



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 2, 2024 – 10:36 PM EST

PDB ID : 5W2C  
Title : Structure of human DNA polymerase kappa in complex with Lucidin-derived DNA adduct and incoming dAMPNPP  
Authors : Jha, V.K.; Ling, H.  
Deposited on : 2017-06-06  
Resolution : 2.50 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 : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

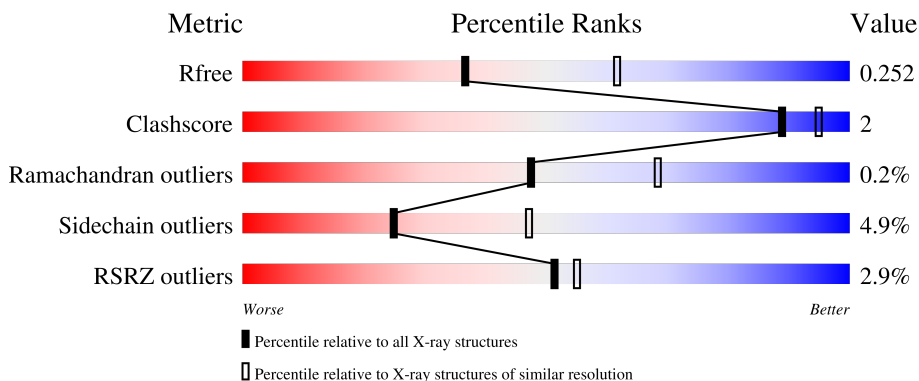
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	A	551	
1	B	551	
2	C	9	
2	P	9	
3	D	13	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	T	13	

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	GOL	B	604	-	-	-	X

## 2 Entry composition i

There are 9 unique types of molecules in this entry. The entry contains 7962 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 DNA polymerase kappa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	431	3385	2140	595	631	19	0	1	0
1	B	431	3332	2109	585	618	20	0	1	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	initiating methionine	UNP Q9UBT6
A	-23	SER	-	expression tag	UNP Q9UBT6
A	-22	TYR	-	expression tag	UNP Q9UBT6
A	-21	TYR	-	expression tag	UNP Q9UBT6
A	-20	HIS	-	expression tag	UNP Q9UBT6
A	-19	HIS	-	expression tag	UNP Q9UBT6
A	-18	HIS	-	expression tag	UNP Q9UBT6
A	-17	HIS	-	expression tag	UNP Q9UBT6
A	-16	HIS	-	expression tag	UNP Q9UBT6
A	-15	HIS	-	expression tag	UNP Q9UBT6
A	-14	ASP	-	expression tag	UNP Q9UBT6
A	-13	TYR	-	expression tag	UNP Q9UBT6
A	-12	ASP	-	expression tag	UNP Q9UBT6
A	-11	ILE	-	expression tag	UNP Q9UBT6
A	-10	PRO	-	expression tag	UNP Q9UBT6
A	-9	THR	-	expression tag	UNP Q9UBT6
A	-8	THR	-	expression tag	UNP Q9UBT6
A	-7	GLU	-	expression tag	UNP Q9UBT6
A	-6	ASN	-	expression tag	UNP Q9UBT6
A	-5	LEU	-	expression tag	UNP Q9UBT6
A	-4	TYR	-	expression tag	UNP Q9UBT6
A	-3	PHE	-	expression tag	UNP Q9UBT6
A	-2	GLN	-	expression tag	UNP Q9UBT6
A	-1	GLY	-	expression tag	UNP Q9UBT6
A	0	ALA	-	expression tag	UNP Q9UBT6

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-24	MET	-	initiating methionine	UNP Q9UBT6
B	-23	SER	-	expression tag	UNP Q9UBT6
B	-22	TYR	-	expression tag	UNP Q9UBT6
B	-21	TYR	-	expression tag	UNP Q9UBT6
B	-20	HIS	-	expression tag	UNP Q9UBT6
B	-19	HIS	-	expression tag	UNP Q9UBT6
B	-18	HIS	-	expression tag	UNP Q9UBT6
B	-17	HIS	-	expression tag	UNP Q9UBT6
B	-16	HIS	-	expression tag	UNP Q9UBT6
B	-15	HIS	-	expression tag	UNP Q9UBT6
B	-14	ASP	-	expression tag	UNP Q9UBT6
B	-13	TYR	-	expression tag	UNP Q9UBT6
B	-12	ASP	-	expression tag	UNP Q9UBT6
B	-11	ILE	-	expression tag	UNP Q9UBT6
B	-10	PRO	-	expression tag	UNP Q9UBT6
B	-9	THR	-	expression tag	UNP Q9UBT6
B	-8	THR	-	expression tag	UNP Q9UBT6
B	-7	GLU	-	expression tag	UNP Q9UBT6
B	-6	ASN	-	expression tag	UNP Q9UBT6
B	-5	LEU	-	expression tag	UNP Q9UBT6
B	-4	TYR	-	expression tag	UNP Q9UBT6
B	-3	PHE	-	expression tag	UNP Q9UBT6
B	-2	GLN	-	expression tag	UNP Q9UBT6
B	-1	GLY	-	expression tag	UNP Q9UBT6
B	0	ALA	-	expression tag	UNP Q9UBT6

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	9	Total	C	N	O	P	0	0	0
			185	87	36	53	9			
2	P	9	Total	C	N	O	P	0	0	0
			185	87	36	53	9			

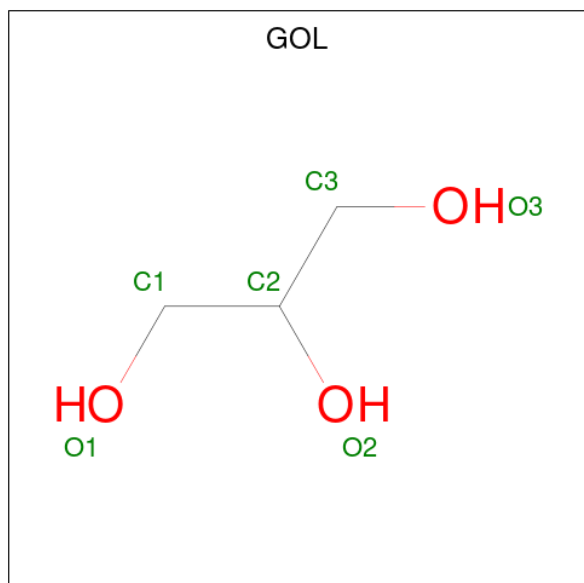
- Molecule 3 is a DNA chain called DNA (5'-D(\*CP\*TP\*AP\*TP\*(LDG)P\*TP\*CP\*GP\*AP\*TP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	13	Total	C	N	O	P	0	0	0
			281	143	45	81	12			
3	T	13	Total	C	N	O	P	0	0	0
			281	143	45	81	12			

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

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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

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



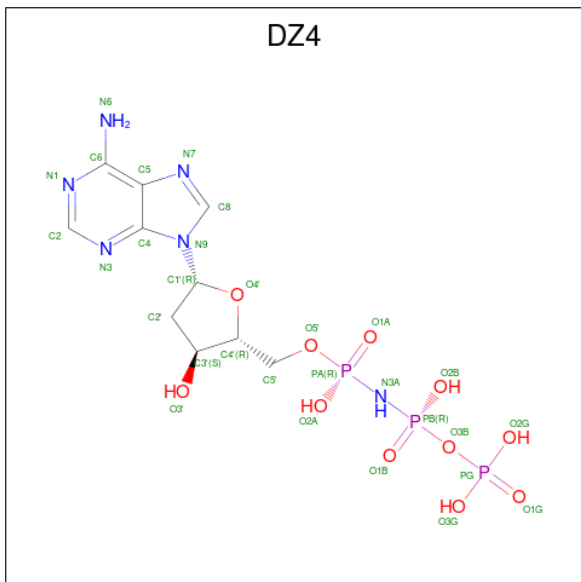
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 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
7	A	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is 2'-deoxy-5'-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]amino}phosphoryl]adenosine (three-letter code: DZ4) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>11</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
8	A	1	Total	C	N	O	P	0	0
			30	10	6	11	3		
8	B	1	Total	C	N	O	P	0	0
			30	10	6	11	3		

- Molecule 9 is water.

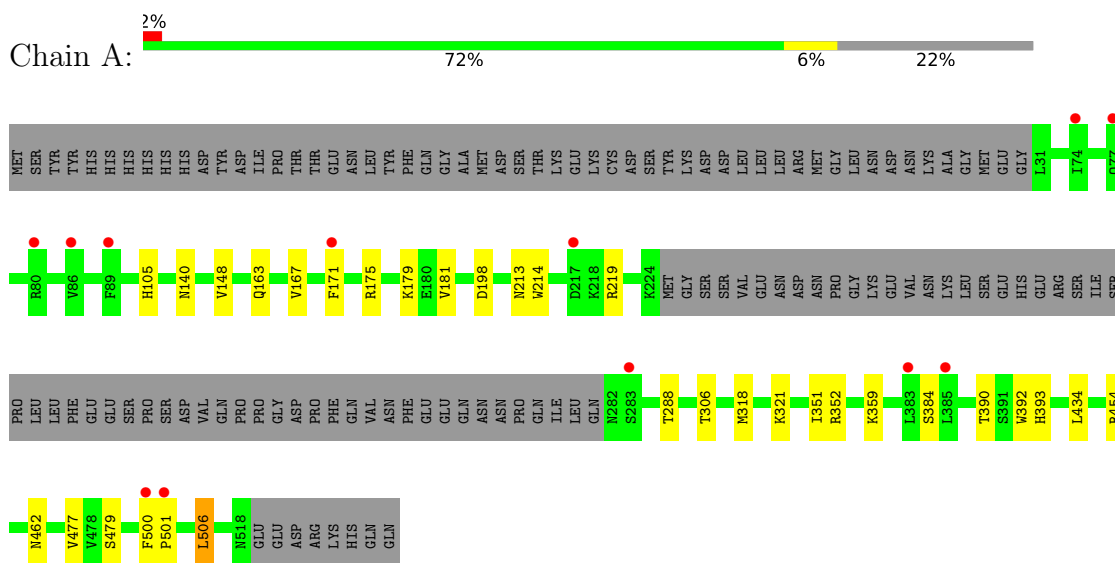
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	86	Total	O	0	0
			86	86		
9	B	80	Total	O	0	0
			80	80		
9	C	9	Total	O	0	0
			9	9		
9	D	12	Total	O	0	0
			12	12		
9	P	12	Total	O	0	0
			12	12		
9	T	14	Total	O	0	0
			14	14		



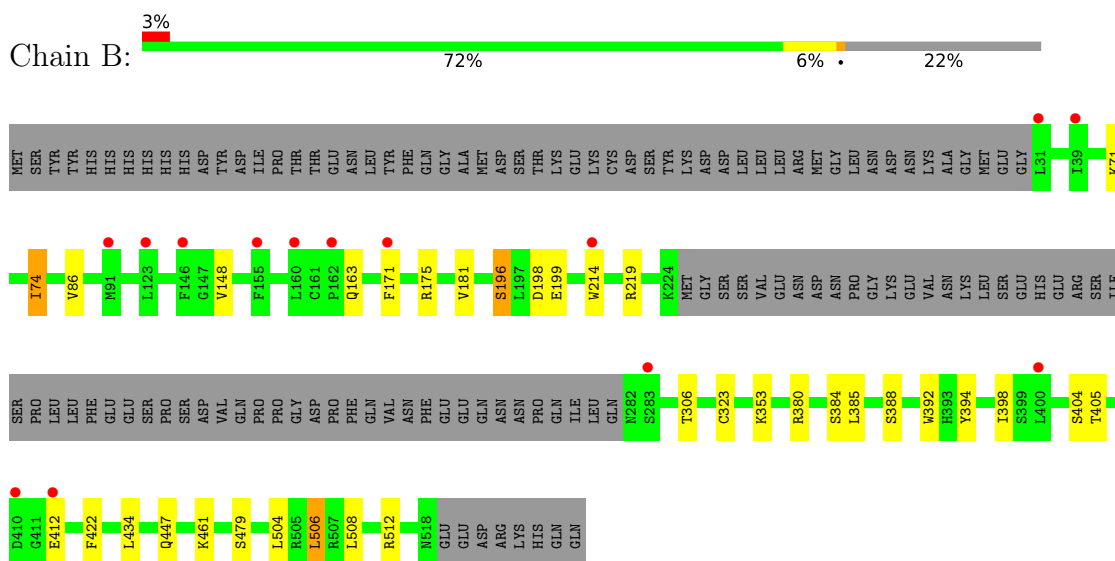
### 3 Residue-property plots [i](#)

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 polymerase kappa



- Molecule 1: DNA polymerase kappa



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

Chain C:  67% 33%




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

Chain P:  67% 33%



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

Chain D:  77% 23%



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

Chain T:  62% 38%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.71Å 128.92Å 167.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 2.50 47.21 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.01-2.50) 99.3 (47.21-2.50)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.62 (at 2.51Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.211 , 0.253 0.215 , 0.252	Depositor DCC
$R_{free}$ test set	1195 reflections (2.47%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.5	Xtrriage
Anisotropy	0.889	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 55.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7962	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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: DZ4, PEG, GOL, MG, LDG, EDO

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.47	0/3441	0.71	0/4640
1	B	0.48	0/3389	0.71	0/4577
2	C	0.35	0/207	0.82	0/317
2	P	0.44	0/207	0.83	0/317
3	D	0.64	0/264	0.86	1/403 (0.2%)
3	T	0.57	0/264	0.89	1/403 (0.2%)
All	All	0.48	0/7772	0.73	2/10657 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	T	7	DT	C1'-O4'-C4'	-7.90	102.20	110.10
3	D	7	DT	C1'-O4'-C4'	-7.32	102.78	110.10

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	A	3385	0	3394	11	0
1	B	3332	0	3288	12	0
2	C	185	0	101	2	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	P	185	0	101	1	1
3	D	281	0	135	1	1
3	T	281	0	135	4	2
4	A	3	0	0	0	0
4	B	3	0	0	0	0
5	A	6	0	8	1	0
5	B	6	0	8	3	0
6	A	8	0	12	2	0
7	A	7	0	10	0	0
7	B	7	0	10	0	0
8	A	30	0	13	0	0
8	B	30	0	13	0	0
9	A	86	0	0	2	0
9	B	80	0	0	0	0
9	C	9	0	0	0	0
9	D	12	0	0	0	0
9	P	12	0	0	0	0
9	T	14	0	0	0	0
All	All	7962	0	7228	29	3

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 2.

All (29) 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:P:12:DA:H2''	2:P:13:DC:H5'	1.76	0.67
1:A:105:HIS:NE2	9:A:701:HOH:O	2.27	0.66
1:A:321:LYS:NZ	2:C:13:DC:OP1	2.32	0.60
1:B:388:SER:OG	3:T:12:DC:OP1	2.12	0.60
1:A:140:ASN:HB2	6:A:605:EDO:H22	1.83	0.59
2:C:12:DA:H2''	2:C:13:DC:H5'	1.84	0.59
1:A:214:TRP:O	1:A:219:ARG:NH1	2.39	0.56
1:A:140:ASN:HB2	6:A:605:EDO:C2	2.40	0.52
1:B:71:LYS:O	1:B:74:ILE:HG23	2.09	0.52
1:A:352:ARG:NH2	9:A:702:HOH:O	2.43	0.50
1:B:74:ILE:HD13	1:B:385:LEU:HD13	1.92	0.50
5:B:604:GOL:H2	3:T:6:LDG:C81	2.43	0.49
1:B:171[B]:PHE:CZ	5:B:604:GOL:H31	2.47	0.48
1:B:214:TRP:O	1:B:219:ARG:NH1	2.48	0.47
1:A:384:SER:HA	1:A:392:TRP:CD1	2.49	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:461:LYS:HB3	1:B:508:LEU:HB3	1.97	0.47
5:B:604:GOL:H2	3:T:6:LDG:C9	2.45	0.47
3:T:13:DC:H2''	3:T:14:DG:OP2	2.15	0.45
1:A:171[B]:PHE:CE1	5:A:604:GOL:H2	2.52	0.45
1:A:462:ASN:HA	1:A:506:LEU:HA	2.00	0.44
1:A:500:PHE:CD1	1:A:501:PRO:HA	2.54	0.42
1:B:196:SER:HB3	1:B:199:GLU:HG2	2.01	0.42
1:B:74:ILE:HD13	1:B:385:LEU:CD1	2.50	0.41
1:B:384:SER:HA	1:B:392:TRP:CD1	2.56	0.41
1:B:394:TYR:CZ	1:B:398:ILE:HD11	2.56	0.41
1:A:454:ARG:HG3	1:A:477:VAL:HG12	2.03	0.41
1:B:504:LEU:HB3	1:B:506:LEU:HD13	2.02	0.41
3:D:6:LDG:O19	3:D:6:LDG:C15	2.69	0.41
1:B:86:VAL:HG11	1:B:380:ARG:HG3	2.03	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:5:DC:P	3:T:14:DG:O3'[3_555]	1.56	0.64
2:C:5:DC:OP2	3:T:14:DG:O3'[3_555]	1.99	0.21
3:D:14:DG:O3'	2:P:5:DC:P[3_555]	2.03	0.17

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	428/551 (78%)	417 (97%)	10 (2%)	1 (0%)	47 68
1	B	428/551 (78%)	412 (96%)	15 (4%)	1 (0%)	47 68
All	All	856/1102 (78%)	829 (97%)	25 (3%)	2 (0%)	47 68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	479	SER
1	B	479	SER

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	363/495 (73%)	346 (95%)	17 (5%)	26	49
1	B	348/495 (70%)	330 (95%)	18 (5%)	23	44
All	All	711/990 (72%)	676 (95%)	35 (5%)	25	47

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	148	VAL
1	A	163	GLN
1	A	167	VAL
1	A	175	ARG
1	A	179	LYS
1	A	181	VAL
1	A	198	ASP
1	A	213	ASN
1	A	288	THR
1	A	306	THR
1	A	318	MET
1	A	351	ILE
1	A	359	LYS
1	A	390	THR
1	A	393	HIS
1	A	434	LEU
1	A	506	LEU
1	B	74	ILE
1	B	148	VAL
1	B	163	GLN
1	B	175	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	181	VAL
1	B	196	SER
1	B	198	ASP
1	B	306	THR
1	B	323	CYS
1	B	353	LYS
1	B	404	SER
1	B	405	THR
1	B	412	GLU
1	B	422	PHE
1	B	434	LEU
1	B	447	GLN
1	B	506	LEU
1	B	512	ARG

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

Mol	Chain	Res	Type
1	A	212	GLN
1	A	444	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](#)

Of 14 ligands modelled in this entry, 6 are monoatomic - leaving 8 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	GOL	B	604	-	5,5,5	0.75	0	5,5,5	0.94	0
8	DZ4	A	608	4	29,32,32	2.50	4 (13%)	33,50,50	0.88	1 (3%)
7	PEG	B	605	-	6,6,6	0.62	0	5,5,5	0.42	0
6	EDO	A	605	-	3,3,3	0.57	0	2,2,2	0.67	0
7	PEG	A	607	-	6,6,6	0.46	0	5,5,5	0.36	0
6	EDO	A	606	-	3,3,3	0.60	0	2,2,2	0.36	0
8	DZ4	B	606	4	29,32,32	0.76	0	33,50,50	0.70	0
5	GOL	A	604	-	5,5,5	0.54	0	5,5,5	0.42	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
5	GOL	B	604	-	-	2/4/4/4	-
8	DZ4	A	608	4	-	3/15/34/34	0/3/3/3
7	PEG	B	605	-	-	2/4/4/4	-
6	EDO	A	605	-	-	1/1/1/1	-
7	PEG	A	607	-	-	3/4/4/4	-
6	EDO	A	606	-	-	1/1/1/1	-
8	DZ4	B	606	4	-	7/15/34/34	0/3/3/3
5	GOL	A	604	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	608	DZ4	PB-O1B	9.77	1.61	1.46
8	A	608	DZ4	PA-O1A	6.74	1.56	1.46
8	A	608	DZ4	PA-O2A	-3.82	1.46	1.56
8	A	608	DZ4	PB-O2B	-3.81	1.46	1.56

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	608	DZ4	C5-C6-N6	2.07	123.50	120.35

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	604	GOL	O1-C1-C2-C3
8	A	608	DZ4	PB-N3A-PA-O5'
8	A	608	DZ4	PB-O3B-PG-O3G
8	B	606	DZ4	PB-N3A-PA-O5'
8	B	606	DZ4	PA-N3A-PB-O1B
8	B	606	DZ4	PB-O3B-PG-O3G
7	A	607	PEG	O1-C1-C2-O2
5	A	604	GOL	O1-C1-C2-C3
7	A	607	PEG	O2-C3-C4-O4
5	B	604	GOL	O1-C1-C2-O2
7	B	605	PEG	O2-C3-C4-O4
5	A	604	GOL	O1-C1-C2-O2
6	A	606	EDO	O1-C1-C2-O2
7	B	605	PEG	C4-C3-O2-C2
6	A	605	EDO	O1-C1-C2-O2
8	B	606	DZ4	PB-O3B-PG-O1G
8	B	606	DZ4	PB-N3A-PA-O1A
7	A	607	PEG	C1-C2-O2-C3
8	A	608	DZ4	PB-O3B-PG-O2G
8	B	606	DZ4	PB-O3B-PG-O2G
8	B	606	DZ4	C5'-O5'-PA-O1A

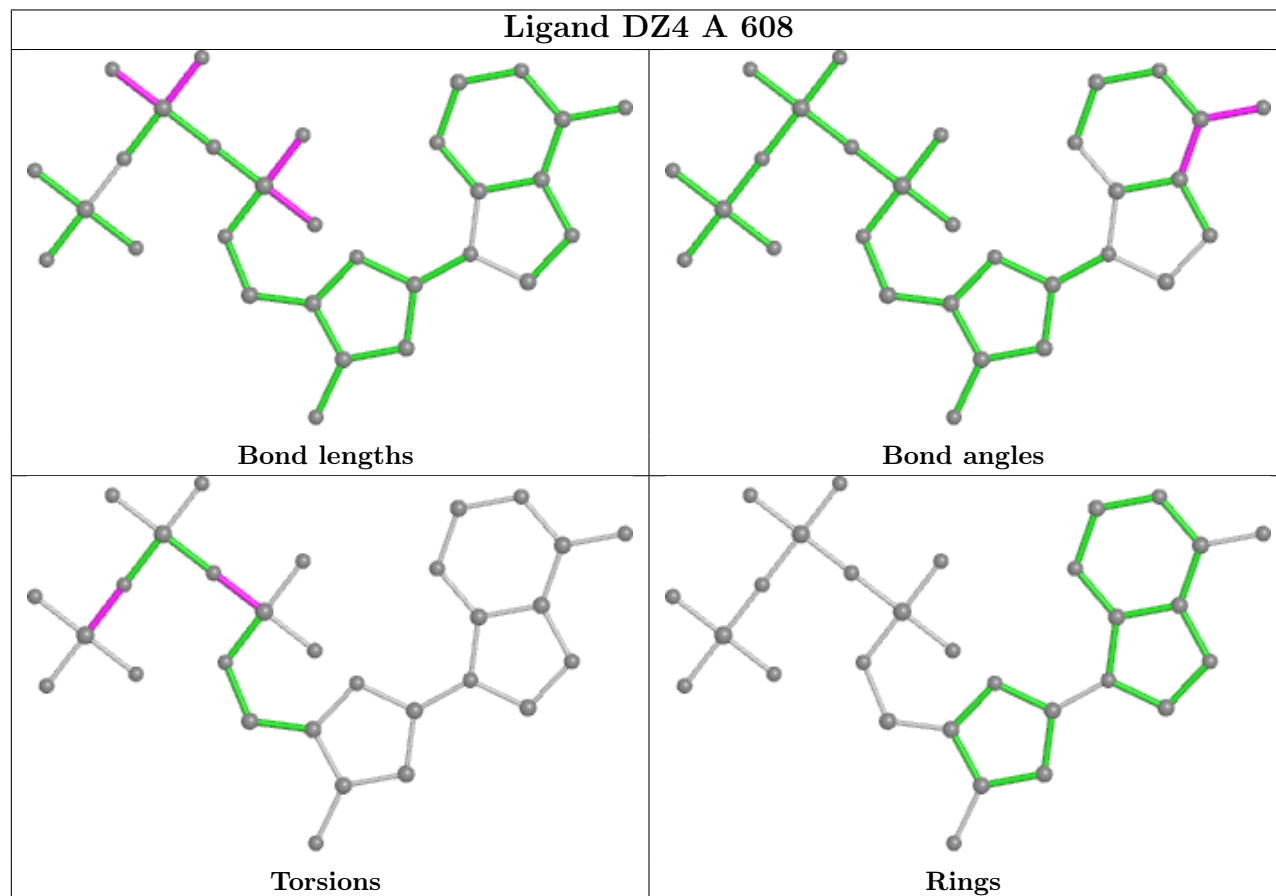
There are no ring outliers.

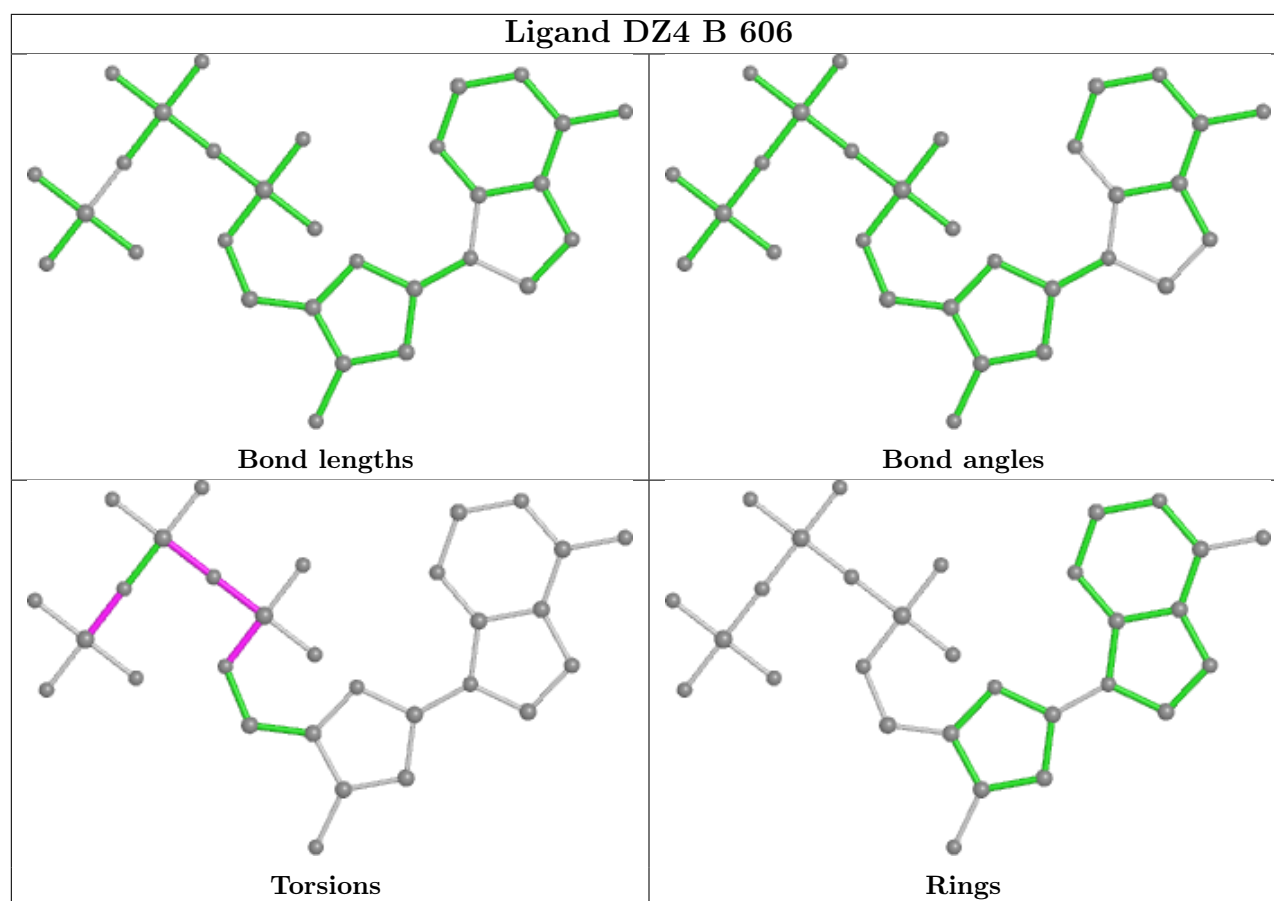
3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	604	GOL	3	0
6	A	605	EDO	2	0
5	A	604	GOL	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	A	431/551 (78%)	0.23	12 (2%) 53 56	52, 80, 123, 147	0
1	B	431/551 (78%)	0.19	14 (3%) 47 51	51, 82, 115, 148	0
2	C	9/9 (100%)	-0.28	0 100 100	65, 68, 72, 74	0
2	P	9/9 (100%)	-0.08	0 100 100	62, 65, 71, 75	0
3	D	12/13 (92%)	0.02	0 100 100	64, 75, 141, 151	0
3	T	12/13 (92%)	0.11	0 100 100	60, 79, 129, 136	0
All	All	904/1146 (78%)	0.20	26 (2%) 51 55	51, 81, 118, 151	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	74	ILE	5.1
1	B	146	PHE	4.4
1	B	31	LEU	4.2
1	B	171[A]	PHE	3.8
1	A	500	PHE	3.5
1	B	91	MET	3.4
1	A	171[A]	PHE	3.3
1	B	214	TRP	3.2
1	A	385	LEU	2.8
1	A	383	LEU	2.8
1	B	162	PRO	2.7
1	A	77	GLN	2.7
1	A	89	PHE	2.6
1	B	160	LEU	2.5
1	B	410	ASP	2.5
1	B	412	GLU	2.5
1	B	123	LEU	2.5
1	A	217	ASP	2.5
1	B	400	LEU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	283	SER	2.3
1	B	155	PHE	2.3
1	B	39	ILE	2.2
1	A	80	ARG	2.2
1	A	86	VAL	2.1
1	B	283	SER	2.1
1	A	501	PRO	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

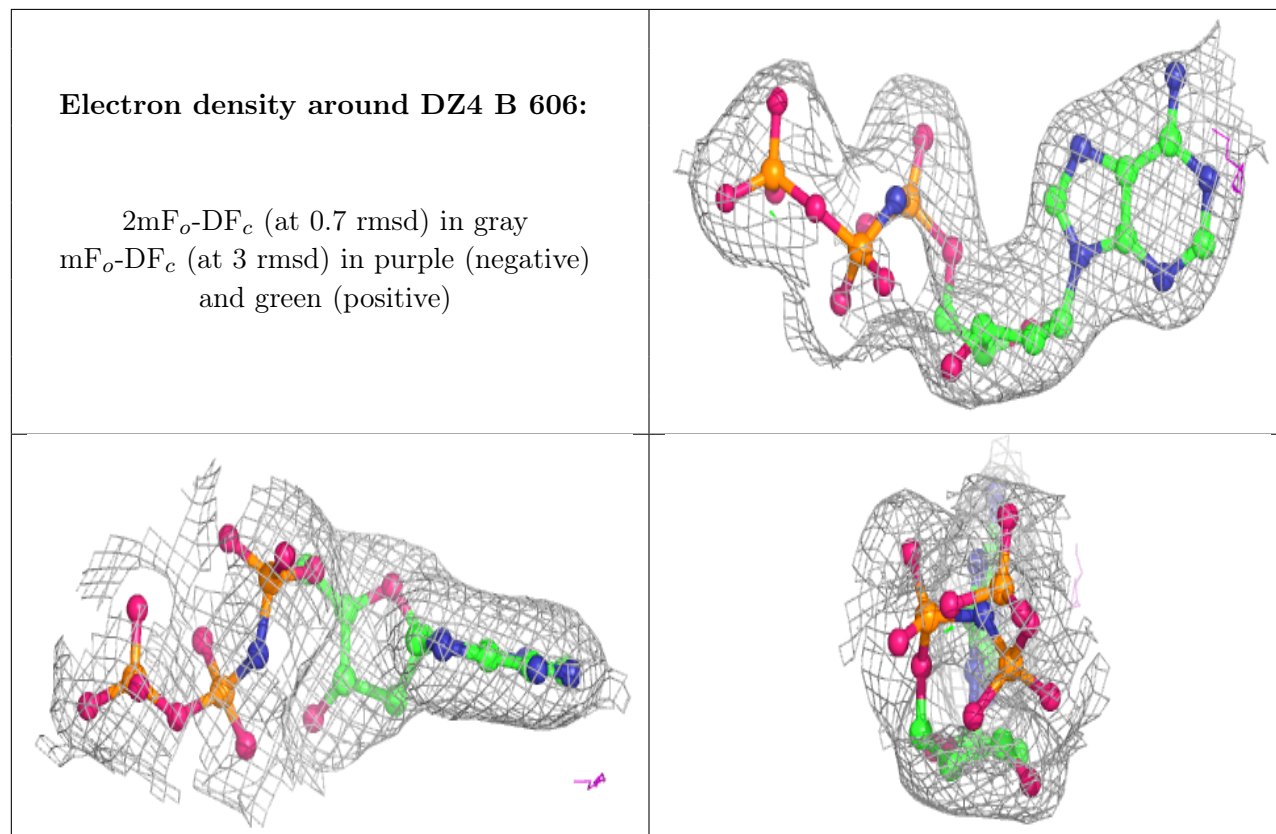
## 6.4 Ligands [i](#)

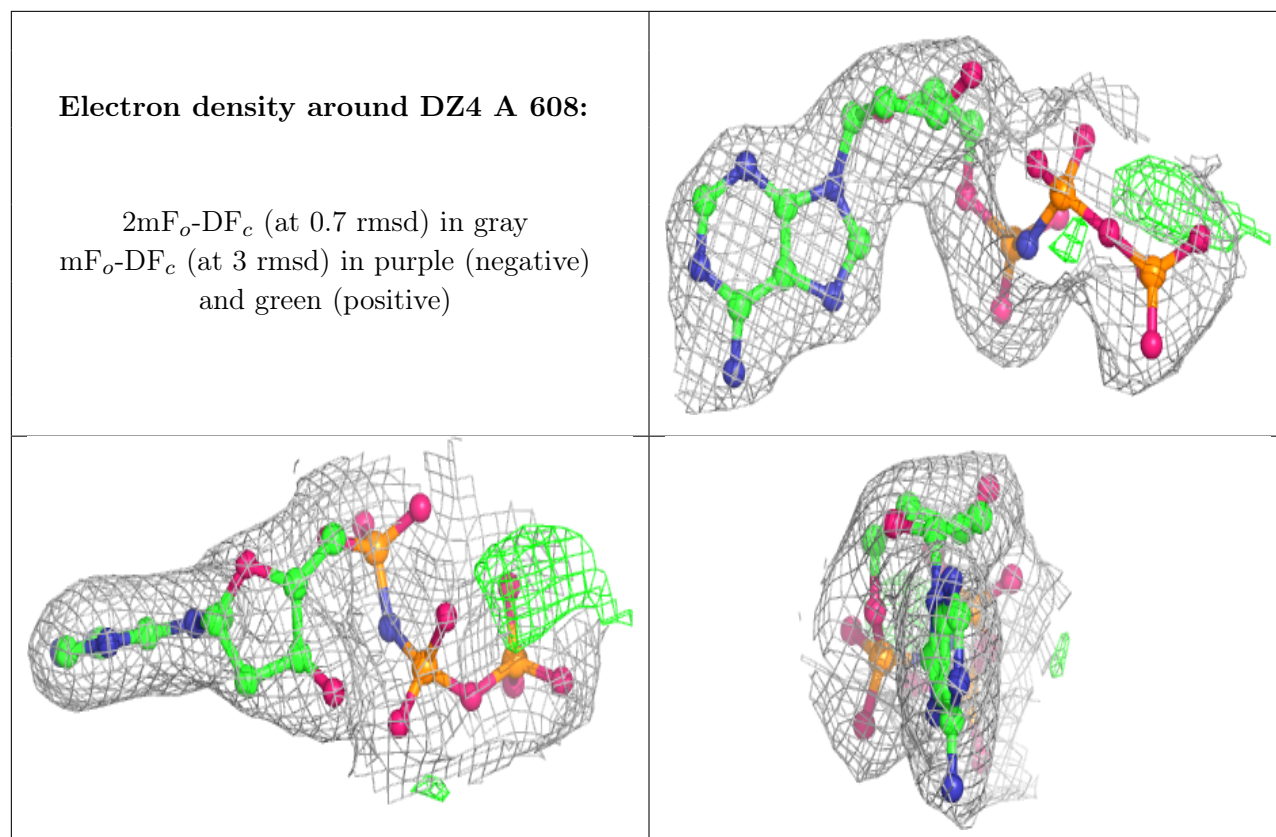
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
7	PEG	B	605	7/7	0.66	0.14	95,105,111,117	0
4	MG	B	603	1/1	0.74	0.20	95,95,95,95	0
6	EDO	A	605	4/4	0.75	0.22	83,84,84,87	0
4	MG	A	602	1/1	0.76	0.06	85,85,85,85	0
6	EDO	A	606	4/4	0.80	0.15	92,93,94,96	0
5	GOL	B	604	6/6	0.80	0.46	75,82,84,84	0
7	PEG	A	607	7/7	0.82	0.18	99,102,110,114	0
4	MG	B	602	1/1	0.85	0.10	82,82,82,82	0
5	GOL	A	604	6/6	0.86	0.47	74,79,83,87	0
4	MG	A	603	1/1	0.95	0.27	72,72,72,72	0
4	MG	B	601	1/1	0.97	0.17	62,62,62,62	0
4	MG	A	601	1/1	0.98	0.23	46,46,46,46	0
8	DZ4	B	606	30/30	0.98	0.19	55,63,70,72	0
8	DZ4	A	608	30/30	0.99	0.18	45,59,63,68	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.