



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 25, 2023 – 03:09 AM EDT

PDB ID : 5WM8  
Title : Structure of the 10R (+)-cis-BP-dG modified Rev1 ternary complex  
Authors : Rechkoblit, O.; Kolbanovsky, A.; Landes, H.; Geacintov, N.E.; Aggarwal, A.K.  
Deposited on : 2017-07-28  
Resolution : 1.92 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

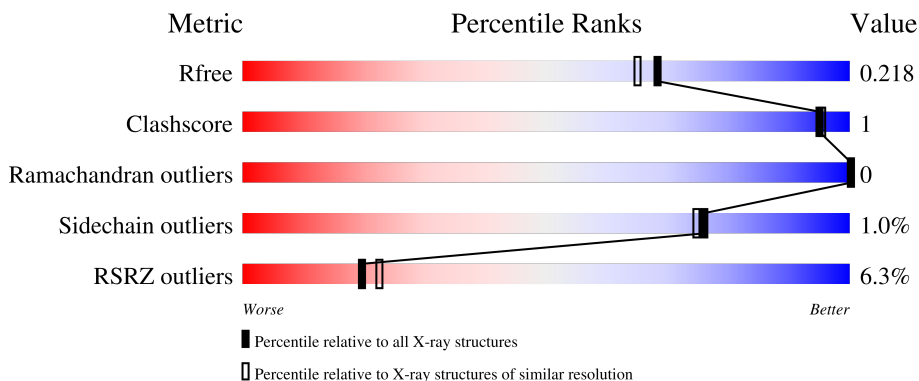
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.92 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	P	12	 17% 92% 8%
2	T	17	 12% 88% 12%
3	A	434	 6% 97% .

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 4514 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 DNA chain called DNA (5'-D(\*GP\*GP\*GP\*GP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*(DDG))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	P	12	274	130	51	80	13	0	2	0

- Molecule 2 is a DNA chain called DNA (5'-D(P\*AP\*TP\*CP\*GP\*CP\*TP\*AP\*CP\*CP\*AP\*CP\*AP\*CP\*CP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	T	17	318	151	56	95	16	0	0	1

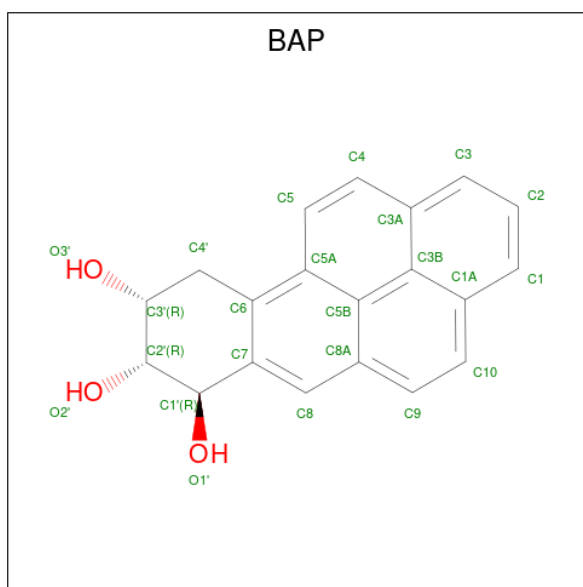
- Molecule 3 is a protein called DNA repair protein REV1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	432	3484	2217	594	647	26	0	9	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

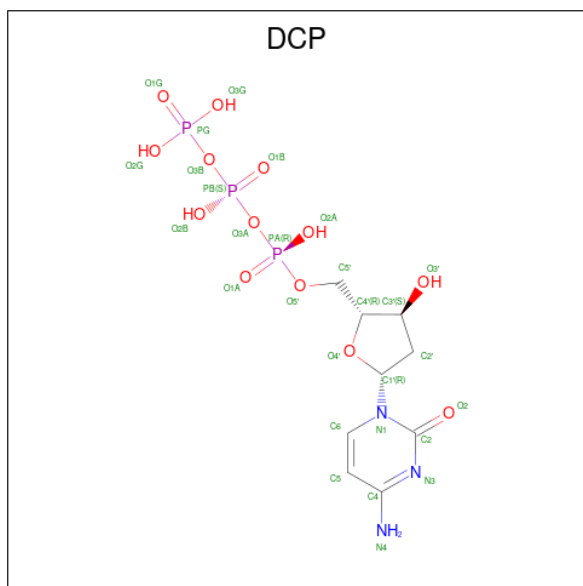
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
4	P	1	1	1	0	0
4	A	3	3	3	0	0

- Molecule 5 is 1,2,3-TRIHIDROXY-1,2,3,4-TETRAHYDROBENZO[A]PYRENE (three-letter code: BAP) (formula: C<sub>20</sub>H<sub>16</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	T	1	Total	C	O	0	0
			23	20	3		

- Molecule 6 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula:  $C_9H_{16}N_3O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			28	9	3	13	3		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



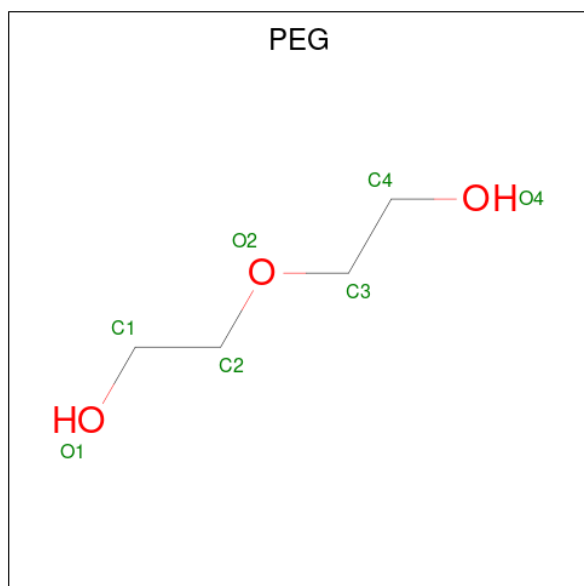
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 8 4 4	0	1
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



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

- Molecule 10 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
10	P	25	Total 25	O 25	0	0
10	T	19	Total 19	O 19	0	0
10	A	267	Total 267	O 267	0	0

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

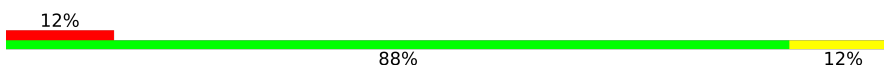
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DNA (5'-D(\*GP\*GP\*GP\*GP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*(DDG))-3')

Chain P: 



- Molecule 2: DNA (5'-D(P\*AP\*TP\*CP\*GP\*CP\*TP\*AP\*CP\*CP\*AP\*CP\*AP\*CP\*CP\*CP\*C)-3')

Chain T: 



- Molecule 3: DNA repair protein REV1

Chain A: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.55Å 180.36Å 54.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.00 – 1.92 43.34 – 1.93	Depositor EDS
% Data completeness (in resolution range)	95.0 (45.00-1.92) 95.0 (43.34-1.93)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 1.92Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.176 , 0.214 0.184 , 0.218	Depositor DCC
$R_{free}$ test set	2290 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.9	Xtrriage
Anisotropy	0.038	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 45.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4514	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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: MG, GOL, DCP, EDO, DDG, BAP, PEG

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	P	0.39	0/284	0.80	0/437
2	T	0.68	1/354 (0.3%)	0.88	0/542
3	A	0.48	0/3569	0.69	2/4818 (0.0%)
All	All	0.49	1/4207 (0.0%)	0.72	2/5797 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	1	DC	O3'-P	-9.92	1.49	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	362	ASP	CB-CG-OD1	-6.21	112.72	118.30
3	A	362	ASP	CB-CG-OD2	5.43	123.19	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	P	274	0	144	0	0
2	T	318	0	178	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	3484	0	3496	9	0
4	A	3	0	0	0	0
4	P	1	0	0	0	0
5	T	23	0	15	0	0
6	A	28	0	12	0	0
7	A	18	0	24	0	0
8	A	40	0	60	1	0
9	A	14	0	20	0	0
10	A	267	0	0	2	0
10	P	25	0	0	0	0
10	T	19	0	0	1	0
All	All	4514	0	3949	10	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 1.

All (10) 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)
3:A:457:ILE:CD1	3:A:489:LEU:HD11	2.16	0.74
3:A:457:ILE:HD12	3:A:489:LEU:HD21	1.82	0.62
3:A:457:ILE:HD12	3:A:489:LEU:HD11	1.90	0.52
3:A:495:GLN:OE1	10:A:901:HOH:O	2.19	0.52
10:T:201:HOH:O	8:A:815:EDO:H22	2.10	0.50
3:A:648[B]:GLN:NE2	10:A:905:HOH:O	2.48	0.46
2:T:2:DA:N1	3:A:307:ILE:HD11	2.31	0.45
3:A:457:ILE:HD13	3:A:489:LEU:HD11	1.96	0.45
3:A:416[B]:MET:HE3	3:A:416[B]:MET:HB3	1.90	0.42
3:A:455:LEU:HD13	3:A:457:ILE:HD11	2.03	0.40

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	
3	A	439/434 (101%)	432 (98%)	7 (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	
3	A	394/391 (101%)	390 (99%)	4 (1%)	76	75

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

Mol	Chain	Res	Type
3	A	324	ARG
3	A	362	ASP
3	A	382	CYS
3	A	722	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	495	GLN
3	A	499	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](#)

2 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
1	DDG	P	12[A]	1	17,23,24	0.95	1 (5%)	15,33,36	1.44	3 (20%)
1	DDG	P	12[B]	1	17,23,24	0.97	1 (5%)	15,33,36	1.75	5 (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
1	DDG	P	12[A]	1	-	0/3/18/19	0/3/3/3
1	DDG	P	12[B]	1	-	0/3/18/19	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	12[A]	DDG	C6-N1	-2.71	1.33	1.37
1	P	12[B]	DDG	C6-N1	-2.71	1.33	1.37

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	12[A]	DDG	O6-C6-C5	-2.88	118.75	124.37
1	P	12[B]	DDG	O6-C6-C5	-2.88	118.75	124.37
1	P	12[B]	DDG	C2'-C3'-C4'	2.68	107.75	102.72
1	P	12[B]	DDG	O4'-C1'-C2'	2.62	109.51	106.67
1	P	12[A]	DDG	C5-C6-N1	2.55	118.45	113.95
1	P	12[B]	DDG	C5-C6-N1	2.55	118.45	113.95
1	P	12[A]	DDG	N2-C2-N1	2.00	120.98	116.71
1	P	12[B]	DDG	N2-C2-N1	2.00	120.98	116.71

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 21 ligands modelled in this entry, 4 are monoatomic - leaving 17 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
8	EDO	A	813	-	3,3,3	0.44	0	2,2,2	0.34	0
8	EDO	A	812[A]	-	3,3,3	0.35	0	2,2,2	0.44	0
6	DCP	A	804	4	25,29,29	0.79	1 (4%)	37,45,45	1.56	9 (24%)
5	BAP	T	101	2	27,27,27	2.19	9 (33%)	39,42,42	0.94	3 (7%)
8	EDO	A	809	-	3,3,3	0.46	0	2,2,2	0.27	0
7	GOL	A	805	-	5,5,5	0.41	0	5,5,5	0.64	0
8	EDO	A	811	-	3,3,3	0.40	0	2,2,2	0.43	0
8	EDO	A	815	-	3,3,3	0.53	0	2,2,2	0.17	0
9	PEG	A	817	-	6,6,6	0.46	0	5,5,5	0.26	0
8	EDO	A	812[B]	-	3,3,3	0.41	0	2,2,2	0.51	0
9	PEG	A	818	-	6,6,6	0.47	0	5,5,5	0.25	0
8	EDO	A	810	-	3,3,3	0.48	0	2,2,2	0.49	0
8	EDO	A	814	-	3,3,3	0.47	0	2,2,2	0.23	0
8	EDO	A	816	-	3,3,3	0.47	0	2,2,2	0.29	0
8	EDO	A	808	-	3,3,3	0.48	0	2,2,2	0.02	0
7	GOL	A	807	-	5,5,5	0.44	0	5,5,5	0.49	0
7	GOL	A	806	-	5,5,5	0.29	0	5,5,5	0.23	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
8	EDO	A	813	-	-	1/1/1/1	-
8	EDO	A	812[A]	-	-	1/1/1/1	-
6	DCP	A	804	4	-	6/22/34/34	0/2/2/2
8	EDO	A	809	-	-	1/1/1/1	-
8	EDO	A	811	-	-	1/1/1/1	-
7	GOL	A	805	-	-	0/4/4/4	-
8	EDO	A	815	-	-	0/1/1/1	-
9	PEG	A	817	-	-	1/4/4/4	-
9	PEG	A	818	-	-	1/4/4/4	-
8	EDO	A	812[B]	-	-	0/1/1/1	-
5	BAP	T	101	2	-	-	0/5/5/5
8	EDO	A	810	-	-	0/1/1/1	-
8	EDO	A	814	-	-	1/1/1/1	-
8	EDO	A	816	-	-	1/1/1/1	-
8	EDO	A	808	-	-	0/1/1/1	-
7	GOL	A	807	-	-	2/4/4/4	-
7	GOL	A	806	-	-	2/4/4/4	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	T	101	BAP	C6-C7	7.30	1.49	1.37
5	T	101	BAP	C6-C5A	3.62	1.49	1.43
5	T	101	BAP	C5B-C8A	2.84	1.48	1.42
5	T	101	BAP	C5A-C5B	2.68	1.48	1.42
5	T	101	BAP	C8-C7	2.46	1.40	1.36
5	T	101	BAP	C3B-C1A	2.45	1.47	1.42
5	T	101	BAP	C7-C1'	-2.41	1.48	1.51
5	T	101	BAP	C3B-C3A	2.35	1.47	1.42
5	T	101	BAP	C10-C9	2.31	1.41	1.35
6	A	804	DCP	C6-C5	2.27	1.40	1.35

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	804	DCP	O3G-PG-O2G	3.50	121.01	107.64
6	A	804	DCP	O2A-PA-O5'	3.12	122.26	107.75
6	A	804	DCP	O2A-PA-O1A	2.79	126.05	112.24
5	T	101	BAP	O1'-C1'-C7	-2.78	105.01	110.49
6	A	804	DCP	O5'-PA-O1A	-2.66	98.67	109.07
6	A	804	DCP	O2B-PB-O1B	2.65	125.36	112.24
6	A	804	DCP	C2'-C1'-N1	-2.56	107.88	113.77

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	T	101	BAP	C4'-C3'-C2'	2.30	113.07	110.30
5	T	101	BAP	C5A-C5B-C8A	2.27	122.93	119.94
6	A	804	DCP	O3B-PG-O1G	-2.22	98.88	111.19
6	A	804	DCP	PB-O3A-PA	-2.10	125.61	132.83
6	A	804	DCP	N4-C4-N3	2.09	121.63	117.97

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	804	DCP	PB-O3B-PG-O3G
7	A	806	GOL	C1-C2-C3-O3
7	A	806	GOL	O2-C2-C3-O3
8	A	809	EDO	O1-C1-C2-O2
8	A	812[A]	EDO	O1-C1-C2-O2
8	A	816	EDO	O1-C1-C2-O2
9	A	817	PEG	O2-C3-C4-O4
8	A	813	EDO	O1-C1-C2-O2
6	A	804	DCP	PA-O3A-PB-O1B
8	A	814	EDO	O1-C1-C2-O2
9	A	818	PEG	O2-C3-C4-O4
6	A	804	DCP	PA-O3A-PB-O2B
7	A	807	GOL	O1-C1-C2-C3
6	A	804	DCP	PB-O3B-PG-O1G
8	A	811	EDO	O1-C1-C2-O2
6	A	804	DCP	PB-O3B-PG-O2G
6	A	804	DCP	C5'-O5'-PA-O3A
7	A	807	GOL	O1-C1-C2-O2

There are no ring outliers.

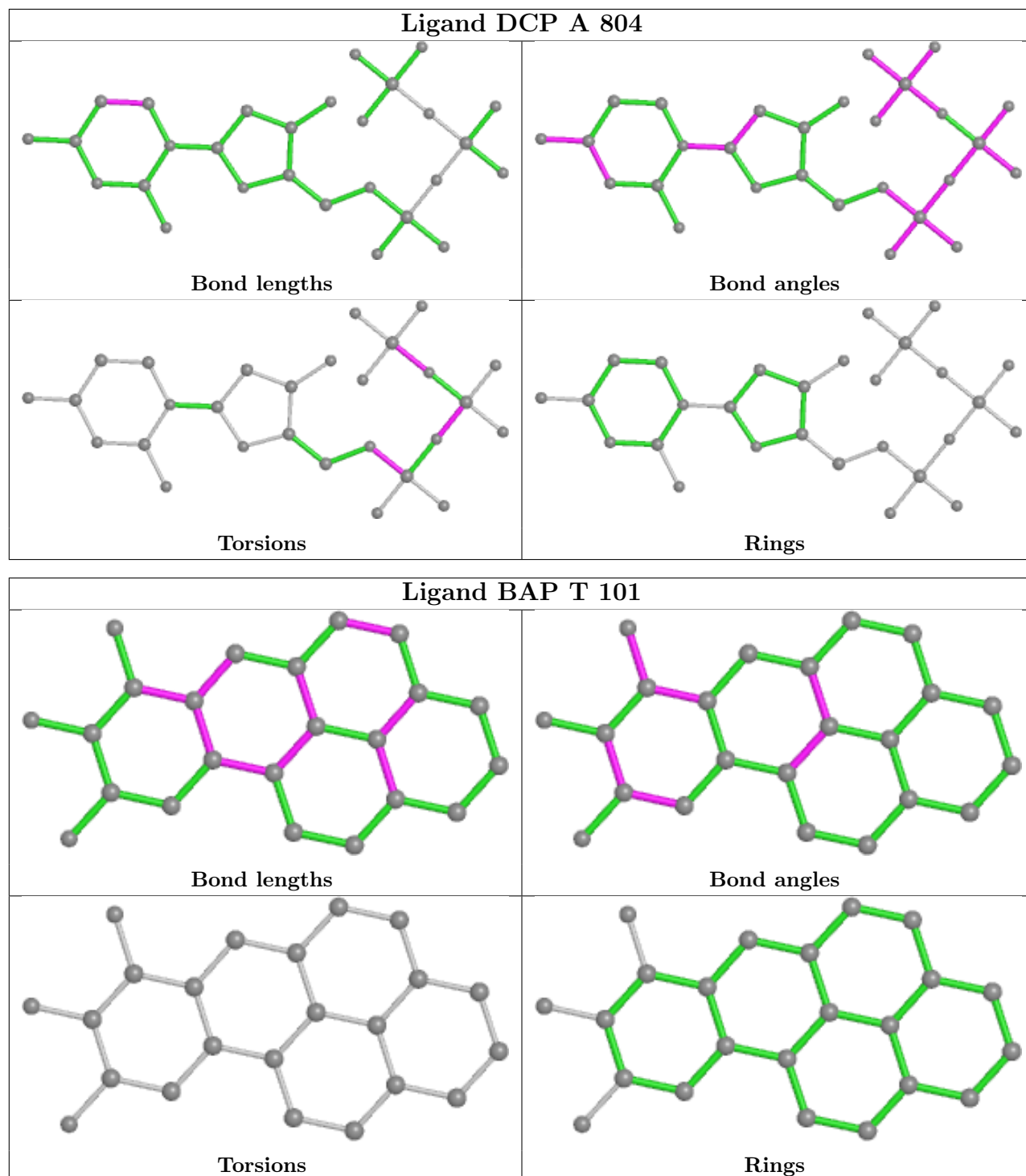
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	815	EDO	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](#)

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<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	P	11/12 (91%)	0.55	2 (18%) <b>1</b> <b>1</b>	26, 41, 54, 54	0
2	T	17/17 (100%)	0.31	2 (11%) <b>4</b> <b>5</b>	20, 41, 55, 59	0
3	A	432/434 (99%)	0.23	25 (5%) <b>23</b> <b>26</b>	16, 29, 51, 84	0
All	All	460/463 (99%)	0.24	29 (6%) <b>20</b> <b>22</b>	16, 29, 54, 84	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	477	ASP	5.0
3	A	347	TYR	4.7
3	A	478	ASN	4.2
3	A	479	ILE	4.2
3	A	699	ILE	3.8
3	A	381	ALA	3.6
2	T	17	DC	3.5
3	A	348	THR	3.3
3	A	377	SER	3.0
3	A	698	GLY	2.9
3	A	738	VAL	2.9
3	A	476	PRO	2.9
3	A	380	SER	2.9
2	T	1	DC	2.8
3	A	345	HIS	2.7
1	P	1	DG	2.6
3	A	540	GLU	2.5
3	A	569	HIS	2.3
3	A	480	HIS	2.3
3	A	376	SER	2.3
3	A	457	ILE	2.3
3	A	677	ILE	2.2
3	A	674	ASP	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	A	378	SER	2.2
3	A	481	ASN	2.2
1	P	2	DG	2.1
3	A	566	ASP	2.1
3	A	349	LYS	2.1
3	A	344	ILE	2.1

## 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
1	DDG	P	12[A]	21/22	0.96	0.16	21,24,31,33	10
1	DDG	P	12[B]	21/22	0.96	0.16	19,22,24,24	10

## 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.

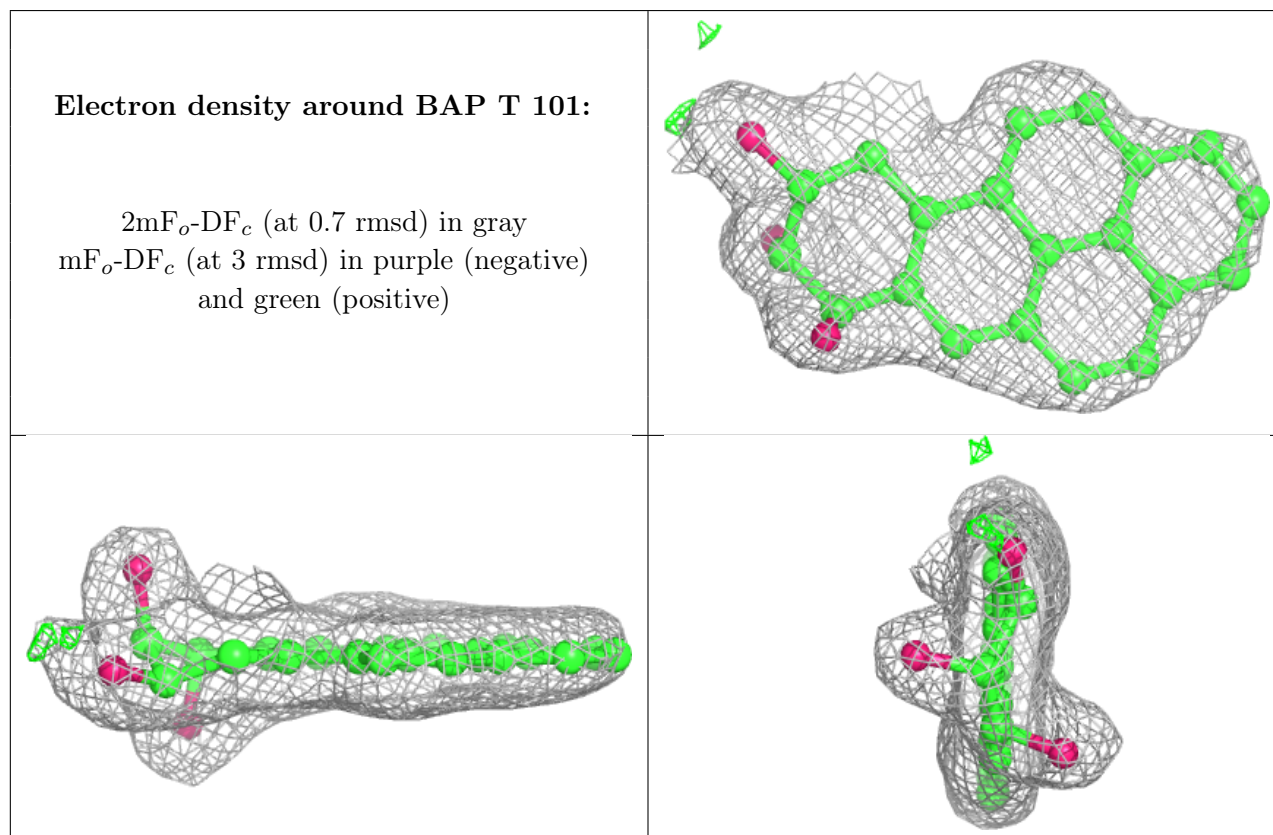
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
8	EDO	A	816	4/4	0.73	0.22	59,61,62,62	0
9	PEG	A	818	7/7	0.74	0.23	69,74,77,77	0
9	PEG	A	817	7/7	0.75	0.14	54,56,57,61	0
8	EDO	A	815	4/4	0.77	0.21	47,49,50,50	0
8	EDO	A	809	4/4	0.86	0.18	52,53,54,54	0
8	EDO	A	814	4/4	0.86	0.42	44,49,49,52	0
4	MG	A	801	1/1	0.86	0.15	29,29,29,29	0
8	EDO	A	812[B]	4/4	0.89	0.24	45,46,47,48	4
8	EDO	A	812[A]	4/4	0.89	0.24	25,26,27,28	4
5	BAP	T	101	23/23	0.90	0.13	30,40,43,44	0

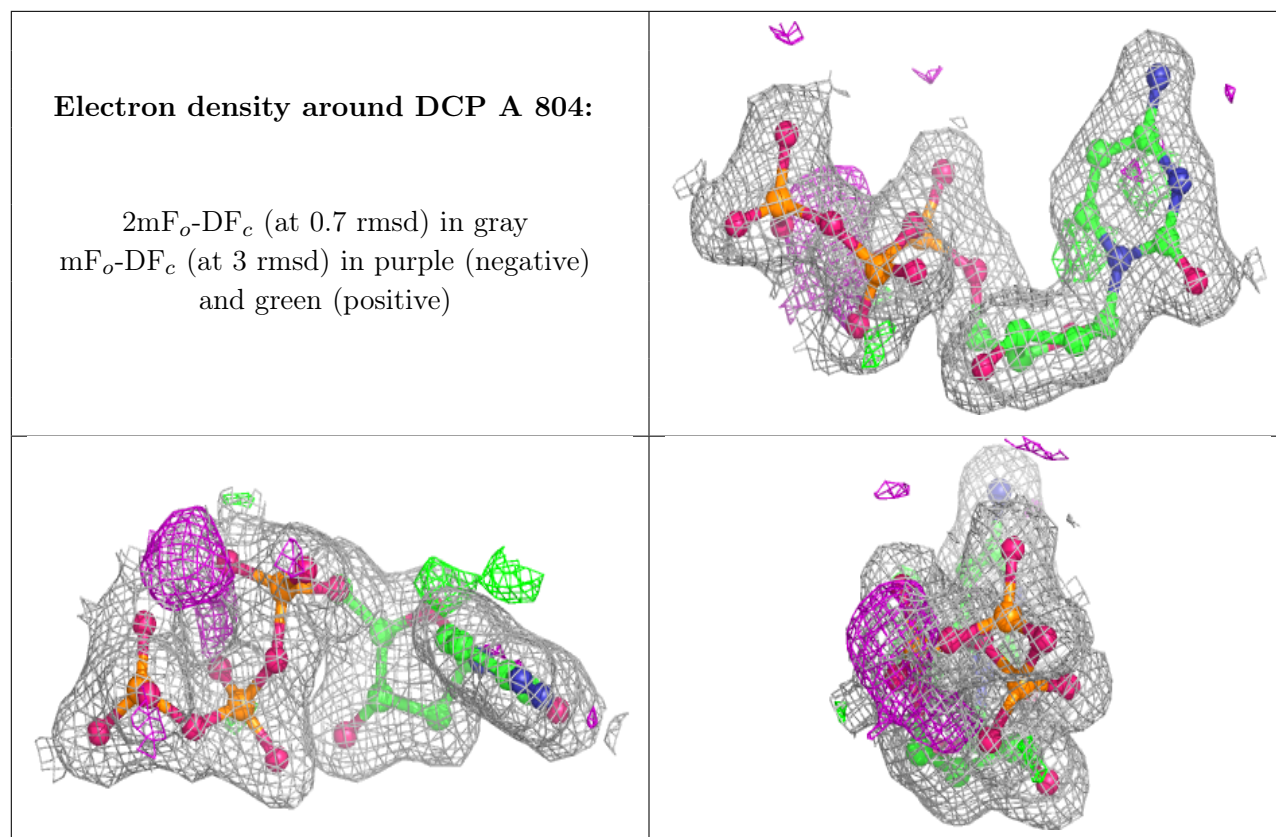
*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MG	A	803	1/1	0.90	0.10	22,22,22,22	0
8	EDO	A	813	4/4	0.90	0.18	30,35,36,41	0
8	EDO	A	811	4/4	0.90	0.13	37,42,44,45	0
8	EDO	A	808	4/4	0.93	0.16	31,34,35,40	0
7	GOL	A	805	6/6	0.93	0.15	35,37,37,40	0
8	EDO	A	810	4/4	0.93	0.12	33,36,36,40	0
7	GOL	A	807	6/6	0.93	0.13	36,40,41,44	0
7	GOL	A	806	6/6	0.94	0.14	48,50,51,52	0
4	MG	P	101	1/1	0.95	0.08	30,30,30,30	0
4	MG	A	802	1/1	0.97	0.07	16,16,16,16	0
6	DCP	A	804	28/28	0.98	0.11	17,18,25,26	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.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.