wwPDB X-ray Structure Validation Summary Report (i)

Nov 27, 2023 – 12:30 pm GMT

PDB ID	:	7QH9
Title	:	TarM(Se)_G117R-4RboP
Authors	:	Guo, Y.; Stehle, T.
Deposited on	:	2021-12-10
Resolution	:	2.69 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:

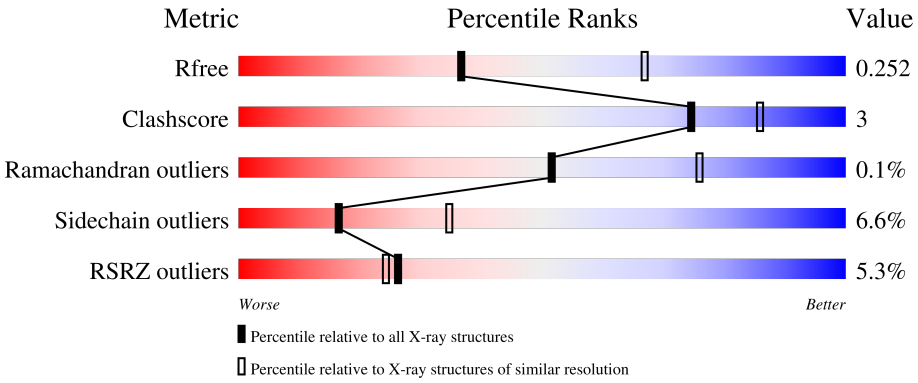
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.69 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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	508	6% 86% 10% ·
1	BBB	508	4% 86% 10% ·
1	CCC	508	6% 87% 9% · ·
1	DDD	508	4% 87% 9% ·

2 Entry composition (i)

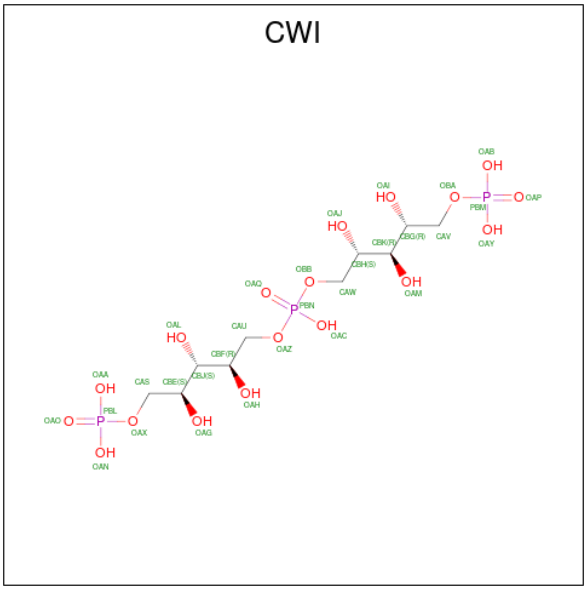
There are 5 unique types of molecules in this entry. The entry contains 14372 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 TarM(Se)_G117R-4RboP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	492	Total	C	N	O	S	0	0	0
			3518	2239	590	676	13			
1	BBB	492	Total	C	N	O	S	0	0	0
			3508	2220	588	684	16			
1	CCC	492	Total	C	N	O	S	0	0	0
			3503	2223	590	675	15			
1	DDD	491	Total	C	N	O	S	0	0	0
			3468	2204	582	669	13			

- Molecule 2 is [(2 {R},3 {S},4 {S})-2,3,4-tris(oxidanyl)-5-phosphonooxy-pentyl] [(2 {S},3 {R},4 {R})-2,3,4-tris(oxidanyl)-5-phosphonooxy-pentyl] hydrogen phosphate (three-letter code: CWI) (formula: $C_{10}H_{25}O_{18}P_3$) (labeled as "Ligand of Interest" by depositor).



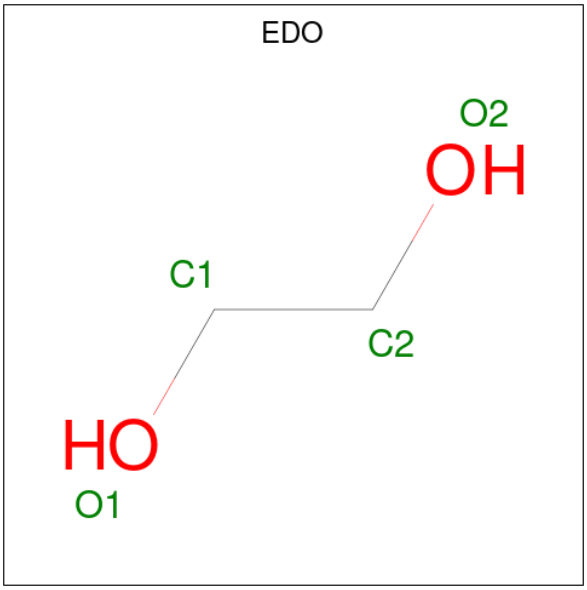
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total	C	O	P	0	0
			31	10	18	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	BBB	1	Total	C	O	P	0	0
			31	10	18	3		
2	CCC	1	Total	C	O	P	0	0
			31	10	18	3		
2	DDD	1	Total	C	O	P	0	0
			31	10	18	3		

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



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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	CCC	2	Total	Cl	0	0
			2	2		
4	DDD	1	Total	Cl	0	0
			1	1		

- Molecule 5 is water.

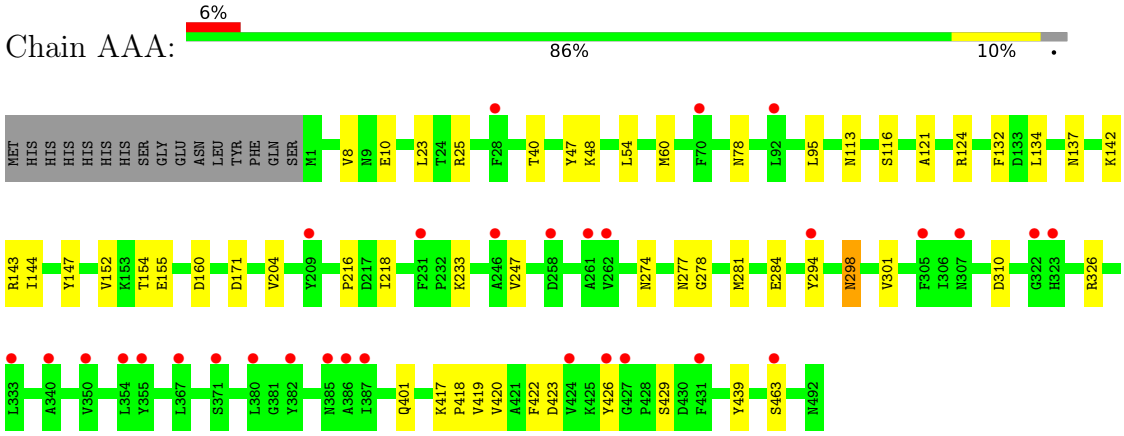
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	54	Total O 54 54	0	0
5	BBB	71	Total O 71 71	0	0
5	CCC	54	Total O 54 54	0	0
5	DDD	61	Total O 61 61	0	0

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.

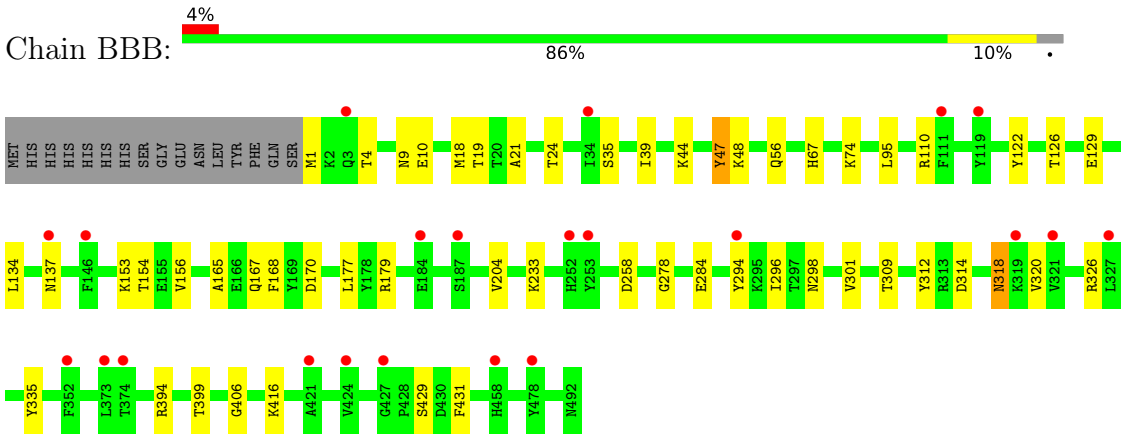
• Molecule 1: TarM(Se)_G117R-4RboP

Chain AAA:



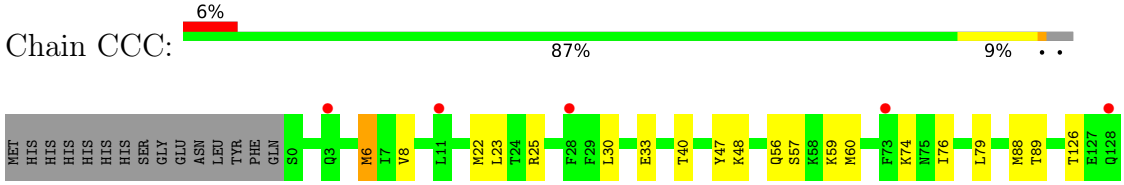
• Molecule 1: TarM(Se)_G117R-4RboP

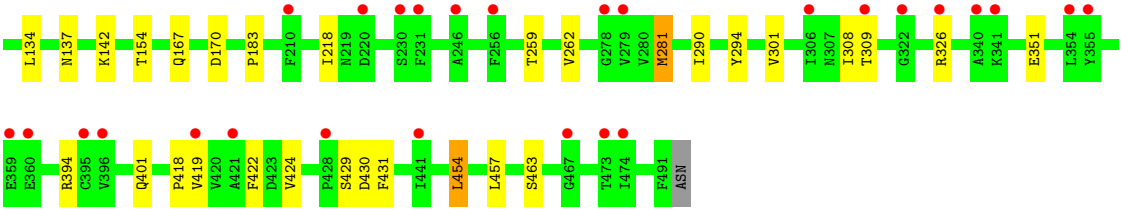
Chain BBB:



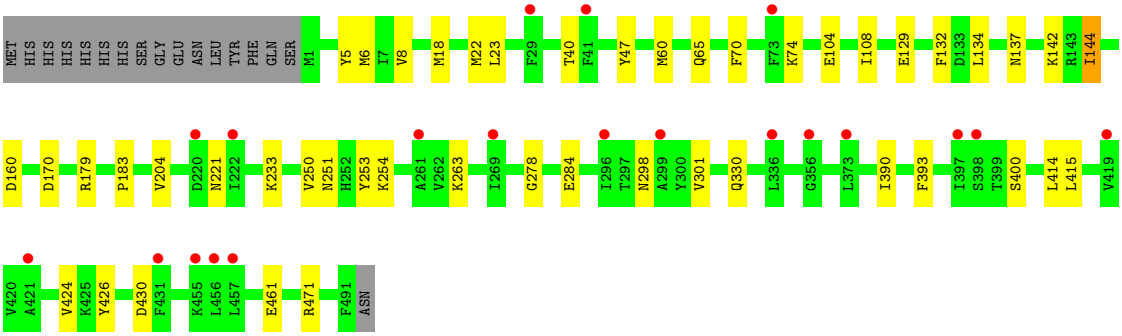
• Molecule 1: TarM(Se)_G117R-4RboP

Chain CCC:





- Molecule 1: TarM(Se)_G117R-4RboP



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	58.87Å 70.57Å 137.38Å 89.97° 90.01° 90.02°	Depositor
Resolution (Å)	49.23 – 2.69 49.21 – 2.69	Depositor EDS
% Data completeness (in resolution range)	95.1 (49.23-2.69) 95.1 (49.21-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.243 , 0.253 0.242 , 0.252	Depositor DCC
R_{free} test set	2924 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	78.3	Xtriage
Anisotropy	0.520	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.21 , 59.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.449 for h,-k,-l 0.458 for -h,k,-l 0.457 for -h,-k,l	Xtriage
Reported twinning fraction	0.331 for H, K, L 0.198 for h,-k,-l 0.217 for -h,-k,l 0.254 for -H, K, -L	Depositor
Outliers	0 of 58472 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14372	wwPDB-VP
Average B, all atoms (Å^2)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% 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: CL, EDO, CWI

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	0.68	0/3581	0.68	0/4876
1	BBB	0.69	0/3570	0.69	0/4862
1	CCC	0.69	0/3564	0.69	0/4853
1	DDD	0.69	0/3530	0.69	0/4805
All	All	0.69	0/14245	0.69	0/19396

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	3518	0	2839	24	0
1	BBB	3508	0	2830	20	0
1	CCC	3503	0	2835	19	0
1	DDD	3468	0	2761	19	0
2	AAA	31	0	0	0	0
2	BBB	31	0	0	0	0
2	CCC	31	0	0	0	0
2	DDD	31	0	0	0	0
3	BBB	4	0	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	DDD	4	0	6	0	0
4	CCC	2	0	0	0	0
4	DDD	1	0	0	0	0
5	AAA	54	0	0	0	0
5	BBB	71	0	0	0	0
5	CCC	54	0	0	0	0
5	DDD	61	0	0	1	0
All	All	14372	0	11277	82	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 3.

The worst 5 of 82 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:294:TYR:CZ	1:AAA:294:TYR:CD1	2.50	0.94
1:DDD:65:GLN:OE1	5:DDD:601:HOH:O	2.12	0.67
1:BBB:278:GLY:HA2	1:BBB:298:ASN:HB2	1.77	0.64
1:CCC:154:THR:OG1	1:CCC:167:GLN:HB2	2.03	0.59
1:AAA:147:TYR:HB3	1:AAA:152:VAL:HG11	1.84	0.58

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	490/508 (96%)	457 (93%)	33 (7%)	0	100	100
1	BBB	490/508 (96%)	442 (90%)	48 (10%)	0	100	100
1	CCC	490/508 (96%)	449 (92%)	40 (8%)	1 (0%)	47	71
1	DDD	489/508 (96%)	446 (91%)	42 (9%)	1 (0%)	47	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1959/2032 (96%)	1794 (92%)	163 (8%)	2 (0%)	51 76

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	DDD	183	PRO
1	CCC	183	PRO

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	269/471 (57%)	250 (93%)	19 (7%)	14 31
1	BBB	272/471 (58%)	253 (93%)	19 (7%)	15 32
1	CCC	268/471 (57%)	247 (92%)	21 (8%)	12 27
1	DDD	260/471 (55%)	248 (95%)	12 (5%)	27 51
All	All	1069/1884 (57%)	998 (93%)	71 (7%)	16 35

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	431	PHE
1	DDD	18	MET
1	DDD	179	ARG
1	BBB	56	GLN
1	BBB	48	LYS

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, 3 are monoatomic - leaving 6 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	$\#\|Z\| > 2$
2	CWI	AAA	501	-	30,30,30	1.52	4 (13%)	43,44,44	1.02	2 (4%)
3	EDO	BBB	501	-	3,3,3	0.25	0	2,2,2	0.22	0
3	EDO	DDD	502	-	3,3,3	0.29	0	2,2,2	0.25	0
2	CWI	DDD	501	-	30,30,30	1.72	8 (26%)	43,44,44	1.01	3 (6%)
2	CWI	BBB	502	-	30,30,30	1.58	5 (16%)	43,44,44	1.07	3 (6%)
2	CWI	CCC	501	-	30,30,30	1.51	4 (13%)	43,44,44	1.05	2 (4%)

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	CWI	AAA	501	-	-	24/40/40/40	-
3	EDO	BBB	501	-	-	1/1/1/1	-
3	EDO	DDD	502	-	-	1/1/1/1	-
2	CWI	DDD	501	-	-	18/40/40/40	-
2	CWI	BBB	502	-	-	16/40/40/40	-
2	CWI	CCC	501	-	-	29/40/40/40	-

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	DDD	501	CWI	PBN-OAQ	3.76	1.64	1.50
2	DDD	501	CWI	PBM-OAP	3.75	1.62	1.50
2	BBB	502	CWI	PBL-OAO	3.63	1.62	1.50
2	CCC	501	CWI	PBM-OAP	3.59	1.62	1.50
2	BBB	502	CWI	PBM-OAP	3.57	1.62	1.50

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	DDD	501	CWI	CAV-CBG-CBK	2.80	117.62	112.20
2	DDD	501	CWI	OAA-PBL-OAX	2.53	113.48	106.73
2	BBB	502	CWI	CAU-CBF-CBJ	-2.41	107.55	112.20
2	CCC	501	CWI	CAU-CBF-CBJ	-2.32	107.72	112.20
2	AAA	501	CWI	CAV-CBG-CBK	2.31	116.67	112.20

There are no chirality outliers.

5 of 89 torsion outliers are listed below:

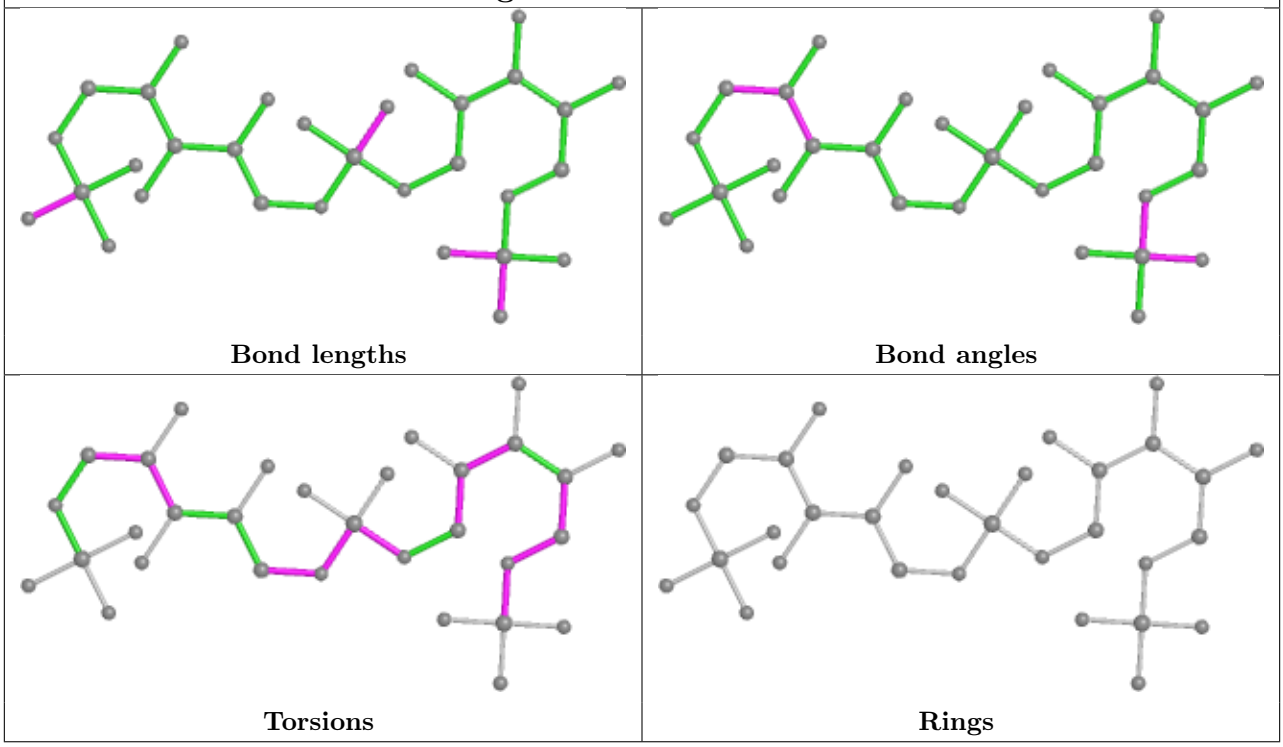
Mol	Chain	Res	Type	Atoms
2	AAA	501	CWI	CAS-OAX-PBL-OAN
2	AAA	501	CWI	CAS-OAX-PBL-OAA
2	AAA	501	CWI	OAX-CAS-CBE-CBJ
2	AAA	501	CWI	OAH-CBF-CBJ-CBE
2	AAA	501	CWI	CAU-CBF-CBJ-CBE

There are no ring outliers.

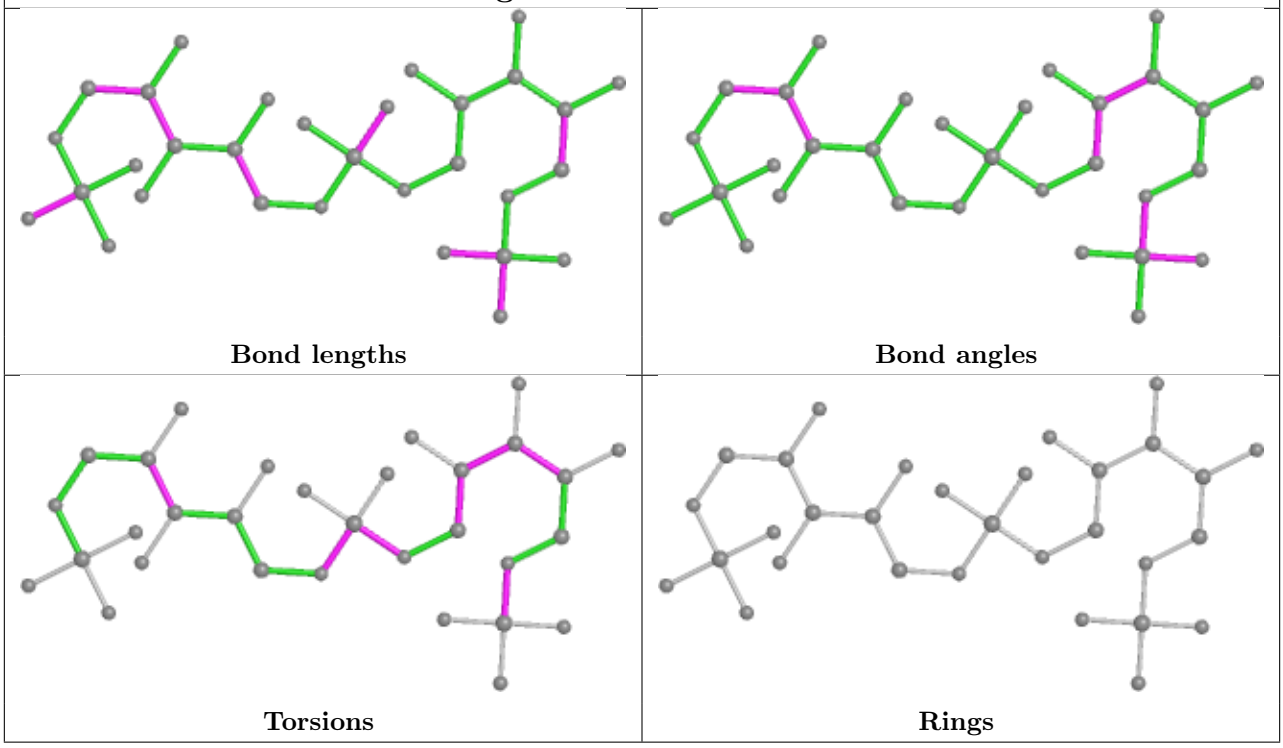
No monomer is involved in short contacts.

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.

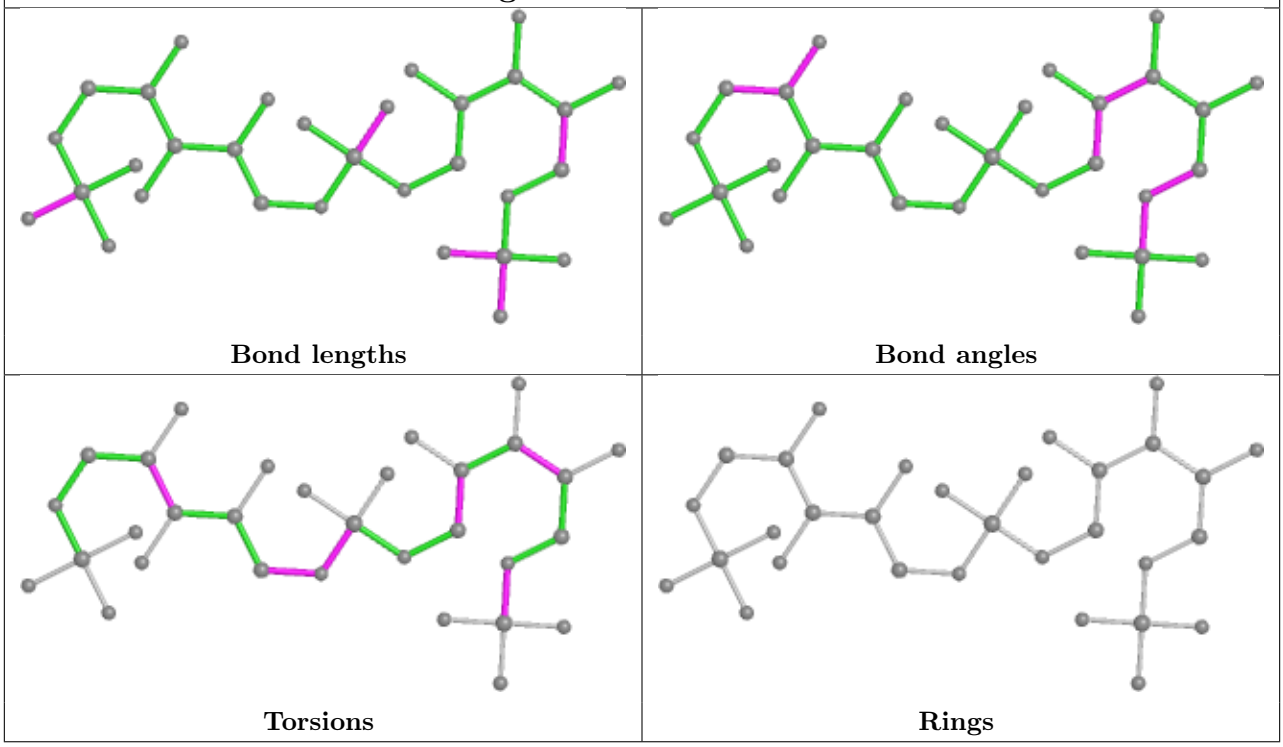
Ligand CWI AAA 501



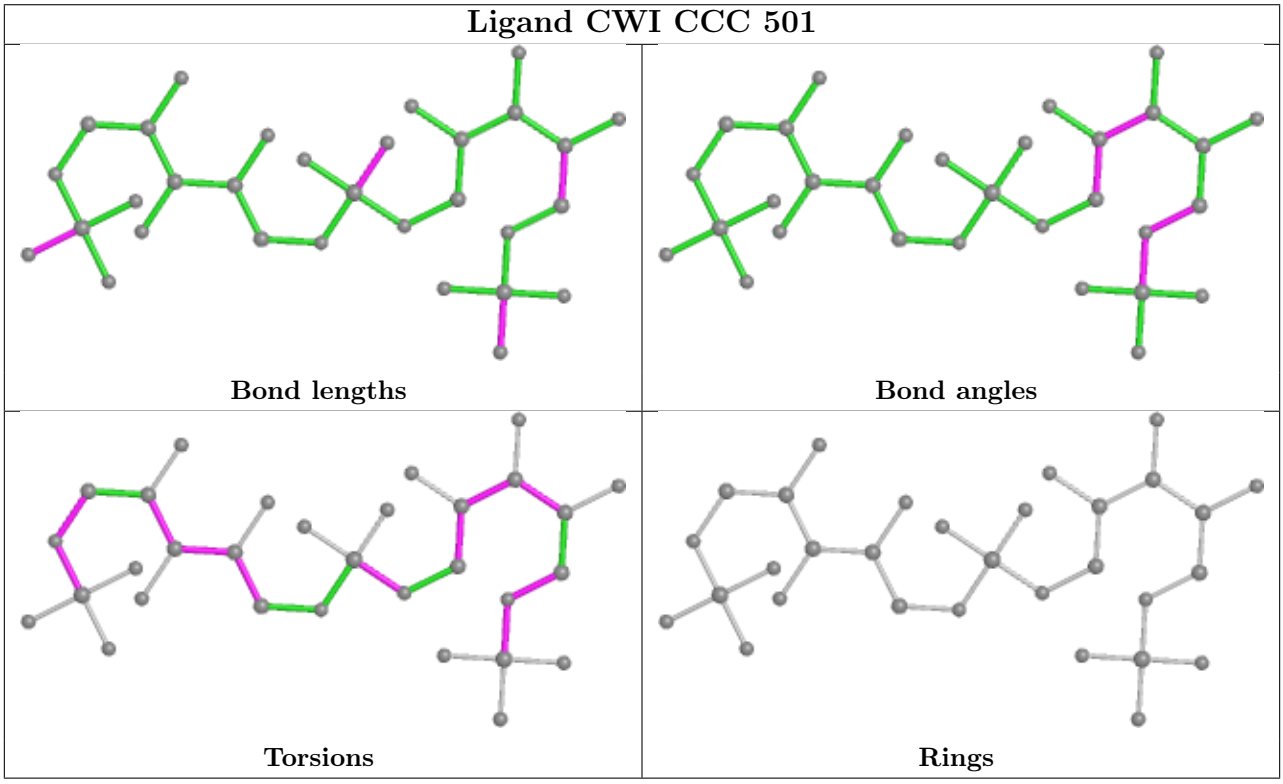
Ligand CWI DDD 501



Ligand CWI BBB 502



Ligand CWI CCC 501



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	492/508 (96%)	0.20	31 (6%)	20	18	54, 98, 142, 167	0
1	BBB	492/508 (96%)	0.05	22 (4%)	33	31	54, 99, 132, 161	0
1	CCC	492/508 (96%)	0.08	32 (6%)	18	16	60, 99, 137, 182	0
1	DDD	491/508 (96%)	0.02	20 (4%)	37	35	43, 98, 130, 172	0
All	All	1967/2032 (96%)	0.09	105 (5%)	26	24	43, 99, 135, 182	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	386	ALA	7.8
1	AAA	350	VAL	6.8
1	CCC	355	TYR	6.7
1	AAA	322	GLY	6.0
1	AAA	424	VAL	5.9

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

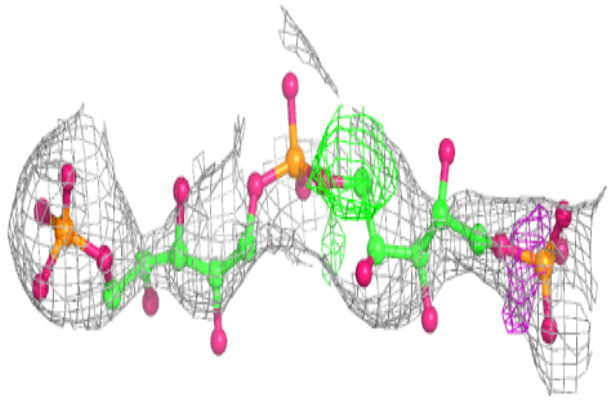
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

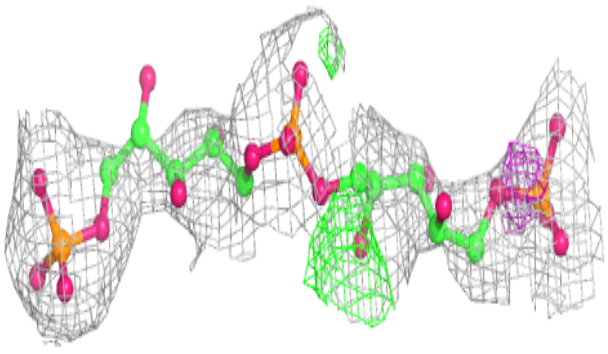
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CWI	BBB	502	31/31	0.83	0.27	69,96,118,127	0
2	CWI	AAA	501	31/31	0.85	0.29	82,113,135,147	0
2	CWI	DDD	501	31/31	0.85	0.20	70,87,113,120	0
2	CWI	CCC	501	31/31	0.88	0.20	70,84,142,144	0
3	EDO	DDD	502	4/4	0.88	0.20	24,29,31,32	0
3	EDO	BBB	501	4/4	0.91	0.15	24,25,27,31	0
4	CL	CCC	502	1/1	0.97	0.07	37,37,37,37	0
4	CL	CCC	503	1/1	0.98	0.18	40,40,40,40	0
4	CL	DDD	503	1/1	0.98	0.08	37,37,37,37	0

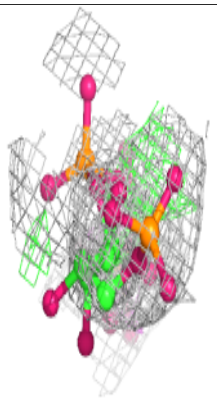
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CWI BBB 502:

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

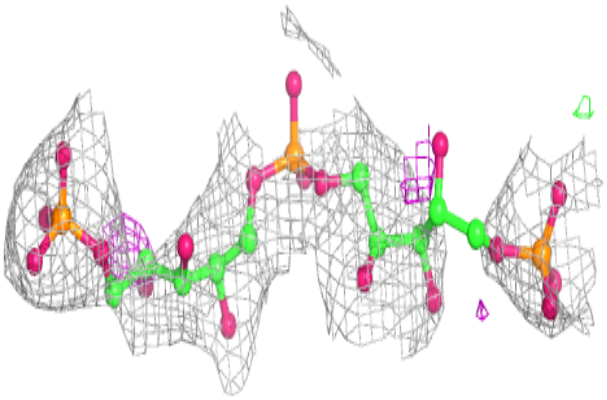


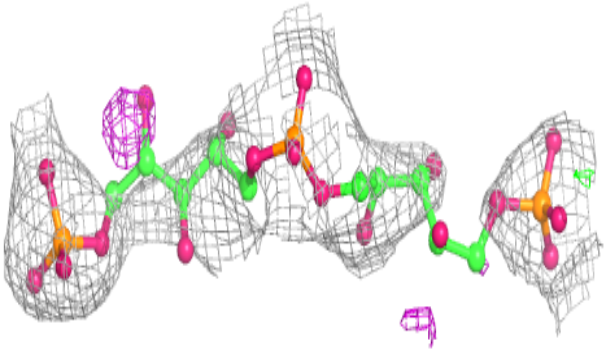


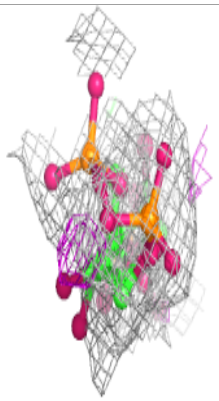


Electron density around CWI AAA 501:

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

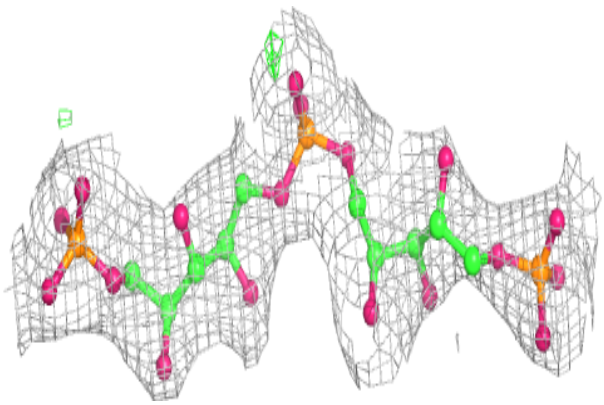


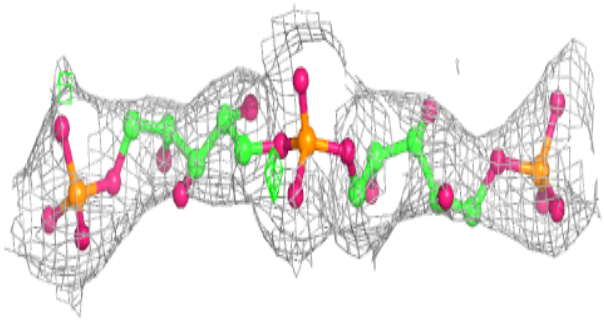


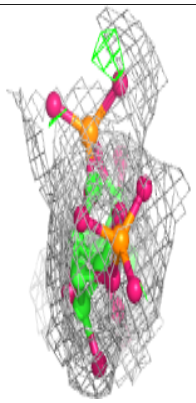


Electron density around CWI DDD 501:

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

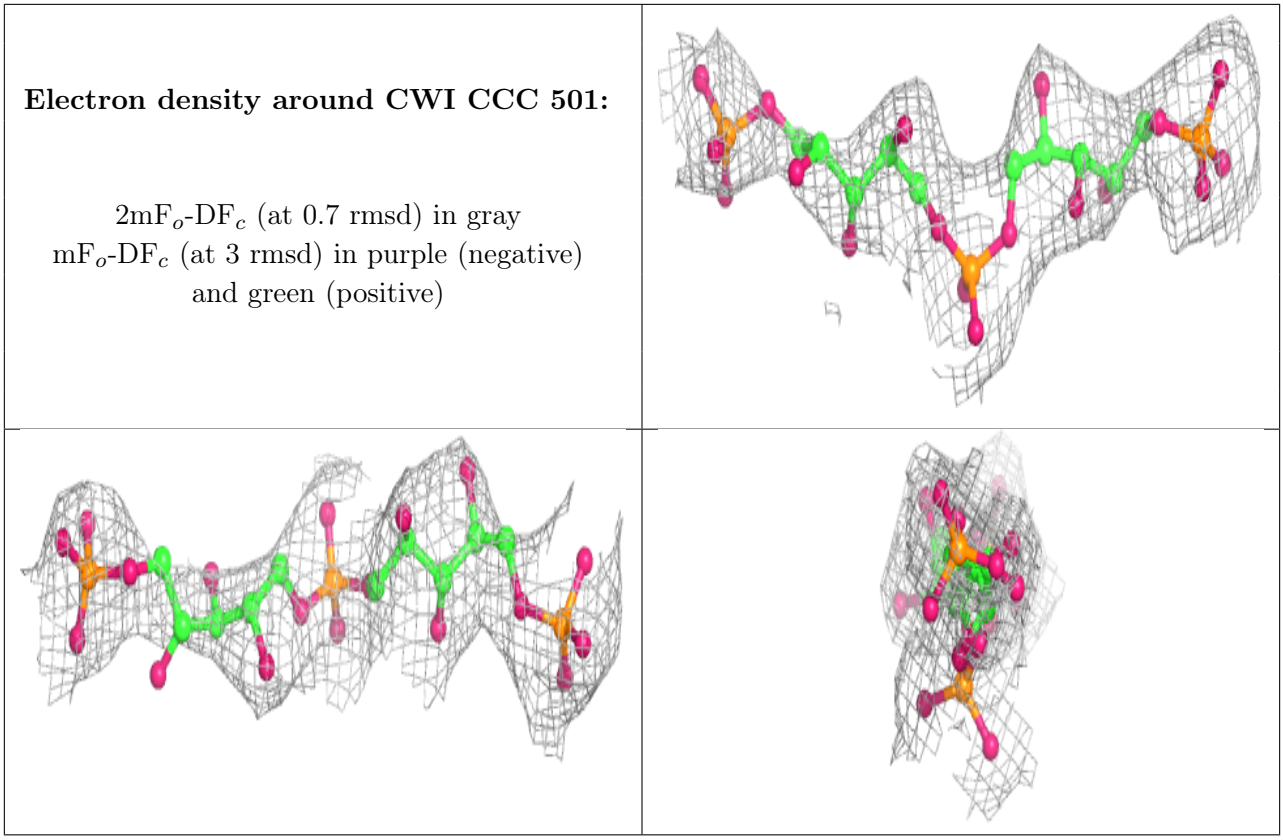






Electron density around CWI CCC 501:

$2mF_o$-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.