



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 1, 2021 – 07:27 am BST

PDB ID : 7AKS  
Title : Human ADP-ribosylserine hydrolase ARH3 mutant E41A in complex with H2B-S7-mar peptide  
Authors : Ariza, A.  
Deposited on : 2020-10-02  
Resolution : 1.86 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](mailto: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 following versions of software and data (see [references ⓘ](#)) 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	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

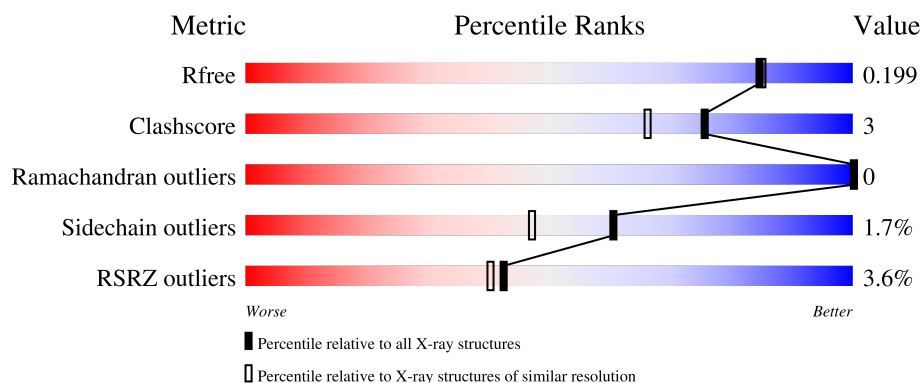
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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.



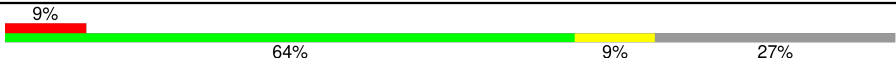


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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  $\geq 3$ , 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	AAA	349	<div> <div>5%</div> <div>88%</div> <div>5% • 6%</div> </div>
1	CCC	349	<div> <div>%</div> <div>88%</div> <div>7% •</div> </div>
1	EEE	349	<div> <div>4%</div> <div>87%</div> <div>7% • 5%</div> </div>
1	GGG	349	<div> <div>3%</div> <div>89%</div> <div>6% 5%</div> </div>
2	BaB	11	<div> <div>9%</div> <div>64%</div> <div>9% 27%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	DaD	11	
2	FaF	11	
2	HaH	11	

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
5	ACT	GGG	407	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11678 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 ADP-ribose glycohydrolase ARH3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	327	Total	C	N	O	S	0	7	0
			2536	1596	433	493	14			
1	CCC	334	Total	C	N	O	S	0	11	0
			2614	1643	443	513	15			
1	EEE	330	Total	C	N	O	S	0	13	0
			2605	1636	449	505	15			
1	GGG	333	Total	C	N	O	S	0	9	0
			2595	1633	440	508	14			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	15	GLY	-	expression tag	UNP Q9NX46
AAA	16	PRO	-	expression tag	UNP Q9NX46
AAA	17	GLY	-	expression tag	UNP Q9NX46
AAA	18	SER	-	expression tag	UNP Q9NX46
AAA	41	ALA	GLU	engineered mutation	UNP Q9NX46
CCC	15	GLY	-	expression tag	UNP Q9NX46
CCC	16	PRO	-	expression tag	UNP Q9NX46
CCC	17	GLY	-	expression tag	UNP Q9NX46
CCC	18	SER	-	expression tag	UNP Q9NX46
CCC	41	ALA	GLU	engineered mutation	UNP Q9NX46
EEE	15	GLY	-	expression tag	UNP Q9NX46
EEE	16	PRO	-	expression tag	UNP Q9NX46
EEE	17	GLY	-	expression tag	UNP Q9NX46
EEE	18	SER	-	expression tag	UNP Q9NX46
EEE	41	ALA	GLU	engineered mutation	UNP Q9NX46
GGG	15	GLY	-	expression tag	UNP Q9NX46
GGG	16	PRO	-	expression tag	UNP Q9NX46
GGG	17	GLY	-	expression tag	UNP Q9NX46
GGG	18	SER	-	expression tag	UNP Q9NX46
GGG	41	ALA	GLU	engineered mutation	UNP Q9NX46

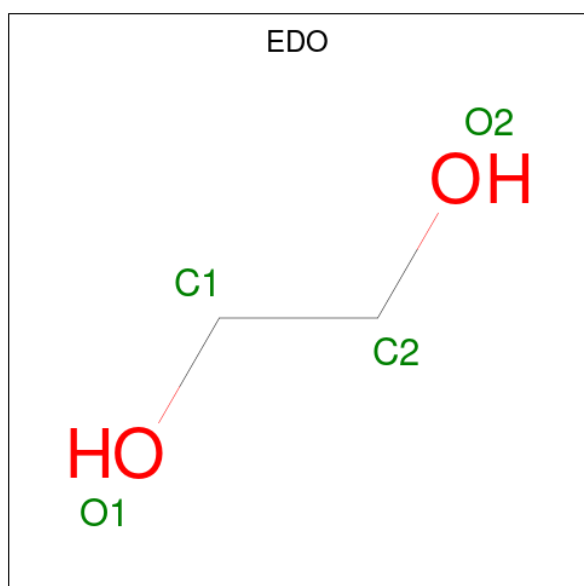
- Molecule 2 is a protein called modified peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	BaB	8	Total	C	N	O	0	0	0
			54	35	9	10			
2	DaD	8	Total	C	N	O	0	0	0
			54	35	9	10			
2	FaF	7	Total	C	N	O	0	0	0
			47	30	8	9			
2	HaH	7	Total	C	N	O	0	0	0
			47	30	8	9			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	2	Total	Mg	0	0
			2	2		
3	CCC	3	Total	Mg	0	0
			3	3		
3	EEE	2	Total	Mg	0	0
			2	2		
3	GGG	2	Total	Mg	0	0
			2	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	C	O	0	0
			4	2	2		
4	AAA	1	Total	C	O	0	0
			4	2	2		
4	CCC	1	Total	C	O	0	0
			4	2	2		
4	CCC	1	Total	C	O	0	0
			4	2	2		
4	CCC	1	Total	C	O	0	0
			4	2	2		
4	CCC	1	Total	C	O	0	0
			4	2	2		
4	CCC	1	Total	C	O	0	0
			4	2	2		
4	EEE	1	Total	C	O	0	0
			4	2	2		
4	EEE	1	Total	C	O	0	0
			4	2	2		
4	EEE	1	Total	C	O	0	0
			4	2	2		
4	EEE	1	Total	C	O	0	0
			4	2	2		
4	EEE	1	Total	C	O	0	0
			4	2	2		
4	GGG	1	Total	C	O	0	0
			4	2	2		
4	GGG	1	Total	C	O	0	0
			4	2	2		
4	GGG	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	1
			8	4	4		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		
5	CCC	1	Total	C	O	0	0
			4	2	2		

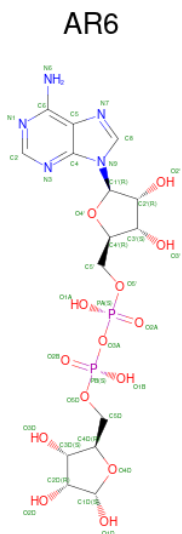
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		
5	GGG	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 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<sub>15</sub>H<sub>23</sub>N<sub>5</sub>O<sub>14</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	BaB	1	Total 35	C 15	N 5	O 13	P 2	0	0
6	DaD	1	Total 35	C 15	N 5	O 13	P 2	0	0
6	FaF	1	Total 35	C 15	N 5	O 13	P 2	0	0
6	HaH	1	Total 35	C 15	N 5	O 13	P 2	0	0

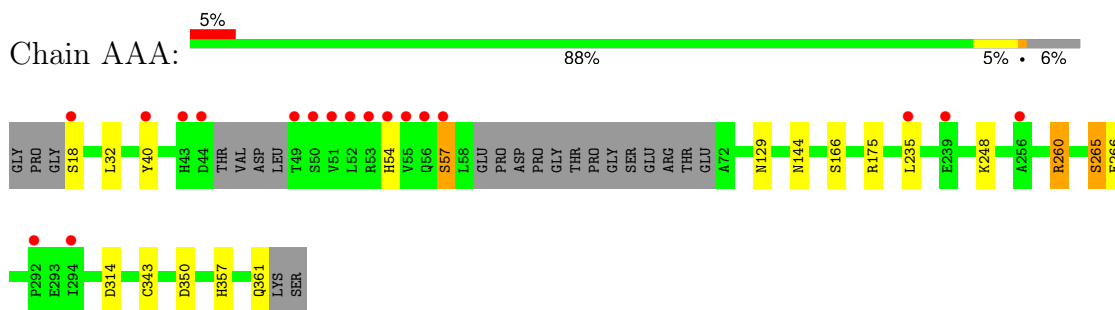
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	AAA	142	Total O 149 149	0	7
7	BaB	8	Total O 9 9	0	1
7	CCC	209	Total O 216 216	0	7
7	DaD	9	Total O 9 9	0	0
7	EEE	196	Total O 206 206	0	10
7	FaF	7	Total O 7 7	0	0
7	GGG	179	Total O 181 181	0	2
7	HaH	8	Total O 8 8	0	0

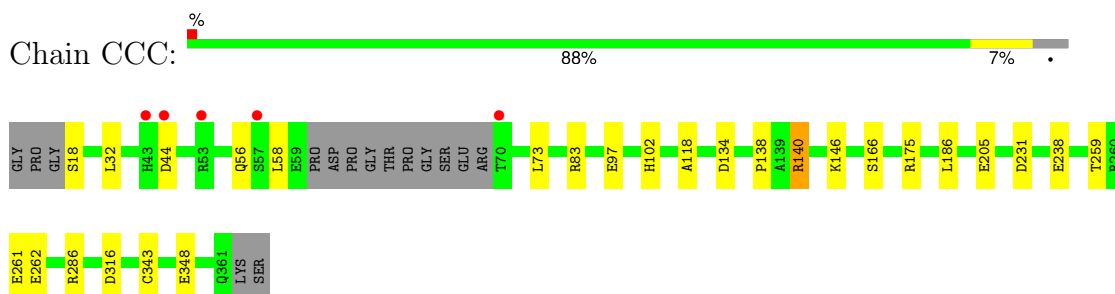
### 3 Residue-property plots [i](#)

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.

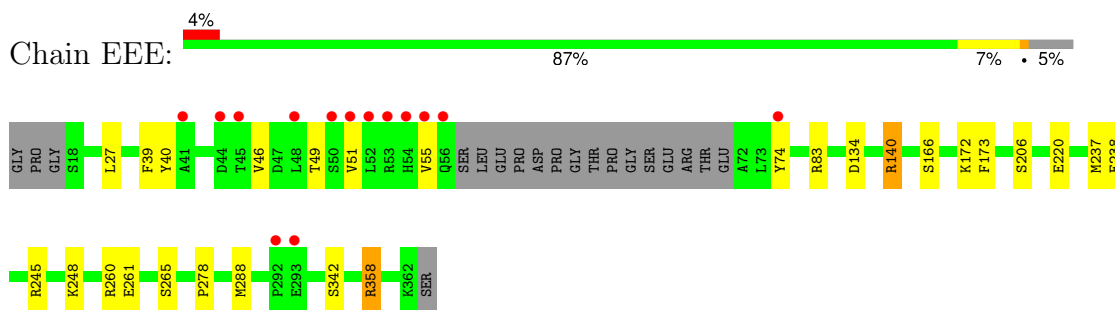
- Molecule 1: ADP-ribose glycohydrolase ARH3



- Molecule 1: ADP-ribose glycohydrolase ARH3

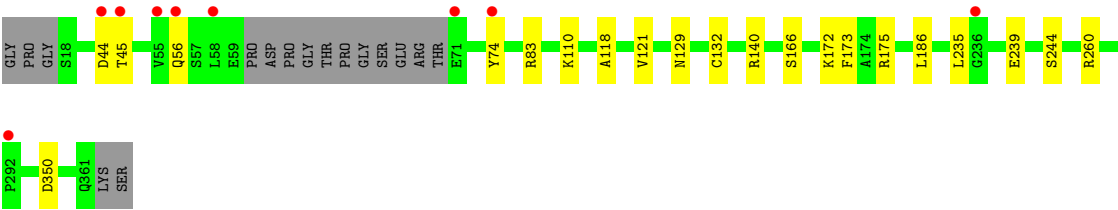


- Molecule 1: ADP-ribose glycohydrolase ARH3

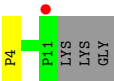


- Molecule 1: ADP-ribose glycohydrolase ARH3

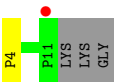




● Molecule 2: modified peptide



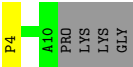
● Molecule 2: modified peptide



● Molecule 2: modified peptide



● Molecule 2: modified peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.36Å 158.64Å 74.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	97.57 – 1.86 97.38 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.2 (97.57-1.86) 99.2 (97.38-1.85)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.56 (at 1.84Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.169 , 0.195 0.178 , 0.199	Depositor DCC
$R_{free}$ test set	6145 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.5	Xtrriage
Anisotropy	0.411	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11678	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, N7P, AR6, MG, ACT

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	0.74	0/2601	0.83	6/3511 (0.2%)
1	CCC	0.78	1/2686 (0.0%)	0.83	4/3628 (0.1%)
1	EEE	0.77	0/2680	0.83	4/3616 (0.1%)
1	GGG	0.75	0/2664	0.81	2/3599 (0.1%)
2	BaB	0.64	0/45	0.96	0/61
2	DaD	0.90	0/45	0.92	0/61
2	FaF	0.79	0/37	1.15	0/49
2	HaH	0.83	0/37	1.17	0/49
All	All	0.76	1/10795 (0.0%)	0.83	16/14574 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	CCC	97	GLU	CD-OE1	7.06	1.33	1.25

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	CCC	140	ARG	NE-CZ-NH2	-7.74	116.43	120.30
1	EEE	140	ARG	NE-CZ-NH2	-6.92	116.84	120.30
1	GGG	83	ARG	NE-CZ-NH2	-6.60	117.00	120.30
1	AAA	260[A]	ARG	NE-CZ-NH1	6.31	123.46	120.30
1	AAA	260[B]	ARG	NE-CZ-NH1	6.31	123.46	120.30

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	2536	0	2490	10	0
1	CCC	2614	0	2566	17	0
1	EEE	2605	0	2567	22	0
1	GGG	2595	0	2549	16	0
2	BaB	54	0	46	0	0
2	DaD	54	0	46	0	0
2	FaF	47	0	39	0	0
2	HaH	47	0	39	0	0
3	AAA	2	0	0	0	0
3	CCC	3	0	0	0	0
3	EEE	2	0	0	0	0
3	GGG	2	0	0	0	0
4	AAA	12	0	18	0	0
4	CCC	24	0	36	4	0
4	EEE	28	0	42	3	0
4	GGG	12	0	18	0	0
5	AAA	24	0	18	2	0
5	CCC	36	0	27	2	0
5	EEE	28	0	21	1	0
5	GGG	28	0	21	4	0
6	BaB	35	0	18	0	0
6	DaD	35	0	18	0	0
6	FaF	35	0	18	0	0
6	HaH	35	0	18	0	0
7	AAA	149	0	0	2	0
7	BaB	9	0	0	0	0
7	CCC	216	0	0	4	0
7	DaD	9	0	0	0	0
7	EEE	206	0	0	6	0
7	FaF	7	0	0	0	0
7	GGG	181	0	0	5	0
7	HaH	8	0	0	0	0
All	All	11678	0	10615	61	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 61 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:CCC:175:ARG:HH22	5:CCC:416:ACT:H2	1.52	0.74
5:CCC:410:ACT:H1	7:CCC:568:HOH:O	1.88	0.72
1:CCC:102:HIS:HD2	7:CCC:698:HOH:O	1.82	0.62
1:GGG:74[B]:TYR:CD2	7:GGG:649:HOH:O	2.51	0.62
1:GGG:110:LYS:HD3	7:GGG:575:HOH:O	2.00	0.60

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	328/349 (94%)	323 (98%)	5 (2%)	0	100	100
1	CCC	341/349 (98%)	335 (98%)	6 (2%)	0	100	100
1	EEE	339/349 (97%)	332 (98%)	7 (2%)	0	100	100
1	GGG	338/349 (97%)	330 (98%)	8 (2%)	0	100	100
2	BaB	6/11 (54%)	6 (100%)	0	0	100	100
2	DaD	6/11 (54%)	6 (100%)	0	0	100	100
2	FaF	5/11 (46%)	5 (100%)	0	0	100	100
2	HaH	5/11 (46%)	5 (100%)	0	0	100	100
All	All	1368/1440 (95%)	1342 (98%)	26 (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	AAA	268/279 (96%)	262 (98%)	6 (2%)	52	36
1	CCC	279/279 (100%)	275 (99%)	4 (1%)	67	55
1	EEE	277/279 (99%)	272 (98%)	5 (2%)	59	45
1	GGG	276/279 (99%)	272 (99%)	4 (1%)	67	55
2	BaB	4/6 (67%)	4 (100%)	0	100	100
2	DaD	4/6 (67%)	4 (100%)	0	100	100
2	FaF	3/6 (50%)	3 (100%)	0	100	100
2	HaH	3/6 (50%)	3 (100%)	0	100	100
All	All	1114/1140 (98%)	1095 (98%)	19 (2%)	60	47

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	EEE	358[B]	ARG
1	GGG	235	LEU
1	GGG	350	ASP
1	GGG	166	SER
1	CCC	166	SER

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 ⓘ

4 non-standard protein/DNA/RNA residues 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	N7P	FaF	4	2	9,10,11	1.63	4 (44%)	9,13,15	1.63	1 (11%)
2	N7P	DaD	4	2	9,10,11	1.94	1 (11%)	9,13,15	1.19	1 (11%)
2	N7P	HaH	4	2	9,10,11	1.81	3 (33%)	9,13,15	1.73	2 (22%)
2	N7P	BaB	4	2	9,10,11	1.49	1 (11%)	9,13,15	1.69	3 (33%)

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	N7P	FaF	4	2	-	0/4/16/18	0/1/1/1
2	N7P	DaD	4	2	-	0/4/16/18	0/1/1/1
2	N7P	HaH	4	2	-	0/4/16/18	0/1/1/1
2	N7P	BaB	4	2	-	0/4/16/18	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	DaD	4	N7P	CA-N	-5.40	1.43	1.48
2	BaB	4	N7P	CA-N	-4.15	1.44	1.48
2	HaH	4	N7P	CA-N	-3.12	1.45	1.48
2	HaH	4	N7P	C1-N	-2.97	1.31	1.35
2	FaF	4	N7P	CD-N	-2.72	1.42	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	FaF	4	N7P	CB-CA-C	-3.99	107.21	112.70
2	HaH	4	N7P	C2-C1-N	-3.28	114.43	117.87
2	BaB	4	N7P	O1-C1-N	2.72	123.61	120.64
2	HaH	4	N7P	CB-CG-CD	2.54	112.20	104.98
2	DaD	4	N7P	CD-N-C1	2.42	131.17	125.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 61 ligands modelled in this entry, 9 are monoatomic - leaving 52 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
5	ACT	AAA	406	-	1,3,3	2.42	1 (100%)	0,3,3	-	-
5	ACT	CCC	414	-	1,3,3	4.61	1 (100%)	0,3,3	-	-
4	EDO	CCC	407	-	3,3,3	0.42	0	2,2,2	0.14	0
5	ACT	AAA	411	-	1,3,3	4.99	1 (100%)	0,3,3	-	-
5	ACT	CCC	413[A]	-	1,3,3	4.22	1 (100%)	0,3,3	-	-
5	ACT	GGG	406	-	1,3,3	4.54	1 (100%)	0,3,3	-	-
6	AR6	FaF	201	3,2	33,38,39	0.89	0	37,58,60	1.10	2 (5%)
5	ACT	GGG	411	-	1,3,3	3.98	1 (100%)	0,3,3	-	-
4	EDO	CCC	405	3	3,3,3	0.40	0	2,2,2	0.30	0
4	EDO	EEE	409	-	3,3,3	0.22	0	2,2,2	0.34	0
5	ACT	CCC	416	-	1,3,3	4.55	1 (100%)	0,3,3	-	-
4	EDO	CCC	409	-	3,3,3	0.30	0	2,2,2	0.39	0
4	EDO	AAA	404	-	3,3,3	0.47	0	2,2,2	0.58	0
5	ACT	AAA	409	-	1,3,3	4.82	1 (100%)	0,3,3	-	-
5	ACT	GGG	412	-	1,3,3	4.75	1 (100%)	0,3,3	-	-
4	EDO	GGG	404	-	3,3,3	0.08	0	2,2,2	0.54	0
4	EDO	AAA	403	-	3,3,3	0.49	0	2,2,2	0.16	0
4	EDO	CCC	408	-	3,3,3	0.79	0	2,2,2	0.77	0
4	EDO	EEE	404	-	3,3,3	0.16	0	2,2,2	0.39	0
5	ACT	GGG	408	-	1,3,3	5.22	1 (100%)	0,3,3	-	-
4	EDO	EEE	405	-	3,3,3	0.72	0	2,2,2	0.50	0
4	EDO	GGG	403	-	3,3,3	0.42	0	2,2,2	0.14	0
6	AR6	DaD	201	3,2	33,38,39	1.01	2 (6%)	37,58,60	1.25	2 (5%)
5	ACT	GGG	410	-	1,3,3	4.24	1 (100%)	0,3,3	-	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	EEE	403	-	3,3,3	0.31	0	2,2,2	0.07	0
4	EDO	EEE	406	-	3,3,3	0.21	0	2,2,2	0.15	0
5	ACT	CCC	417	-	1,3,3	4.43	1 (100%)	0,3,3	-	-
4	EDO	EEE	407	-	3,3,3	0.42	0	2,2,2	0.47	0
5	ACT	EEE	412	-	1,3,3	2.91	1 (100%)	0,3,3	-	-
5	ACT	CCC	411	-	1,3,3	3.52	1 (100%)	0,3,3	-	-
5	ACT	AAA	407	-	1,3,3	3.55	1 (100%)	0,3,3	-	-
4	EDO	CCC	406	-	3,3,3	0.15	0	2,2,2	0.55	0
5	ACT	GGG	407	-	1,3,3	2.81	1 (100%)	0,3,3	-	-
4	EDO	CCC	404	-	3,3,3	0.45	0	2,2,2	0.37	0
4	EDO	AAA	405	-	3,3,3	0.08	0	2,2,2	0.21	0
5	ACT	CCC	412	-	1,3,3	2.76	1 (100%)	0,3,3	-	-
6	AR6	HaH	201	3,2	33,38,39	0.77	0	37,58,60	1.44	4 (10%)
5	ACT	EEE	416	-	1,3,3	5.11	1 (100%)	0,3,3	-	-
6	AR6	BaB	201	3,2	33,38,39	0.80	1 (3%)	37,58,60	1.31	2 (5%)
5	ACT	CCC	415	-	1,3,3	5.22	1 (100%)	0,3,3	-	-
5	ACT	CCC	410	-	1,3,3	3.88	1 (100%)	0,3,3	-	-
4	EDO	GGG	405	-	3,3,3	0.44	0	2,2,2	0.34	0
5	ACT	EEE	413	-	1,3,3	2.96	1 (100%)	0,3,3	-	-
5	ACT	CCC	413[B]	-	1,3,3	3.59	1 (100%)	0,3,3	-	-
5	ACT	GGG	409	-	1,3,3	4.90	1 (100%)	0,3,3	-	-
5	ACT	EEE	414	-	1,3,3	4.11	1 (100%)	0,3,3	-	-
4	EDO	EEE	408	-	3,3,3	0.80	0	2,2,2	0.74	0
5	ACT	AAA	410	-	1,3,3	4.39	1 (100%)	0,3,3	-	-
5	ACT	EEE	411	-	1,3,3	3.82	1 (100%)	0,3,3	-	-
5	ACT	AAA	408	-	1,3,3	3.64	1 (100%)	0,3,3	-	-
5	ACT	EEE	415	-	1,3,3	3.66	1 (100%)	0,3,3	-	-
5	ACT	EEE	410	-	1,3,3	2.64	1 (100%)	0,3,3	-	-

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
4	EDO	CCC	407	-	-	1/1/1/1	-
6	AR6	FaF	201	3,2	-	1/18/51/54	0/4/4/4
4	EDO	CCC	405	3	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	EEE	409	-	-	1/1/1/1	-
4	EDO	CCC	409	-	-	0/1/1/1	-
4	EDO	AAA	404	-	-	0/1/1/1	-
4	EDO	GGG	404	-	-	0/1/1/1	-
4	EDO	AAA	403	-	-	0/1/1/1	-
4	EDO	CCC	408	-	-	1/1/1/1	-
4	EDO	EEE	404	-	-	0/1/1/1	-
4	EDO	EEE	405	-	-	1/1/1/1	-
4	EDO	GGG	403	-	-	1/1/1/1	-
6	AR6	DaD	201	3,2	-	1/18/51/54	0/4/4/4
4	EDO	EEE	403	-	-	0/1/1/1	-
4	EDO	EEE	406	-	-	0/1/1/1	-
4	EDO	EEE	407	-	-	1/1/1/1	-
4	EDO	CCC	406	-	-	1/1/1/1	-
4	EDO	CCC	404	-	-	0/1/1/1	-
4	EDO	AAA	405	-	-	0/1/1/1	-
6	AR6	HaH	201	3,2	-	1/18/51/54	0/4/4/4
6	AR6	BaB	201	3,2	-	1/18/51/54	0/4/4/4
4	EDO	GGG	405	-	-	0/1/1/1	-
4	EDO	EEE	408	-	-	1/1/1/1	-

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	CCC	415	ACT	CH3-C	5.22	1.55	1.48
5	GGG	408	ACT	CH3-C	5.22	1.55	1.48
5	EEE	416	ACT	CH3-C	5.11	1.55	1.48
5	AAA	411	ACT	CH3-C	4.99	1.55	1.48
5	GGG	409	ACT	CH3-C	4.90	1.55	1.48

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	HaH	201	AR6	O2D-C2D-C3D	6.28	123.17	111.27
6	DaD	201	AR6	O2D-C2D-C3D	5.01	120.77	111.27
6	BaB	201	AR6	O2D-C2D-C3D	4.96	120.66	111.27
6	FaF	201	AR6	O2D-C2D-C3D	3.54	117.97	111.27
6	HaH	201	AR6	O3D-C3D-C2D	-2.73	105.45	112.04

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

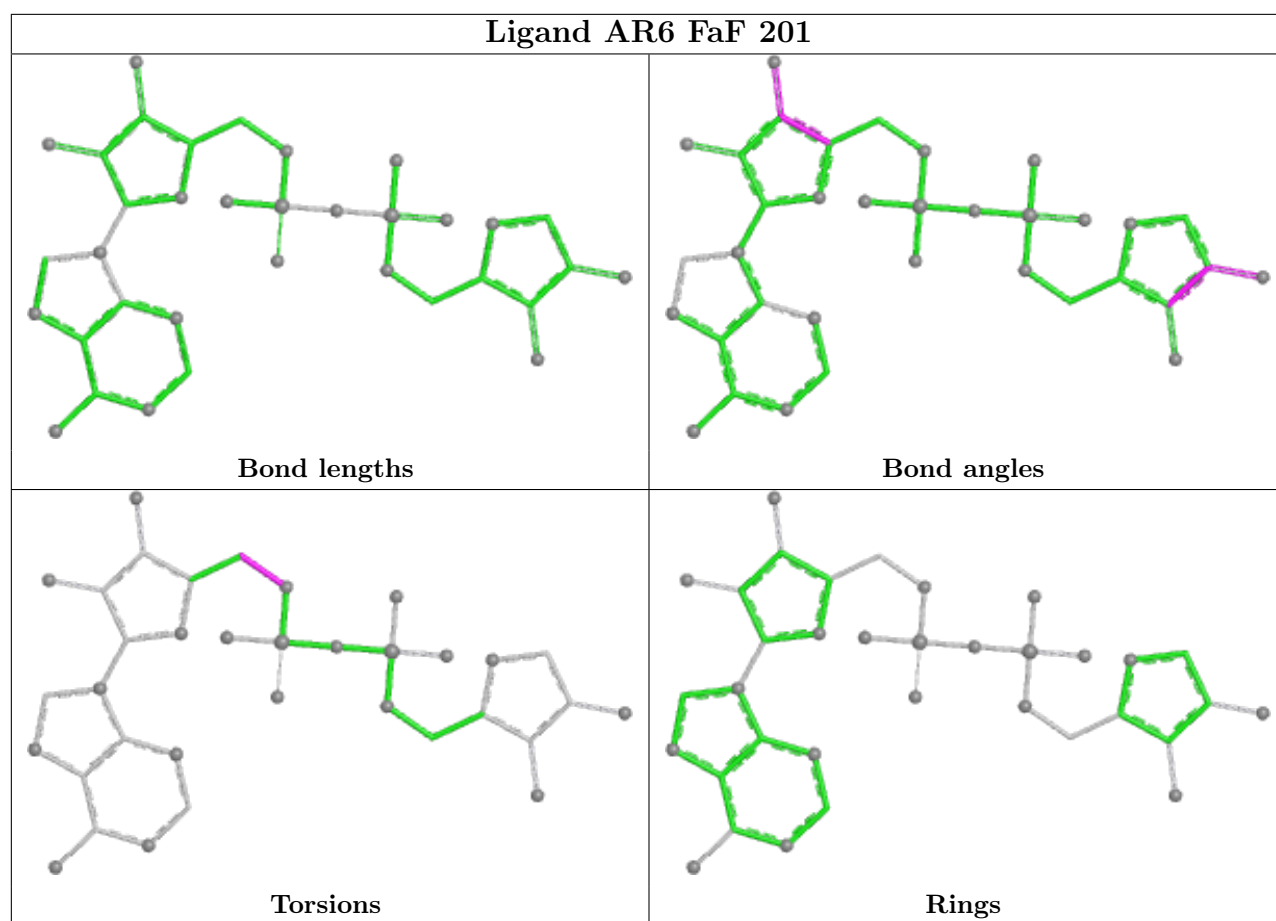
Mol	Chain	Res	Type	Atoms
4	EEE	407	EDO	O1-C1-C2-O2
4	EEE	408	EDO	O1-C1-C2-O2
4	EEE	409	EDO	O1-C1-C2-O2
4	CCC	408	EDO	O1-C1-C2-O2
4	EEE	405	EDO	O1-C1-C2-O2

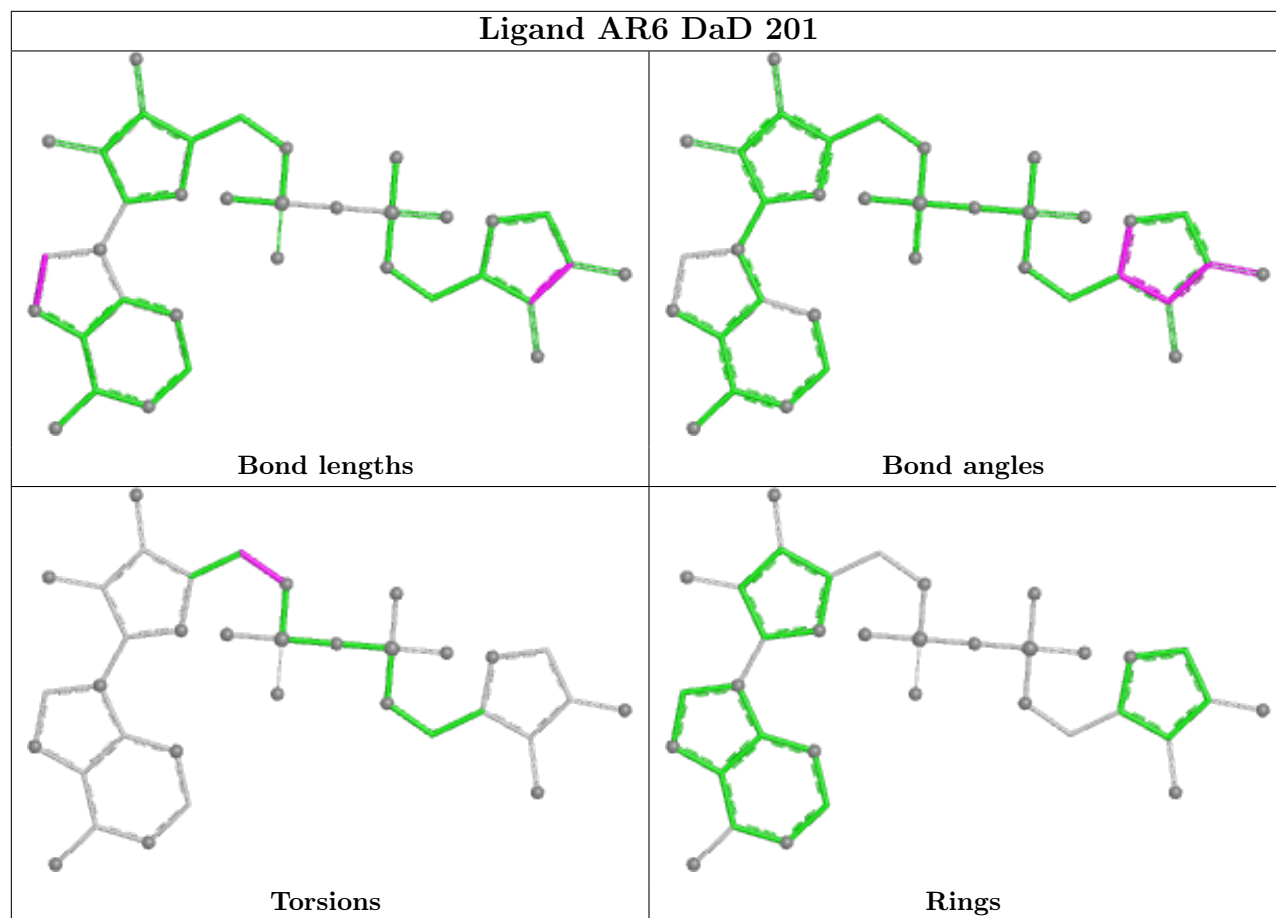
There are no ring outliers.

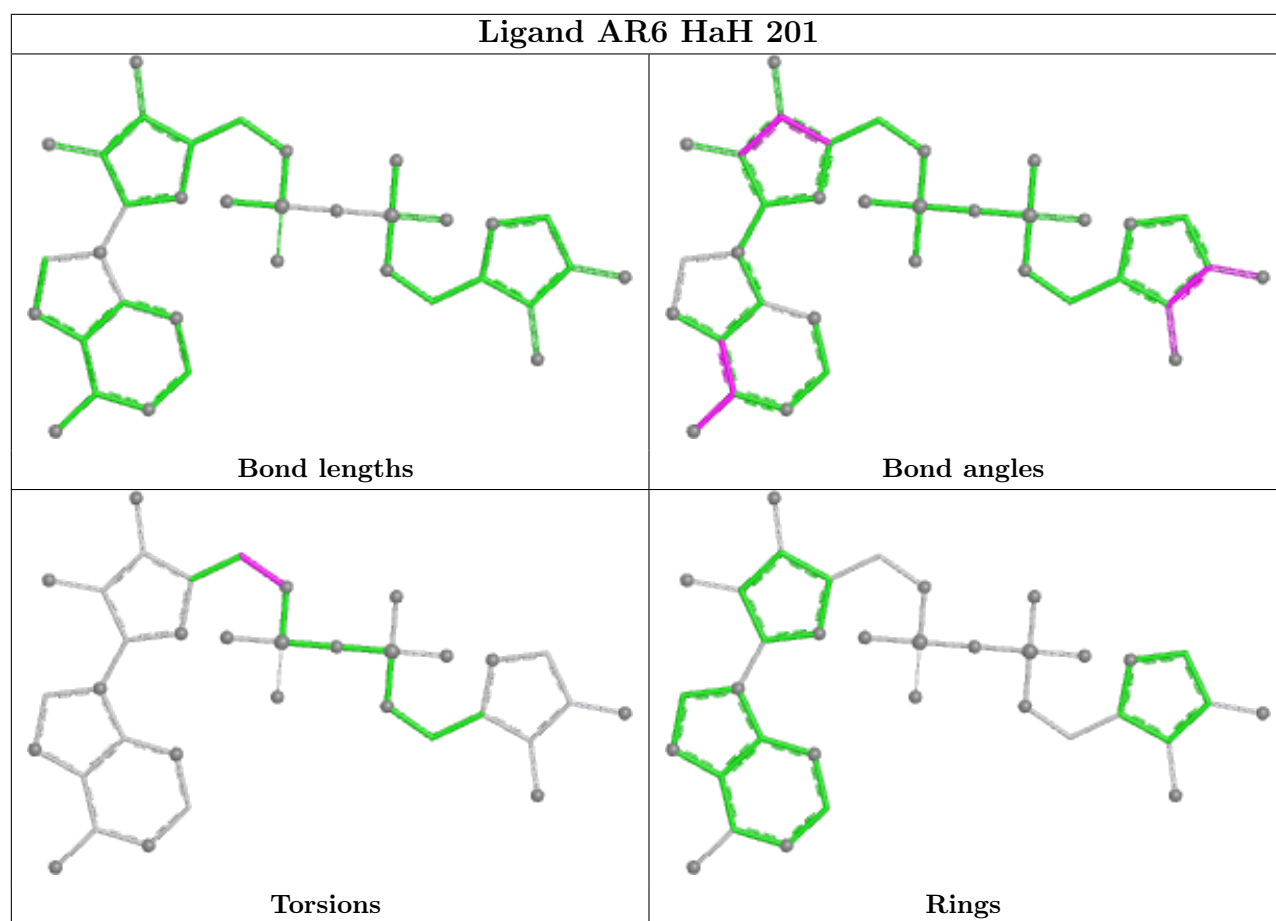
10 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	CCC	407	EDO	1	0
5	AAA	411	ACT	1	0
5	CCC	416	ACT	1	0
5	GGG	412	ACT	1	0
4	CCC	408	EDO	3	0
4	EEE	404	EDO	3	0
5	GGG	407	ACT	3	0
5	CCC	410	ACT	1	0
5	EEE	413	ACT	1	0
5	AAA	408	ACT	1	0

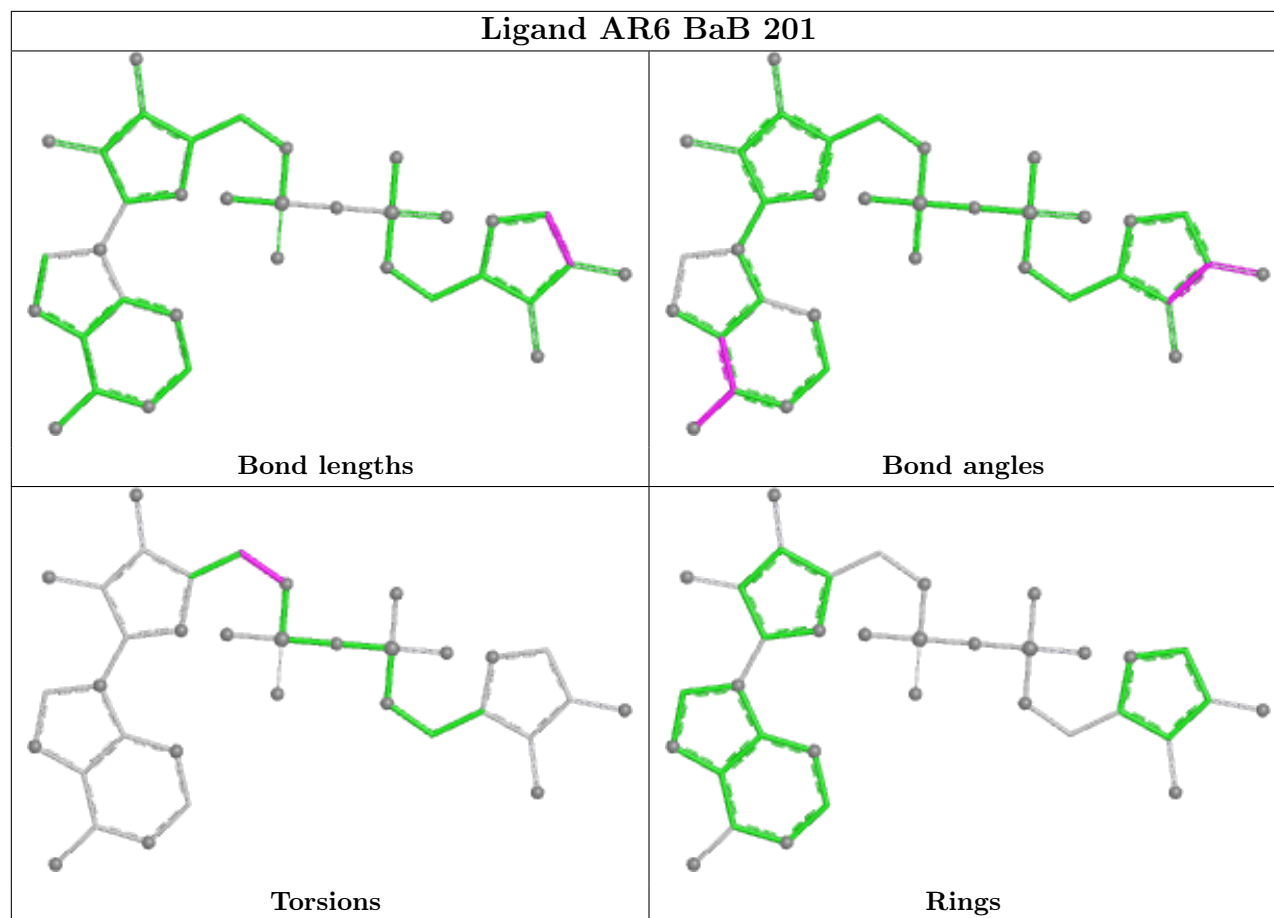
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	DaD	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	DaD	4:N7P	C	5:ALA	N	1.67

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	AAA	327/349 (93%)	-0.10	18 (5%) 25 24	18, 28, 67, 115	0
1	CCC	334/349 (95%)	-0.26	5 (1%) 73 74	15, 24, 54, 78	0
1	EEE	330/349 (94%)	-0.15	14 (4%) 36 34	16, 25, 58, 96	0
1	GGG	333/349 (95%)	-0.16	9 (2%) 54 53	16, 25, 62, 92	0
2	BaB	7/11 (63%)	0.17	1 (14%) 2 2	22, 27, 49, 62	0
2	DaD	7/11 (63%)	0.41	1 (14%) 2 2	23, 27, 50, 64	0
2	FaF	6/11 (54%)	0.00	0 100 100	23, 30, 37, 50	0
2	HaH	6/11 (54%)	-0.43	0 100 100	20, 24, 27, 40	0
All	All	1350/1440 (93%)	-0.16	48 (3%) 42 40	15, 26, 62, 115	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	49	THR	6.2
1	GGG	44	ASP	6.2
1	AAA	44	ASP	5.5
1	EEE	55	VAL	4.8
1	AAA	52	LEU	4.8

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
2	N7P	DaD	4	10/11	0.96	0.08	23,28,29,30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	N7P	HaH	4	10/11	0.96	0.09	21,25,25,26	0
2	N7P	FaF	4	10/11	0.97	0.08	23,29,31,31	0
2	N7P	BaB	4	10/11	0.97	0.08	21,26,27,28	0

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
5	ACT	CCC	415	4/4	0.66	0.25	54,54,56,62	0
5	ACT	CCC	416	4/4	0.69	0.30	47,48,49,52	0
5	ACT	EEE	415	4/4	0.71	0.21	68,70,72,77	0
5	ACT	EEE	410	4/4	0.72	0.28	52,54,55,57	0
4	EDO	EEE	405	4/4	0.72	0.14	41,41,42,48	0
5	ACT	EEE	416	4/4	0.73	0.22	53,56,57,58	0
5	ACT	AAA	411	4/4	0.76	0.25	48,50,52,53	0
4	EDO	CCC	408	4/4	0.77	0.18	37,40,41,44	0
5	ACT	GGG	412	4/4	0.78	0.24	44,45,49,49	0
5	ACT	AAA	408	4/4	0.81	0.26	55,55,57,59	0
5	ACT	EEE	411	4/4	0.83	0.15	31,39,41,45	0
5	ACT	GGG	406	4/4	0.83	0.28	53,59,59,61	0
5	ACT	GGG	410	4/4	0.83	0.21	55,62,64,66	0
5	ACT	GGG	411	4/4	0.83	0.14	47,50,52,52	0
5	ACT	AAA	407	4/4	0.83	0.26	68,70,72,73	0
5	ACT	CCC	414	4/4	0.84	0.12	59,60,62,65	0
5	ACT	GGG	407	4/4	0.85	0.26	58,59,61,63	0
5	ACT	AAA	410	4/4	0.86	0.22	49,58,59,64	0
5	ACT	GGG	409	4/4	0.87	0.17	64,64,66,66	0
4	EDO	EEE	408	4/4	0.87	0.14	42,44,49,51	0
4	EDO	CCC	405	4/4	0.89	0.13	27,31,32,34	0
5	ACT	GGG	408	4/4	0.89	0.13	33,41,41,41	0
5	ACT	CCC	411	4/4	0.89	0.22	43,44,46,46	0
5	ACT	CCC	417	4/4	0.90	0.16	45,54,54,58	0
5	ACT	EEE	412	4/4	0.90	0.16	46,47,49,50	0

*Continued on next page...*

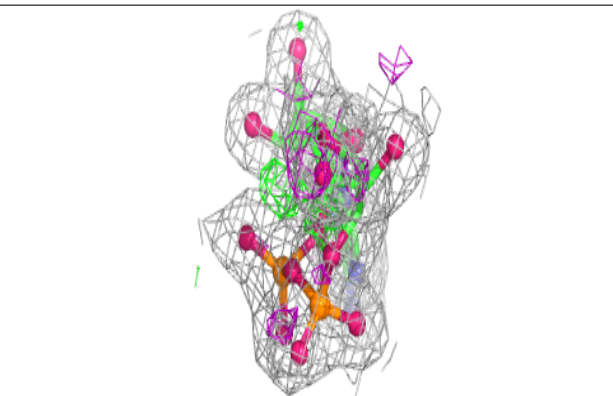
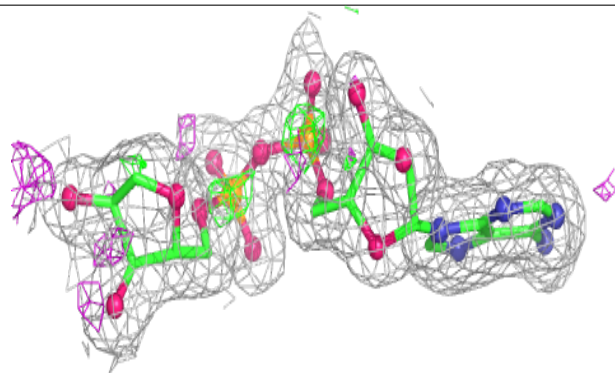
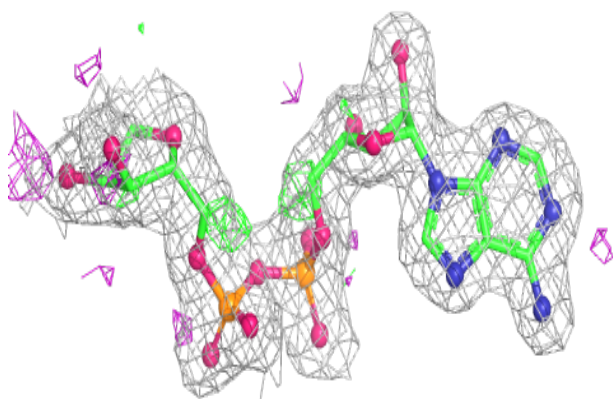
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	ACT	EEE	414	4/4	0.90	0.20	63,63,66,67	0
4	EDO	CCC	407	4/4	0.91	0.18	38,44,47,48	0
4	EDO	EEE	409	4/4	0.91	0.18	51,54,56,56	0
5	ACT	CCC	413[A]	4/4	0.91	0.20	43,45,46,47	4
5	ACT	CCC	413[B]	4/4	0.91	0.20	35,36,36,37	4
4	EDO	AAA	404	4/4	0.91	0.14	31,34,35,38	0
3	MG	CCC	403	1/1	0.91	0.11	37,37,37,37	0
4	EDO	EEE	407	4/4	0.91	0.17	35,42,45,52	0
4	EDO	GGG	404	4/4	0.92	0.10	30,30,30,30	0
5	ACT	CCC	410	4/4	0.92	0.16	34,35,36,37	0
5	ACT	AAA	409	4/4	0.93	0.14	35,41,41,44	0
4	EDO	CCC	404	4/4	0.93	0.11	34,35,35,36	0
4	EDO	GGG	405	4/4	0.94	0.10	33,33,34,36	0
4	EDO	EEE	404	4/4	0.94	0.12	36,36,37,40	0
4	EDO	AAA	405	4/4	0.94	0.12	32,32,32,33	0
4	EDO	GGG	403	4/4	0.95	0.12	39,40,41,44	0
4	EDO	EEE	403	4/4	0.95	0.17	36,38,38,40	0
4	EDO	CCC	406	4/4	0.95	0.22	36,38,39,40	0
5	ACT	AAA	406	4/4	0.95	0.08	46,46,47,50	0
5	ACT	EEE	413	4/4	0.95	0.18	46,53,55,57	0
4	EDO	CCC	409	4/4	0.96	0.11	33,34,35,37	0
3	MG	EEE	402	1/1	0.96	0.08	19,19,19,19	0
4	EDO	EEE	406	4/4	0.97	0.13	36,37,37,38	0
4	EDO	AAA	403	4/4	0.97	0.13	41,41,41,42	0
3	MG	CCC	402	1/1	0.97	0.08	18,18,18,18	0
5	ACT	CCC	412	4/4	0.97	0.07	30,30,32,32	0
6	AR6	DaD	201	35/36	0.97	0.09	20,23,25,27	0
6	AR6	FaF	201	35/36	0.97	0.09	18,23,26,29	0
3	MG	GGG	401	1/1	0.98	0.06	17,17,17,17	0
6	AR6	BaB	201	35/36	0.98	0.08	16,18,23,24	0
3	MG	CCC	401	1/1	0.98	0.11	18,18,18,18	0
3	MG	AAA	401	1/1	0.98	0.07	19,19,19,19	0
6	AR6	HaH	201	35/36	0.98	0.08	16,18,21,22	0
3	MG	EEE	401	1/1	0.99	0.10	18,18,18,18	0
3	MG	GGG	402	1/1	0.99	0.06	18,18,18,18	0
3	MG	AAA	402	1/1	0.99	0.09	20,20,20,20	0

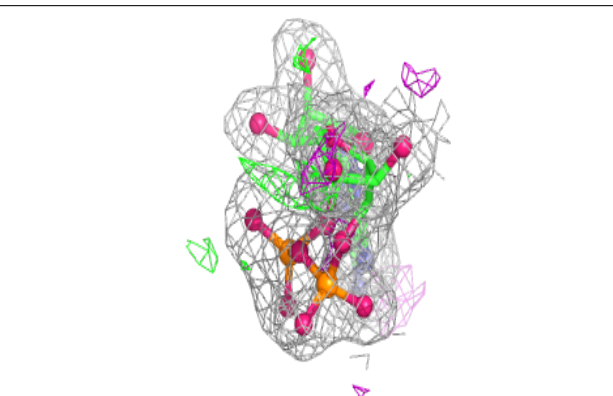
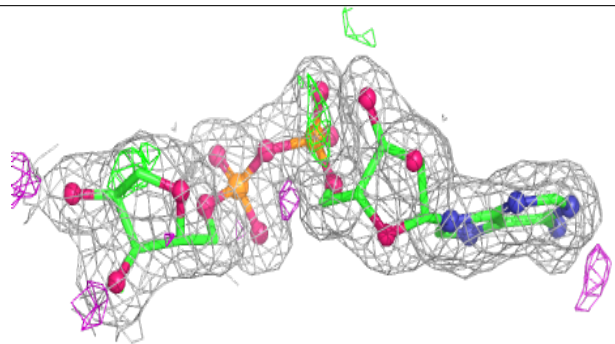
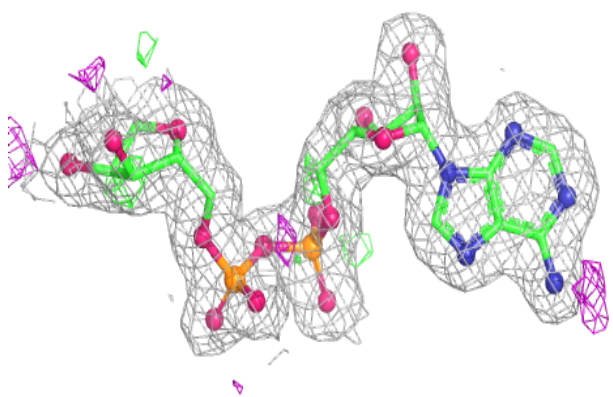
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around AR6 DaD 201:**

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

**Electron density around AR6 FaF 201:**

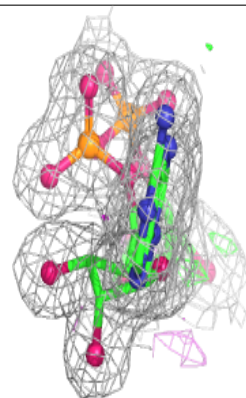
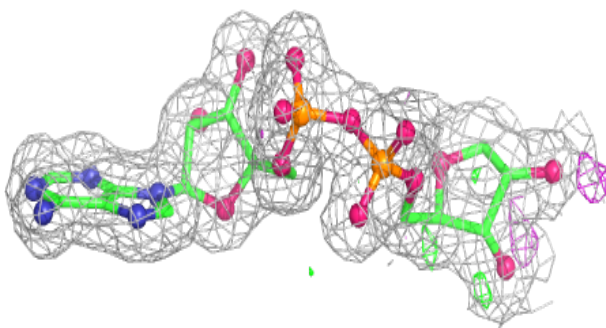
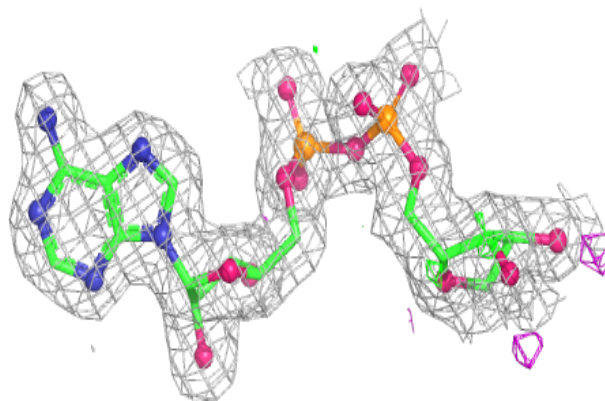
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



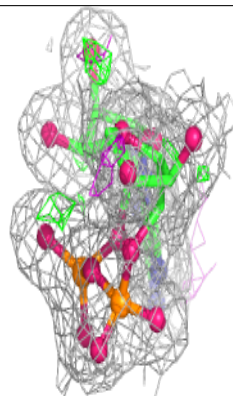
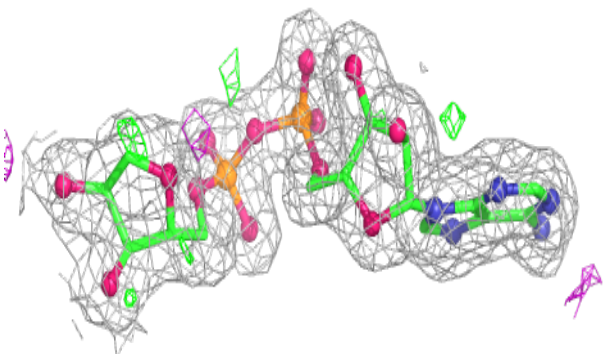
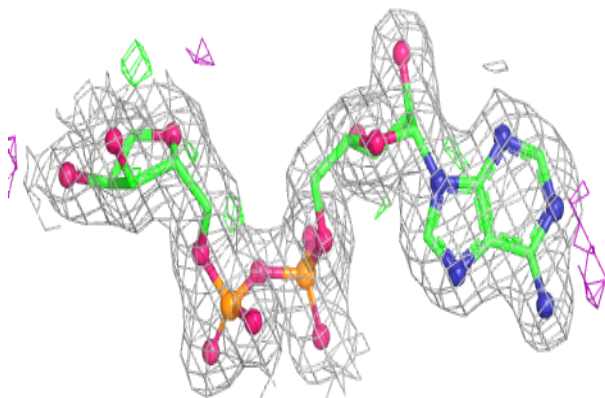


**Electron density around AR6 BaB 201:**

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

**Electron density around AR6 HaH 201:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.