



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 07:09 PM UTC

PDB ID : 4F4Y / pdb\_00004f4y  
Title : Y-family DNA polymerase chimera Dbh-Dpo4-Dbh  
Authors : Pata, J.D.; Wilson, R.C.  
Deposited on : 2012-05-11  
Resolution : 2.34 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 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

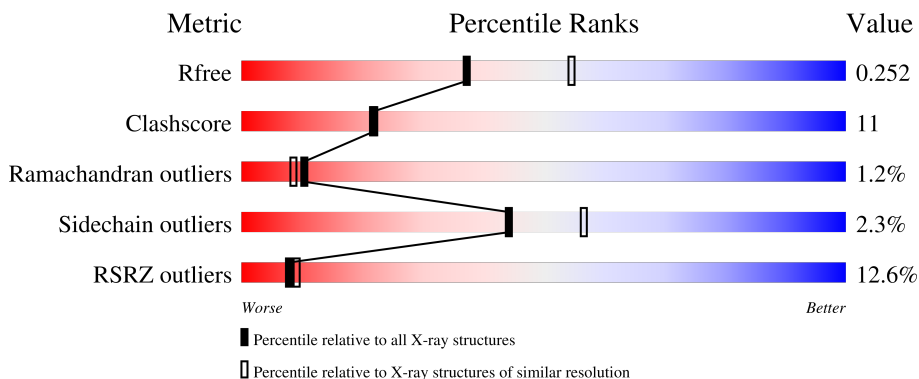
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.34 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))
$R_{free}$	180053	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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	362	 10% 72% 22% • 5%
1	B	362	 15% 68% 25% • 5%
2	C	16	 6% 69% 31%
2	P	16	 62% 38%
3	D	19	 63% 26% 5% 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	T	19	 68% 26% 5%

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 7148 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 IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	343	Total	C	N	O	S	0	1	0
			2736	1758	476	494	8			
1	B	343	Total	C	N	O	S	0	0	0
			2730	1753	475	494	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	355	GLY	-	expression tag	UNP Q4JB80
A	356	GLY	-	expression tag	UNP Q4JB80
A	357	HIS	-	expression tag	UNP Q4JB80
A	358	HIS	-	expression tag	UNP Q4JB80
A	359	HIS	-	expression tag	UNP Q4JB80
A	360	HIS	-	expression tag	UNP Q4JB80
A	361	HIS	-	expression tag	UNP Q4JB80
A	362	HIS	-	expression tag	UNP Q4JB80
B	355	GLY	-	expression tag	UNP Q4JB80
B	356	GLY	-	expression tag	UNP Q4JB80
B	357	HIS	-	expression tag	UNP Q4JB80
B	358	HIS	-	expression tag	UNP Q4JB80
B	359	HIS	-	expression tag	UNP Q4JB80
B	360	HIS	-	expression tag	UNP Q4JB80
B	361	HIS	-	expression tag	UNP Q4JB80
B	362	HIS	-	expression tag	UNP Q4JB80

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	16	Total	C	N	O	P	0	0	0
			325	154	62	94	15			

*Continued on next page...*

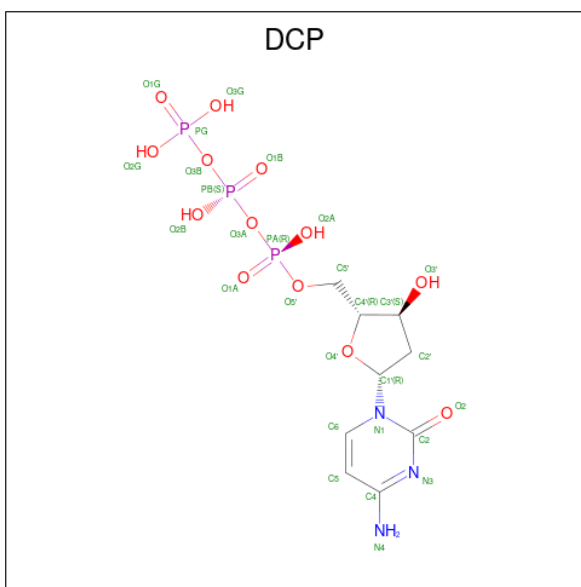
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	P	16	Total 325	C 154	N 62	O 94	P 15	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	18	Total 364	C 173	N 64	O 109	P 18	0	0	0
3	T	18	Total 364	C 173	N 64	O 109	P 18	0	0	0

- Molecule 4 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (CCD ID: DCP) (formula: C<sub>9</sub>H<sub>16</sub>N<sub>3</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	Total 24	C 9	N 3	O 10	P 2	0	0
4	C	1	Total 24	C 9	N 3	O 10	P 2	0	0

- Molecule 5 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	2	Total 2	Ca 2	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total 2	O Ca 2	0	0

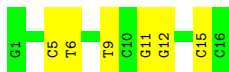
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	114	Total 114	O 114	0	0
6	B	77	Total 77	O 77	0	0
6	C	18	Total 18	O 18	0	0
6	D	13	Total 13	O 13	0	0
6	P	17	Total 17	O 17	0	0
6	T	13	Total 13	O 13	0	0





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



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.08Å 103.39Å 112.99Å 90.00° 101.76° 90.00°	Depositor
Resolution (Å)	29.25 – 2.34 29.25 – 2.34	Depositor EDS
% Data completeness (in resolution range)	94.1 (29.25-2.34) 94.1 (29.25-2.34)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.71 (at 2.34Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.222 , 0.259 0.214 , 0.252	Depositor DCC
$R_{free}$ test set	1893 reflections (3.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.8	Xtrriage
Anisotropy	0.602	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 62.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	7148	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.50 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.1373e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CA, DCP

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	0.42	0/2777	0.82	1/3741 (0.0%)
1	B	0.40	0/2768	0.83	0/3730
2	C	0.23	0/364	0.87	0/560
2	P	0.30	0/364	1.00	0/560
3	D	0.24	0/406	0.88	3/623 (0.5%)
3	T	0.25	0/406	0.92	1/623 (0.2%)
All	All	0.38	0/7085	0.85	5/9837 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	5	DG	O4'-C1'-N9	-6.61	98.49	108.40
1	A	189	SER	N-CA-C	-5.84	106.69	113.88
3	D	5	DG	C4'-C3'-O3'	-5.69	101.47	110.00
3	D	5	DG	O4'-C1'-N9	-5.67	99.89	108.40
3	D	5	DG	O5'-C5'-C4'	-5.13	103.11	110.80

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	2736	0	2917	56	0
1	B	2730	0	2904	67	0
2	C	325	0	180	7	0
2	P	325	0	180	6	0
3	D	364	0	203	10	0
3	T	364	0	203	3	1
4	A	24	0	12	0	0
4	C	24	0	12	1	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
6	A	114	0	0	14	0
6	B	77	0	0	8	1
6	C	18	0	0	3	0
6	D	13	0	0	6	0
6	P	17	0	0	2	0
6	T	13	0	0	0	0
All	All	7148	0	6611	140	1

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 11.

The worst 5 of 140 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:A:198:LEU:O	6:A:588:HOH:O	1.96	0.82
1:B:325:ARG:NH2	2:C:9:DT:OP1	2.13	0.81
2:C:9:DT:O4	6:C:216:HOH:O	1.99	0.80
1:B:182:ASP:O	6:B:560:HOH:O	2.00	0.79
1:A:36:ARG:NH2	1:A:255:ASN:OD1	2.15	0.75

All (1) 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)
3:T:2:DT:O4	6:B:515:HOH:O[1_554]	2.12	0.08

## 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	342/362 (94%)	315 (92%)	22 (6%)	5 (2%)	8	6
1	B	341/362 (94%)	318 (93%)	20 (6%)	3 (1%)	14	13
All	All	683/724 (94%)	633 (93%)	42 (6%)	8 (1%)	10	8

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	TYR
1	A	211	TYR
1	A	210	ASN
1	A	198	LEU
1	A	327	LYS

### 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	299/315 (95%)	293 (98%)	6 (2%)	48	61
1	B	298/315 (95%)	290 (97%)	8 (3%)	39	51
All	All	597/630 (95%)	583 (98%)	14 (2%)	44	56

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	106	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	115	VAL
1	B	324	VAL
1	B	213	GLU
1	B	224	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	ASN
1	B	113	ASN
1	B	275	ASN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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$
4	DCP	C	101	5	24,25,29	2.24	6 (25%)	36,38,45	1.43	6 (16%)
4	DCP	A	401	5	24,25,29	2.33	9 (37%)	36,38,45	1.41	8 (22%)

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	DCP	C	101	5	-	3/16/28/34	0/2/2/2
4	DCP	A	401	5	-	1/16/28/34	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	101	DCP	O2-C2	6.76	1.36	1.23
4	A	401	DCP	O2-C2	6.59	1.36	1.23
4	C	101	DCP	C4-N4	4.52	1.44	1.33
4	A	401	DCP	C4-N4	4.50	1.44	1.33
4	A	401	DCP	C3'-C4'	-3.94	1.42	1.53

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	101	DCP	O2B-PB-O1B	-3.57	96.92	110.83
4	A	401	DCP	O2B-PB-O1B	-3.42	97.51	110.83
4	C	101	DCP	O4'-C1'-N1	3.37	113.84	107.86
4	C	101	DCP	O2-C2-N3	-3.01	117.58	122.33
4	A	401	DCP	O2A-PA-O1A	-2.94	98.75	112.44

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	DCP	PB-O3A-PA-O1A
4	C	101	DCP	C5'-O5'-PA-O1A
4	C	101	DCP	PA-O3A-PB-O1B
4	C	101	DCP	PA-O3A-PB-O3B

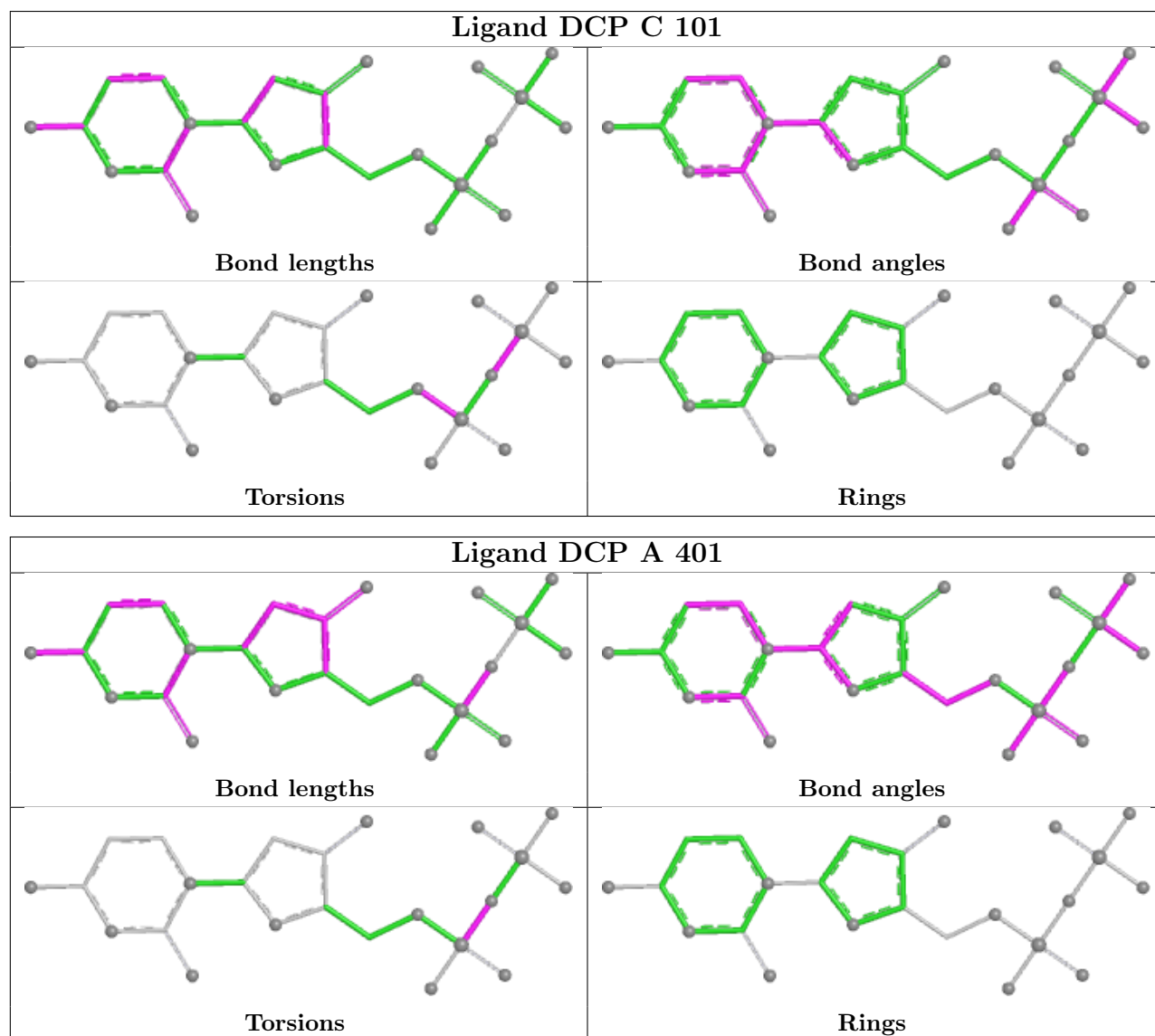
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	101	DCP	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

There are no chain breaks in this entry.

## 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	A	343/362 (94%)	0.87	38 (11%) 10 12	19, 38, 89, 112	1 (0%)
1	B	343/362 (94%)	1.18	56 (16%) 4 5	25, 43, 91, 118	0
2	C	16/16 (100%)	0.92	1 (6%) 26 31	46, 59, 70, 71	0
2	P	16/16 (100%)	0.27	0 100 100	37, 48, 57, 72	0
3	D	18/19 (94%)	0.62	0 100 100	27, 59, 72, 75	0
3	T	18/19 (94%)	0.20	0 100 100	32, 50, 64, 79	0
All	All	754/794 (94%)	0.97	95 (12%) 8 9	19, 42, 90, 118	1 (0%)

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	211	TYR	5.6
1	A	200	ILE	5.1
1	B	198	LEU	5.0
1	A	198	LEU	4.9
1	B	174	ASP	4.6

### 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](#)

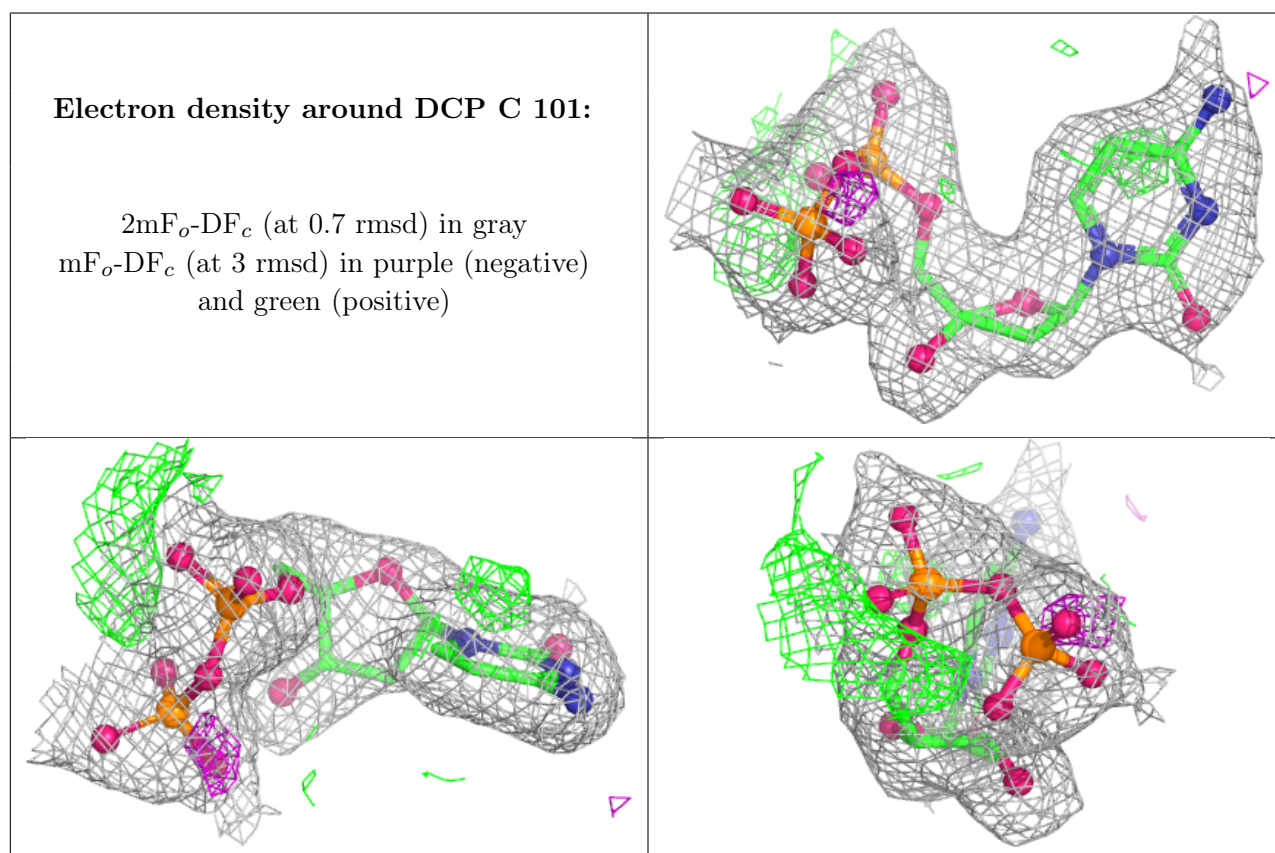
There are no oligosaccharides in this entry.

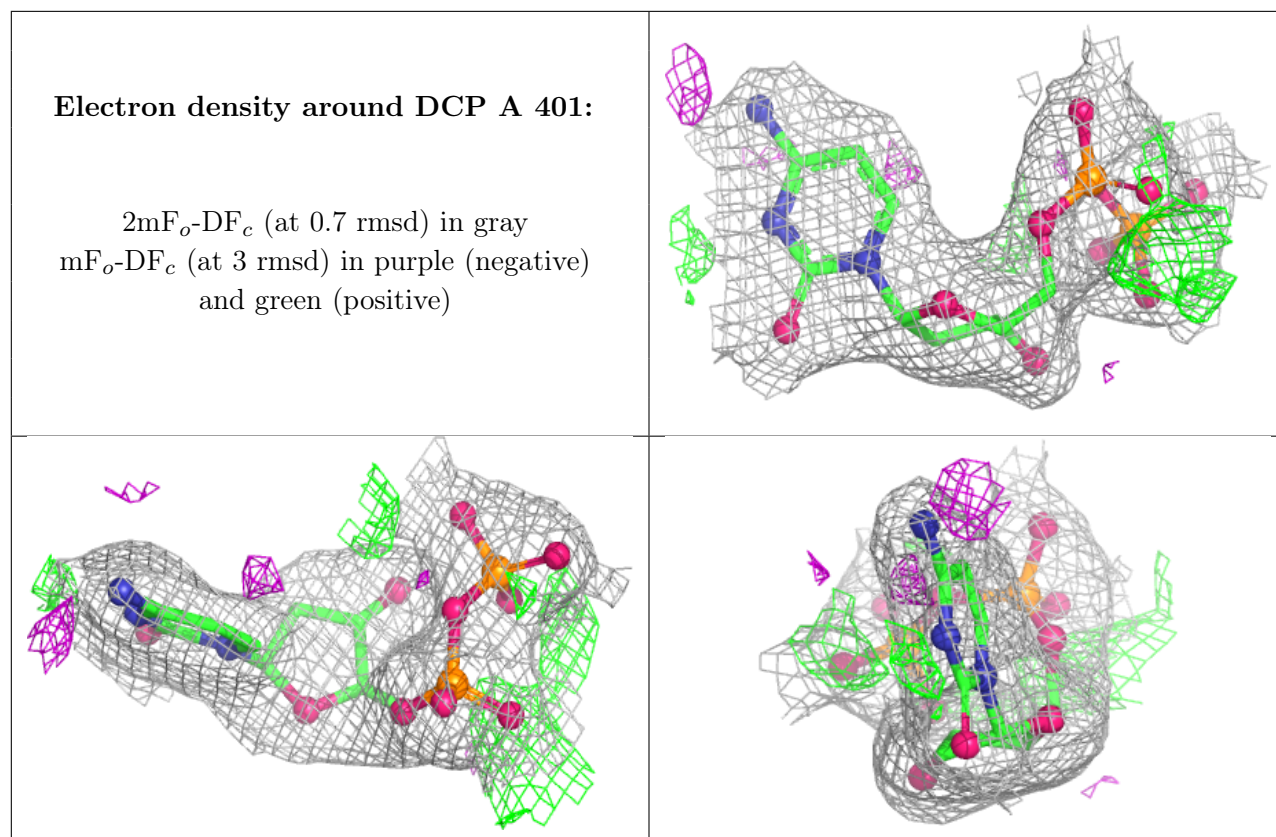
## 6.4 Ligands [i](#)

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
5	CA	A	403	1/1	0.87	0.14	36,36,36,36	0
5	CA	B	401	1/1	0.88	0.18	47,47,47,47	0
5	CA	B	402	1/1	0.90	0.11	35,35,35,35	0
4	DCP	C	101	24/28	0.94	0.10	22,32,45,53	0
4	DCP	A	401	24/28	0.95	0.09	7,21,34,143	0
5	CA	A	402	1/1	0.96	0.10	30,30,30,30	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.