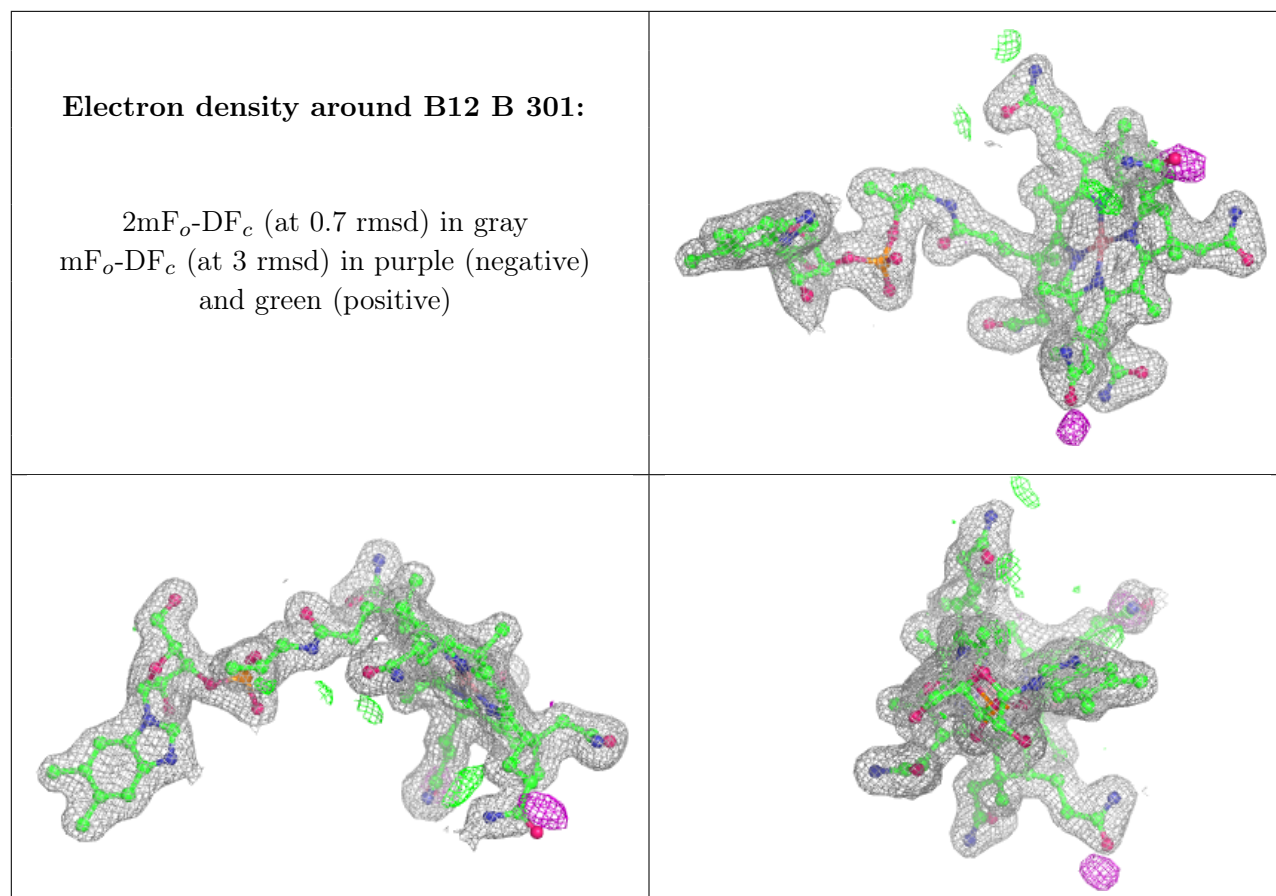
Electron density around B12 D 301:

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

**Electron density around B12 A 301:**

$2mF_o-DF_c$ (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.