wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2020 – 11:22 AM BST

PDB ID	:	5ZDC
Title	:	Crystal structure of poly(ADP-ribose) glycohydrolase (PARG) from Deinococcus radiodurans in complex with ADP-ribose (P32)
Authors	:	Cho, C.C.; Hsu, C.H.
Deposited on	:	2018-02-23
Resolution	:	1.98 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 following versions of software and data (see references (1)) were used in the production of this report:

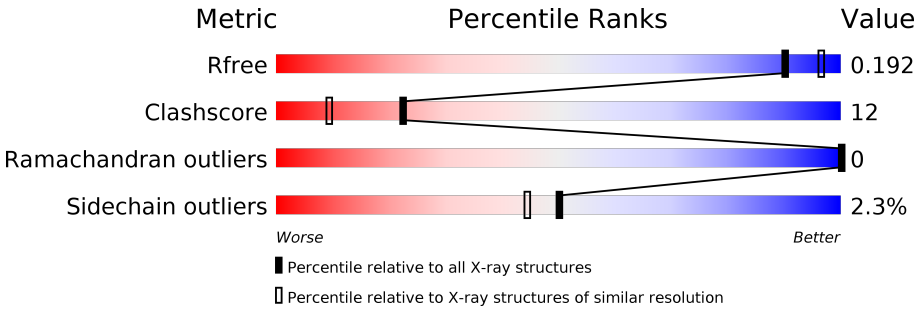
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 1.98 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	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)

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 on 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%

Mol	Chain	Length	Quality of chain
1	C	253	84% 15% •
1	D	253	77% 20% • •
1	F	253	79% 20%
1	H	253	80% 17% ••
1	J	253	75% 22% ••
1	L	253	80% 16% • •

2 Entry composition

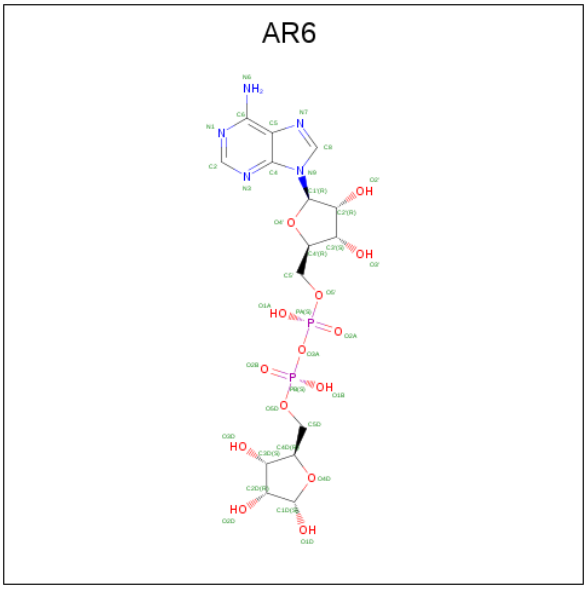
There are 4 unique types of molecules in this entry. The entry contains 12800 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 poly ADP-ribose glycohydrolase.

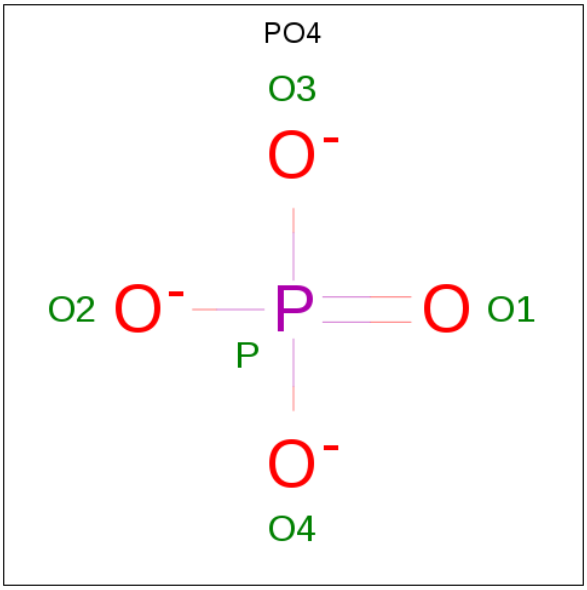
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	253	Total	C	N	O	S	0	0	0
			1944	1218	363	357	6			
1	F	252	Total	C	N	O	S	0	0	0
			1937	1213	362	356	6			
1	H	250	Total	C	N	O	S	0	0	0
			1928	1208	360	354	6			
1	J	250	Total	C	N	O	S	0	0	0
			1928	1208	360	354	6			
1	L	248	Total	C	N	O	S	0	0	0
			1916	1200	358	352	6			
1	D	249	Total	C	N	O	S	0	0	0
			1923	1205	359	353	6			

- Molecule 2 is [(2R,3S,4R,5R)-5-(6-AMINOPURIN-9-YL)-3,4-DIHYDROXY-OXOLAN-2-YL]METHYL [HYDROXY-[[(2R,3S,4R,5S)-3,4,5-TRIHYDROXYOXOLAN-2-YL]METHOXY]PHOSPHORYL] HYDROGEN PHOSPHATE (three-letter code: AR6) (formula: $C_{15}H_{23}N_5O_{14}P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C N O P 36 15 5 14 2	0	0
2	F	1	Total C N O P 36 15 5 14 2	0	0
2	H	1	Total C N O P 36 15 5 14 2	0	0
2	J	1	Total C N O P 36 15 5 14 2	0	0
2	L	1	Total C N O P 36 15 5 14 2	0	0
2	D	1	Total C N O P 36 15 5 14 2	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total O P 5 4 1	0	0
3	F	1	Total O P 5 4 1	0	0
3	D	1	Total O P 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	161	Total O 161 161	0	0

Continued on next page...

Continued from previous page...

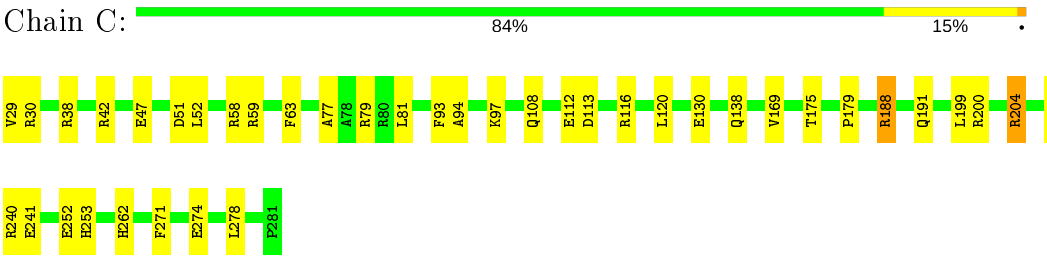
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	F	184	Total O 184 184	0	0
4	H	142	Total O 142 142	0	0
4	J	167	Total O 167 167	0	0
4	L	197	Total O 197 197	0	0
4	D	142	Total O 142 142	0	0

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. 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. 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.

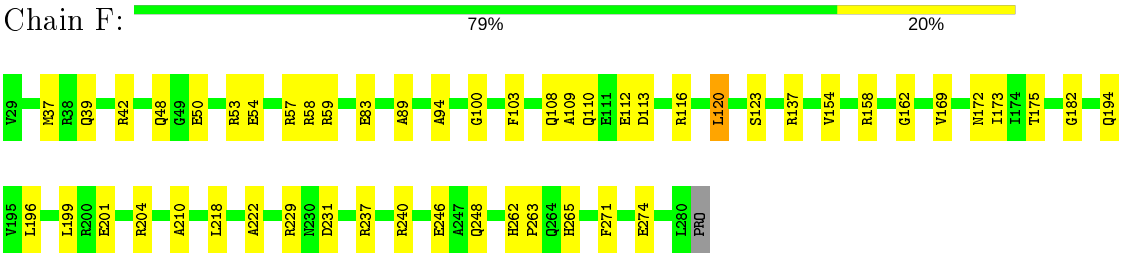
- Molecule 1: poly ADP-ribose glycohydrolase

Chain C:



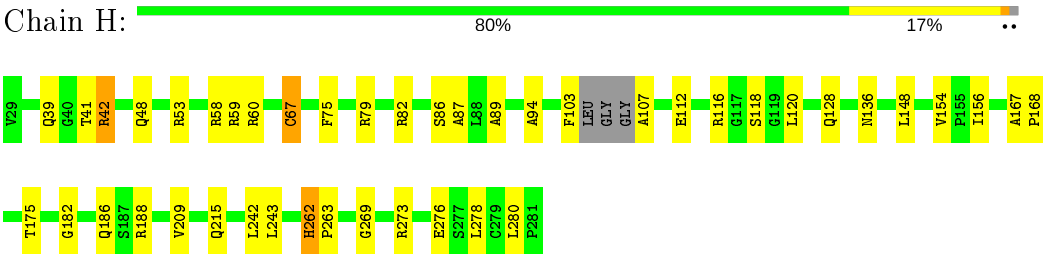
- Molecule 1: poly ADP-ribose glycohydrolase

Chain F:



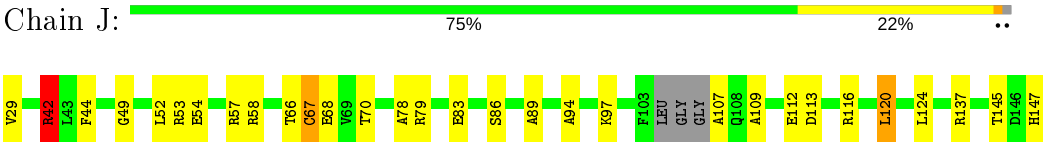
- Molecule 1: poly ADP-ribose glycohydrolase

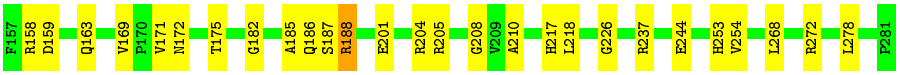
Chain H:



- Molecule 1: poly ADP-ribose glycohydrolase

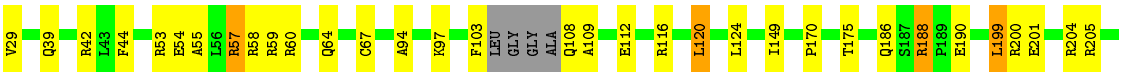
Chain J:

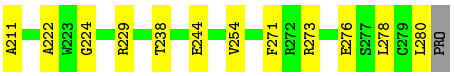




- Molecule 1: poly ADP-ribose glycohydrolase

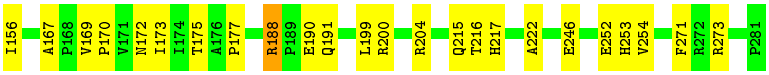
Chain L:





- Molecule 1: poly ADP-ribose glycohydrolase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	108.50Å 108.50Å 130.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.03 – 1.98 26.06 – 1.98	Depositor EDS
% Data completeness (in resolution range)	93.4 (28.03-1.98) 1.4 (26.06-1.98)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	PHENIX (1.13_2998)	Depositor
R, R_{free}	0.144 , 0.191 0.146 , 0.192	Depositor DCC
R_{free} test set	30 reflections (1.75%)	wwPDB-VP
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 25.7	EDS
L-test for twinning[1]	$< \|L\| >$ = (Not available), $< L^2 >$ = (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.60	EDS
Total number of atoms	12800	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

[1] 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: PO4, AR6

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	C	0.40	0/1989	0.60	0/2705
1	D	0.41	0/1967	0.62	1/2674 (0.0%)
1	F	0.44	0/1981	0.60	0/2693
1	H	0.43	0/1972	0.59	1/2681 (0.0%)
1	J	0.40	0/1972	0.60	1/2681 (0.0%)
1	L	0.51	1/1959 (0.1%)	0.60	0/2662
All	All	0.43	1/11840 (0.0%)	0.60	3/16096 (0.0%)

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	H	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	57	ARG	CZ-NH1	10.05	1.46	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	242	LEU	CA-CB-CG	6.80	130.95	115.30
1	D	38	ARG	NE-CZ-NH2	-6.07	117.26	120.30
1	J	42	ARG	CG-CD-NE	5.05	122.41	111.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	262	HIS	Peptide

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	C	1944	0	1917	33	0
1	D	1923	0	1894	46	0
1	F	1937	0	1910	49	0
1	H	1928	0	1899	39	0
1	J	1928	0	1899	58	0
1	L	1916	0	1887	47	0
2	C	36	0	19	1	0
2	D	36	0	19	1	0
2	F	36	0	19	3	0
2	H	36	0	19	1	0
2	J	36	0	19	4	0
2	L	36	0	19	2	0
3	C	5	0	0	0	0
3	D	5	0	0	0	0
3	F	5	0	0	0	0
4	C	161	0	0	12	0
4	D	142	0	0	12	0
4	F	184	0	0	20	0
4	H	142	0	0	10	0
4	J	167	0	0	18	0
4	L	197	0	0	21	0
All	All	12800	0	11520	270	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 12.

The worst 5 of 270 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:AR6:C1'	2:D:301:AR6:O4'	1.64	1.29
1:H:59:ARG:NH1	1:D:86:SER:OG	1.75	1.19

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:L:301:AR6:C1'	2:L:301:AR6:O4'	1.64	1.18
2:J:301:AR6:O4'	2:J:301:AR6:C1'	1.65	1.18
2:H:301:AR6:O4'	2:H:301:AR6:C1'	1.65	1.17

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	C	251/253 (99%)	244 (97%)	7 (3%)	0	100	100
1	D	245/253 (97%)	238 (97%)	7 (3%)	0	100	100
1	F	250/253 (99%)	242 (97%)	8 (3%)	0	100	100
1	H	246/253 (97%)	237 (96%)	9 (4%)	0	100	100
1	J	246/253 (97%)	237 (96%)	9 (4%)	0	100	100
1	L	244/253 (96%)	238 (98%)	6 (2%)	0	100	100
All	All	1482/1518 (98%)	1436 (97%)	46 (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	C	198/198 (100%)	192 (97%)	6 (3%)	41 29
1	D	197/198 (100%)	191 (97%)	6 (3%)	41 29
1	F	197/198 (100%)	193 (98%)	4 (2%)	55 48
1	H	197/198 (100%)	194 (98%)	3 (2%)	65 59
1	J	197/198 (100%)	193 (98%)	4 (2%)	55 48
1	L	196/198 (99%)	192 (98%)	4 (2%)	55 48
All	All	1182/1188 (100%)	1155 (98%)	27 (2%)	50 44

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	120	LEU
1	J	120	LEU
1	D	120	LEU
1	J	42	ARG
1	C	188	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	261	ASN
1	D	138	GLN
1	L	48	GLN
1	F	186	GLN
1	L	186	GLN

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)

9 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	AR6	D	301	-	34,39,39	4.92	12 (35%)	40,60,60	2.25	6 (15%)
3	PO4	D	302	-	4,4,4	1.06	0	6,6,6	0.79	0
2	AR6	F	301	-	34,39,39	4.95	15 (44%)	40,60,60	2.07	4 (10%)
2	AR6	H	301	-	34,39,39	4.85	13 (38%)	40,60,60	2.19	6 (15%)
3	PO4	C	302	-	4,4,4	1.02	0	6,6,6	0.59	0
2	AR6	C	301	-	34,39,39	4.82	14 (41%)	40,60,60	2.42	6 (15%)
2	AR6	J	301	-	34,39,39	4.94	15 (44%)	40,60,60	2.15	5 (12%)
2	AR6	L	301	-	34,39,39	4.71	14 (41%)	40,60,60	2.55	6 (15%)
3	PO4	F	302	-	4,4,4	1.02	0	6,6,6	0.61	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
2	AR6	D	301	-	-	6/18/54/54	0/4/4/4
2	AR6	F	301	-	-	2/18/54/54	0/4/4/4
2	AR6	H	301	-	-	2/18/54/54	0/4/4/4
2	AR6	C	301	-	-	1/18/54/54	0/4/4/4
2	AR6	J	301	-	-	1/18/54/54	0/4/4/4
2	AR6	L	301	-	-	1/18/54/54	0/4/4/4

The worst 5 of 83 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	301	AR6	O4'-C1'	17.26	1.65	1.41
2	H	301	AR6	O4'-C1'	17.15	1.65	1.41
2	F	301	AR6	O4'-C1'	17.14	1.65	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	AR6	O4'-C1'	16.97	1.64	1.41
2	D	301	AR6	O4'-C1'	16.90	1.64	1.41

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	301	AR6	C5-C6-N6	11.30	137.52	120.35
2	C	301	AR6	C5-C6-N6	10.35	136.09	120.35
2	D	301	AR6	C5-C6-N6	9.09	134.17	120.35
2	F	301	AR6	C5-C6-N6	8.98	133.99	120.35
2	H	301	AR6	C5-C6-N6	8.85	133.80	120.35

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	301	AR6	C5D-O5D-PB-O2B
2	C	301	AR6	PA-O3A-PB-O5D
2	D	301	AR6	PA-O3A-PB-O5D
2	D	301	AR6	C5D-O5D-PB-O3A
2	F	301	AR6	PB-O3A-PA-O1A

There are no ring outliers.

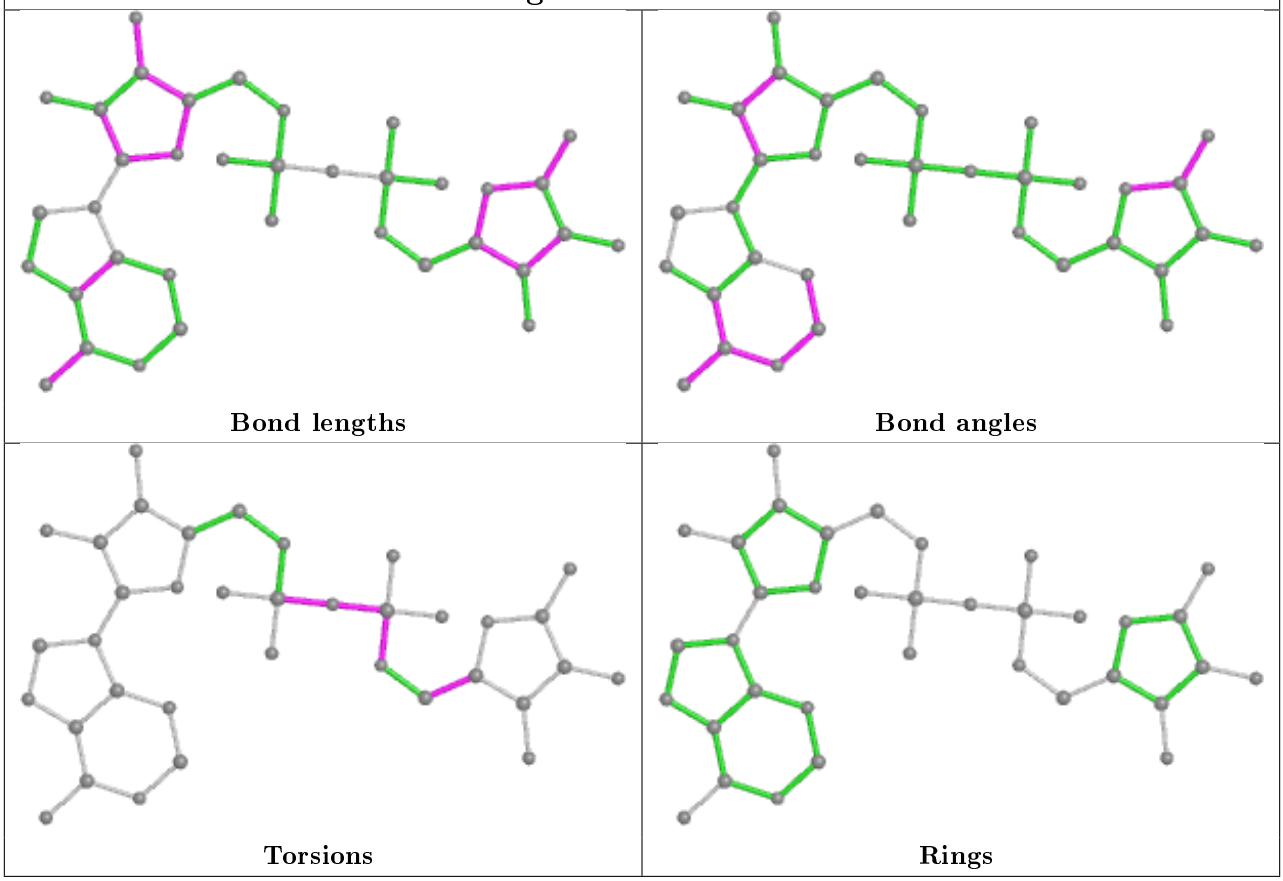
6 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	301	AR6	1	0
2	F	301	AR6	3	0
2	H	301	AR6	1	0
2	C	301	AR6	1	0
2	J	301	AR6	4	0
2	L	301	AR6	2	0

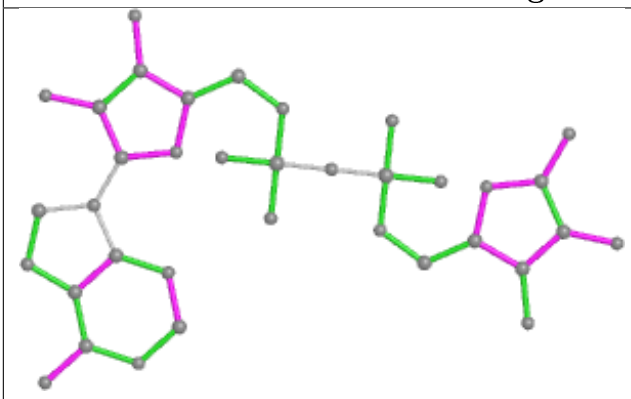
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

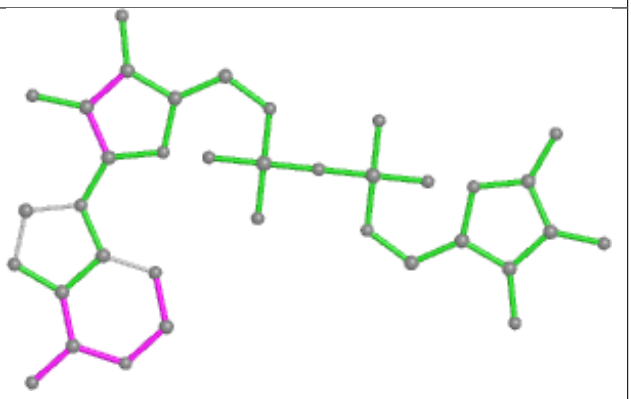
Ligand AR6 D 301



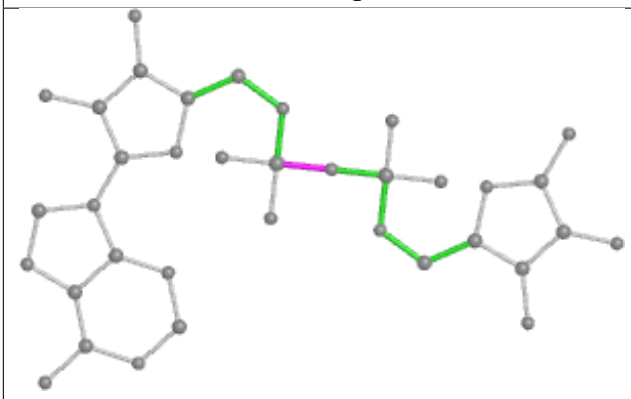
Ligand AR6 F 301



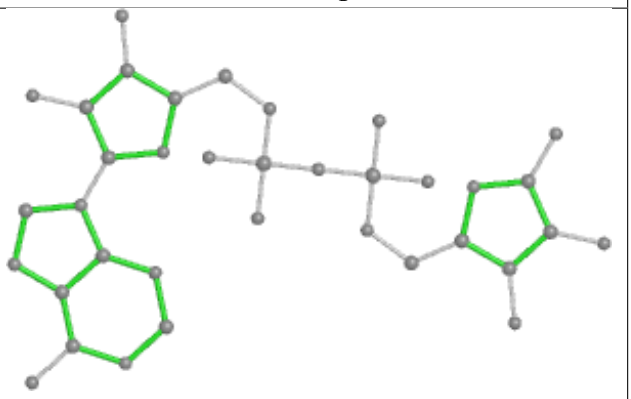
Bond lengths



Bond angles

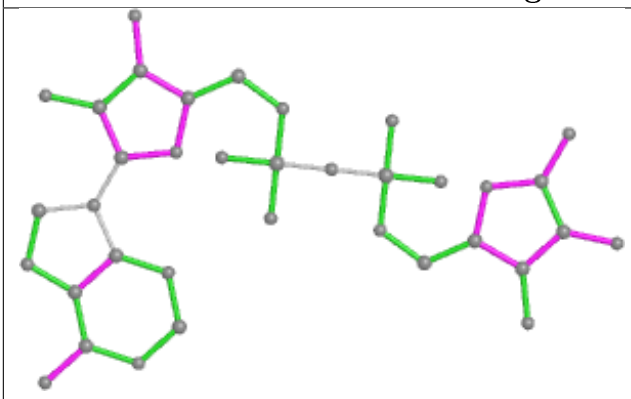


Torsions

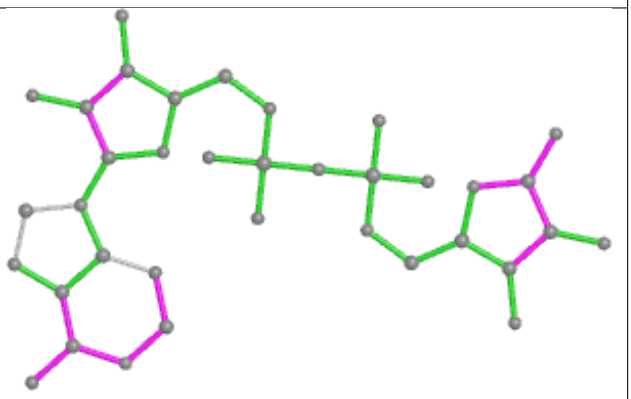


Rings

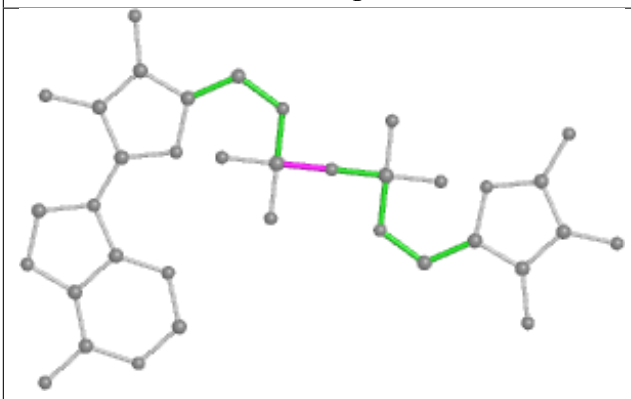
Ligand AR6 H 301



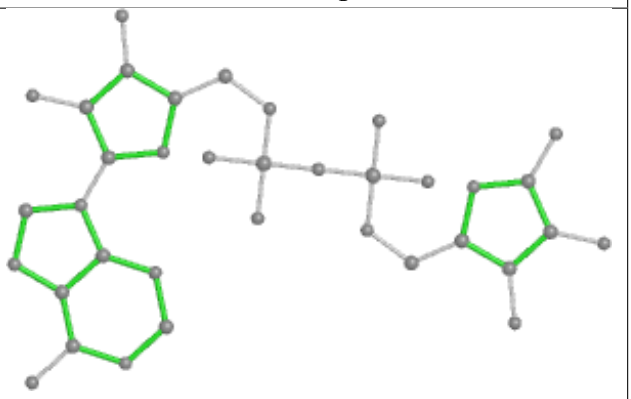
Bond lengths



Bond angles

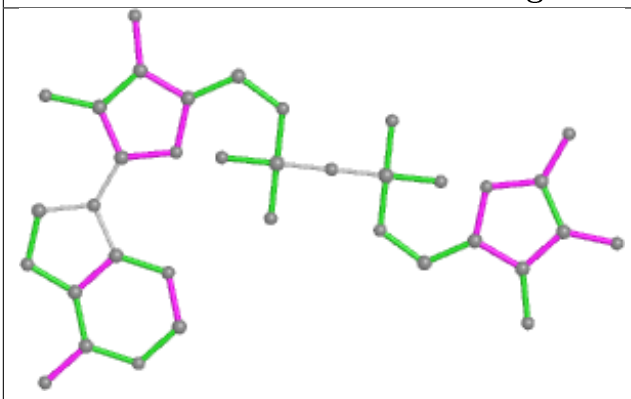


Torsions

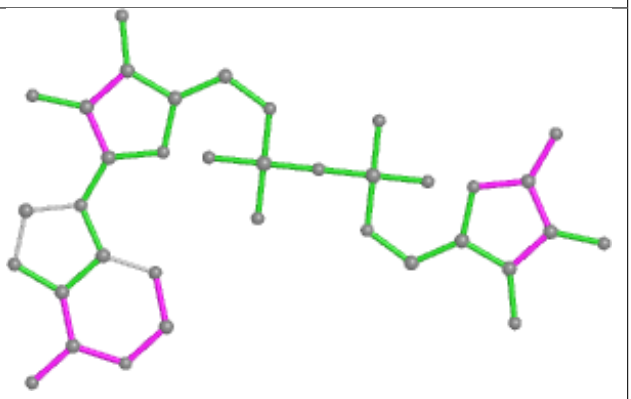


Rings

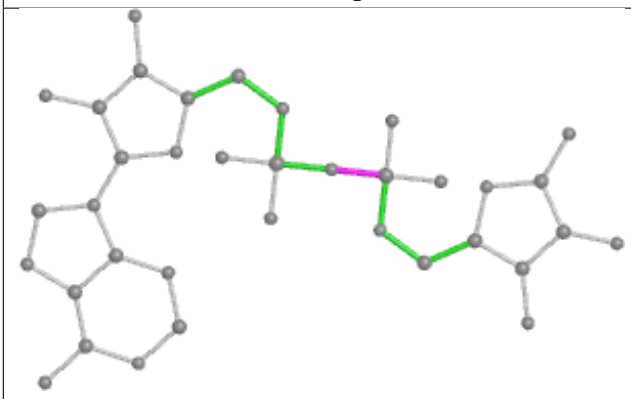
Ligand AR6 C 301



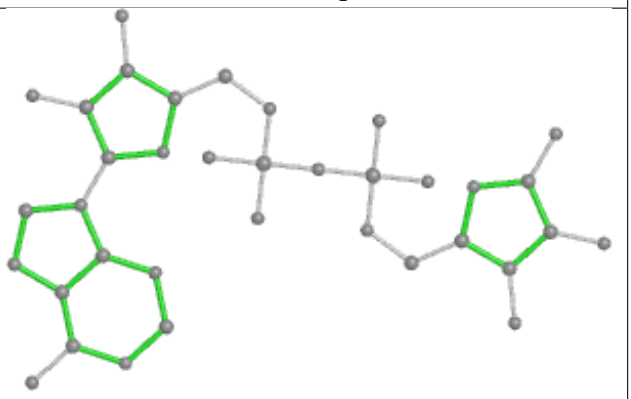
Bond lengths



Bond angles

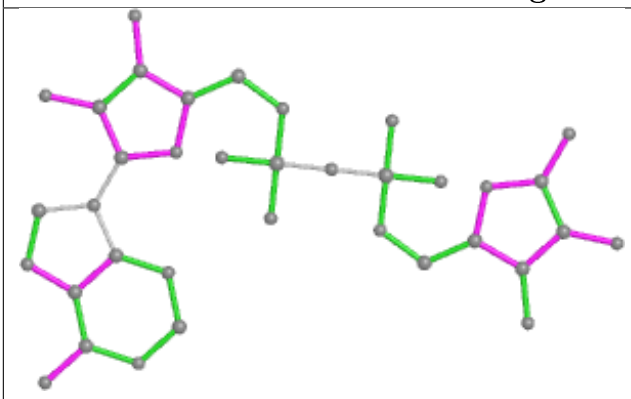


Torsions

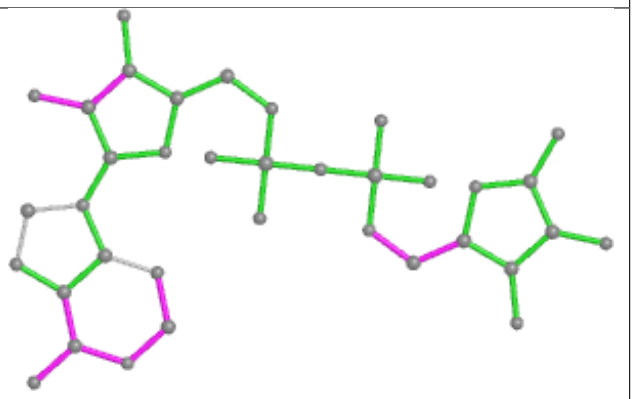


Rings

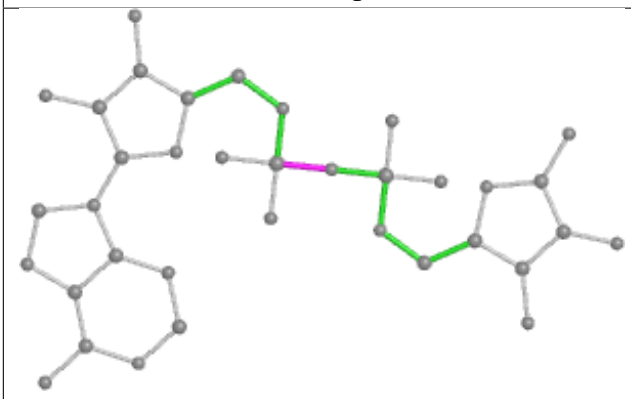
Ligand AR6 J 301



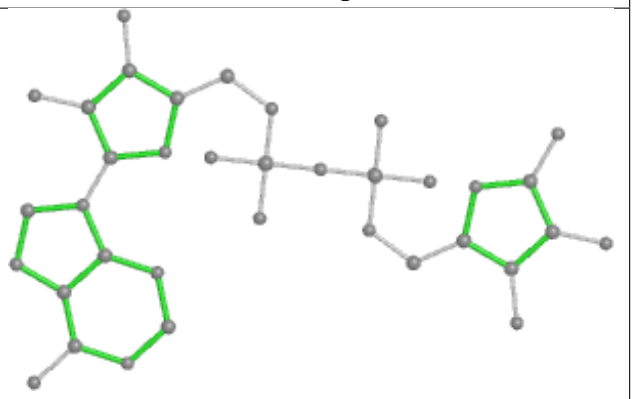
Bond lengths



Bond angles

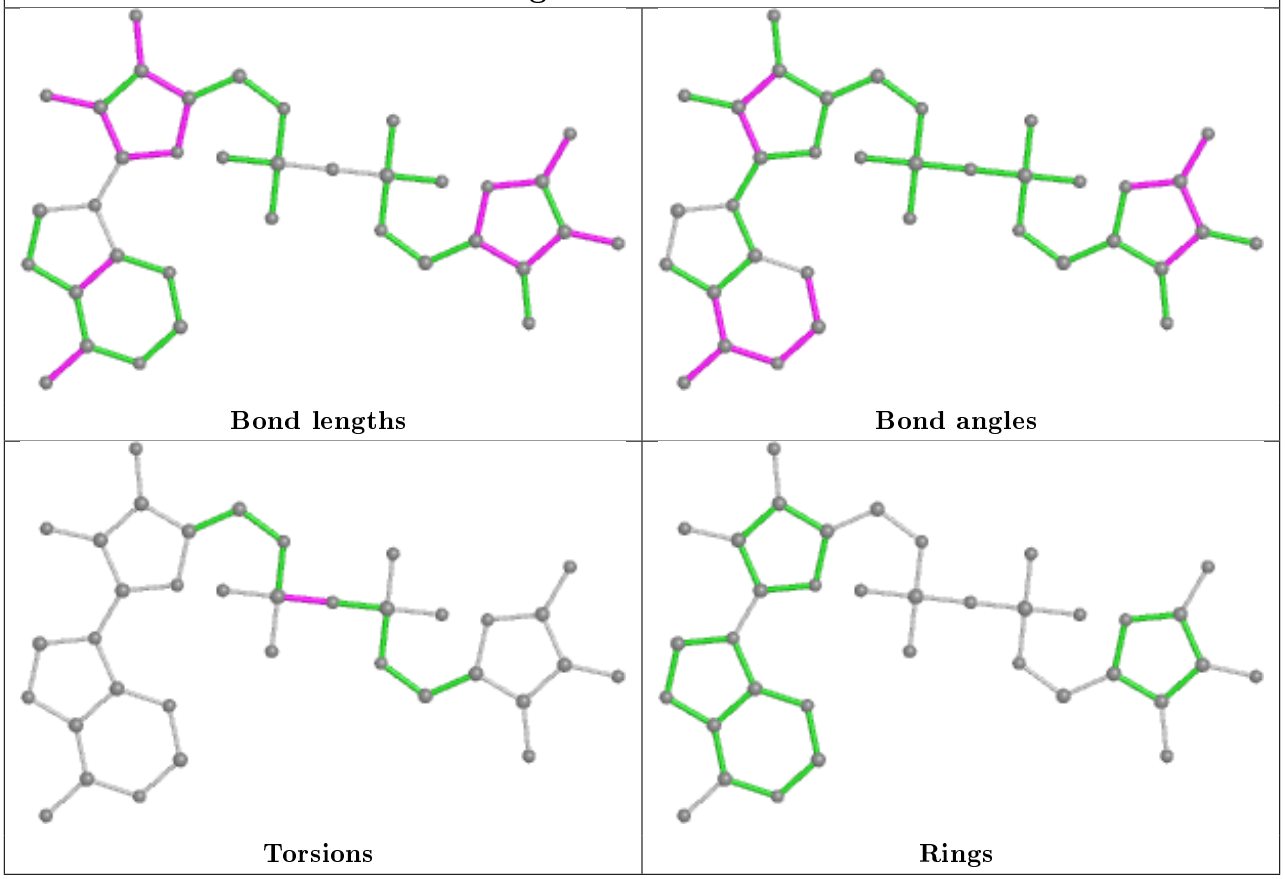


Torsions



Rings

Ligand AR6 L 301



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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

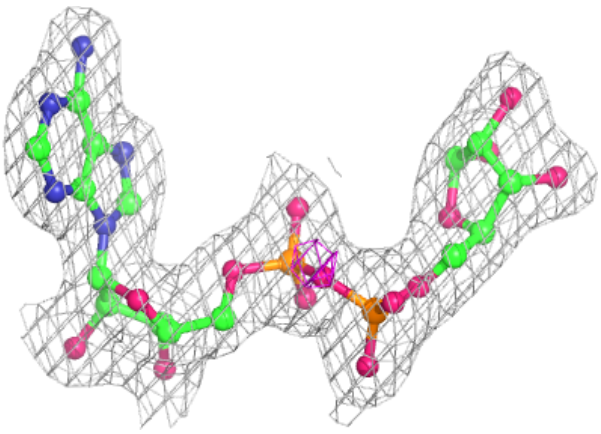
6.4 Ligands

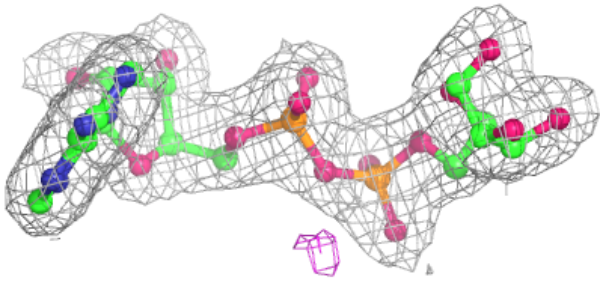
Unable to reproduce the depositors R factor - this section is therefore empty.

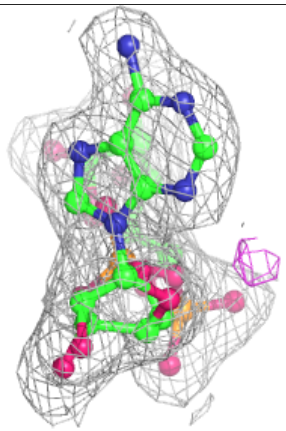
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around AR6 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)

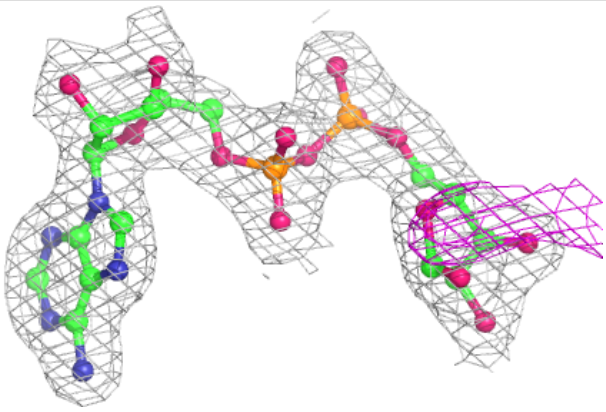


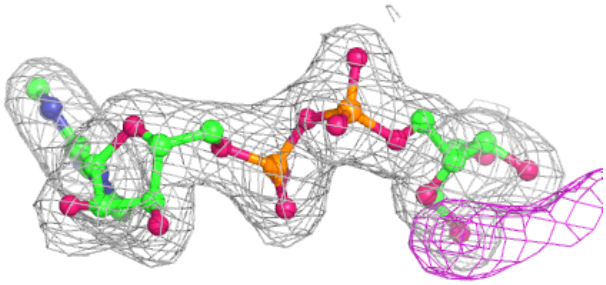


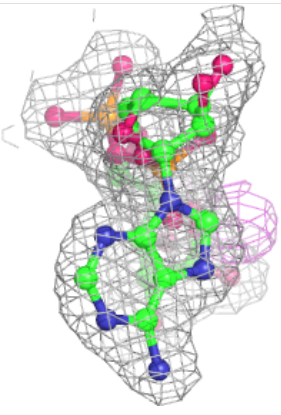


Electron density around AR6 F 301:

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

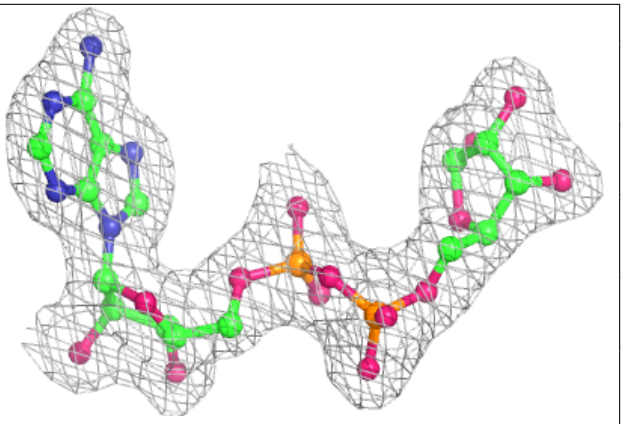


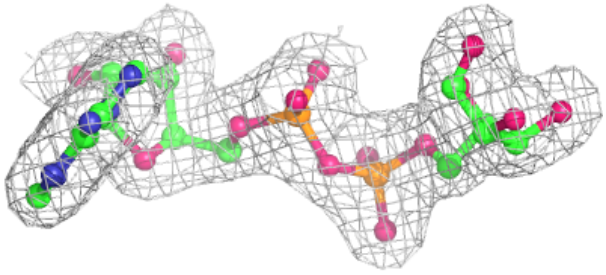


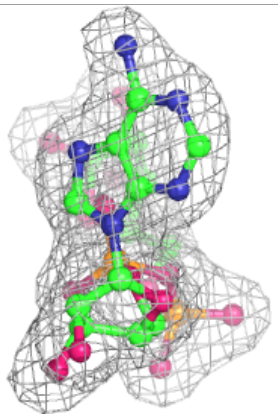


Electron density around AR6 H 301:

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

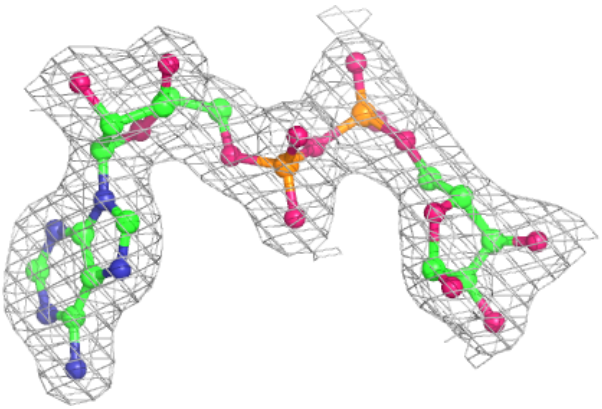


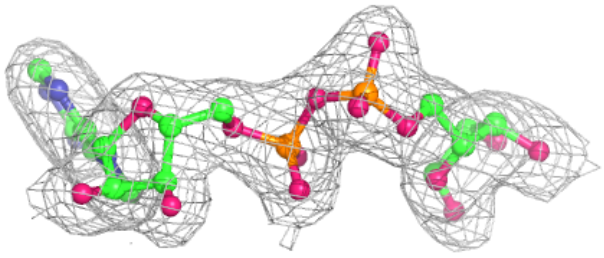


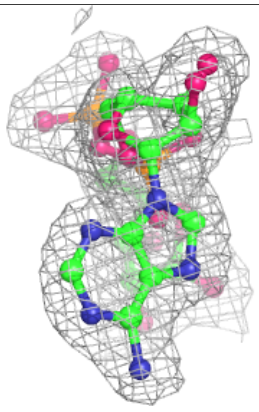


Electron density around AR6 C 301:

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

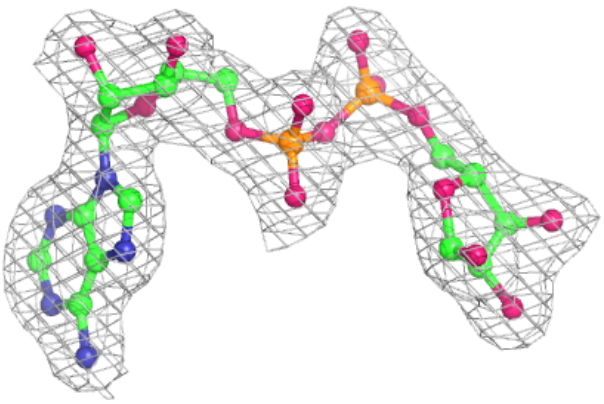


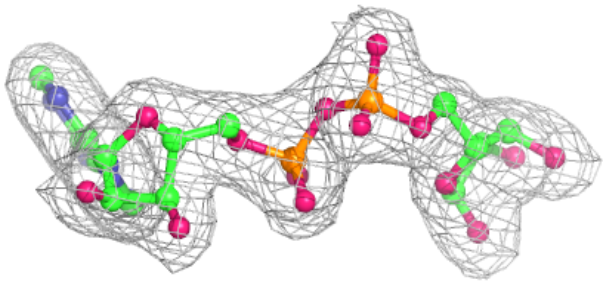


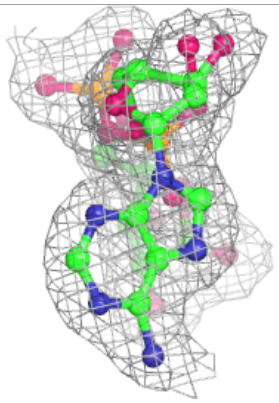


Electron density around AR6 J 301:

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

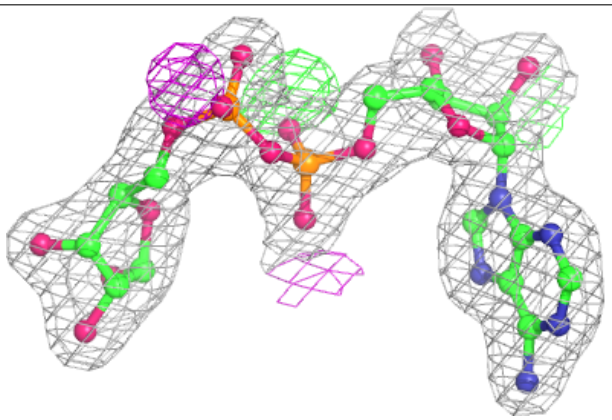


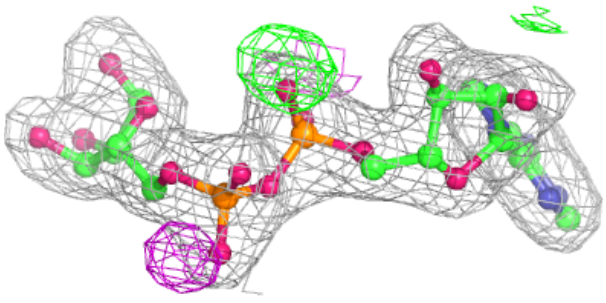


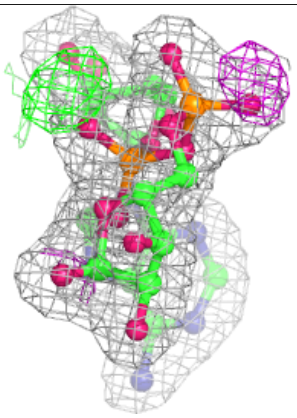


Electron density around AR6 L 301:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.