



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 25, 2024 – 02:38 AM EST

PDB ID : 4ZTZ  
Title : Structural basis for processivity and antiviral drug toxicity in human mitochondrial DNA replicase  
Authors : Szymanski, M.R.; Yin, Y.W.  
Deposited on : 2015-05-15  
Resolution : 3.44 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 : 4.02b-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

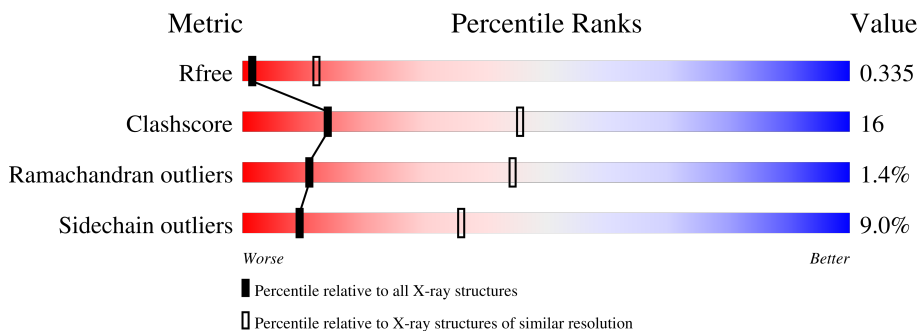
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.44 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	1278 (3.50-3.38)
Clashscore	141614	1361 (3.50-3.38)
Ramachandran outliers	138981	1327 (3.50-3.38)
Sidechain outliers	138945	1328 (3.50-3.38)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	1222	53% (green), 23% (yellow), 20% (grey)
2	B	472	58% (green), 17% (yellow), 23% (grey)
2	C	472	55% (green), 17% (yellow), 24% (grey)
3	T	27	52% (green), 41% (yellow), 7% (grey)
4	P	24	54% (green), 38% (yellow), 8% (grey)

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 14651 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 subunit gamma-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	983	7823	4983	1371	1419	50	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	MET	-	expression tag	UNP P54098
A	198	ALA	ASP	conflict	UNP P54098
A	200	ALA	GLU	conflict	UNP P54098
A	1240	ALA	-	expression tag	UNP P54098
A	1241	ALA	-	expression tag	UNP P54098
A	1242	ALA	-	expression tag	UNP P54098
A	1243	LEU	-	expression tag	UNP P54098
A	1244	GLU	-	expression tag	UNP P54098
A	1245	HIS	-	expression tag	UNP P54098
A	1246	HIS	-	expression tag	UNP P54098
A	1247	HIS	-	expression tag	UNP P54098
A	1248	HIS	-	expression tag	UNP P54098
A	1249	HIS	-	expression tag	UNP P54098
A	1250	HIS	-	expression tag	UNP P54098

- Molecule 2 is a protein called DNA polymerase subunit gamma-2, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	363	2942	1885	520	521	16	0	0	0
2	C	358	2893	1857	509	511	16	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	25	MET	-	expression tag	UNP Q9UHN1
B	486	ALA	-	expression tag	UNP Q9UHN1
B	487	ALA	-	expression tag	UNP Q9UHN1
B	488	ALA	-	expression tag	UNP Q9UHN1
B	489	LEU	-	expression tag	UNP Q9UHN1
B	490	GLU	-	expression tag	UNP Q9UHN1
B	491	HIS	-	expression tag	UNP Q9UHN1
B	492	HIS	-	expression tag	UNP Q9UHN1
B	493	HIS	-	expression tag	UNP Q9UHN1
B	494	HIS	-	expression tag	UNP Q9UHN1
B	495	HIS	-	expression tag	UNP Q9UHN1
B	496	HIS	-	expression tag	UNP Q9UHN1
C	25	MET	-	expression tag	UNP Q9UHN1
C	486	ALA	-	expression tag	UNP Q9UHN1
C	487	ALA	-	expression tag	UNP Q9UHN1
C	488	ALA	-	expression tag	UNP Q9UHN1
C	489	LEU	-	expression tag	UNP Q9UHN1
C	490	GLU	-	expression tag	UNP Q9UHN1
C	491	HIS	-	expression tag	UNP Q9UHN1
C	492	HIS	-	expression tag	UNP Q9UHN1
C	493	HIS	-	expression tag	UNP Q9UHN1
C	494	HIS	-	expression tag	UNP Q9UHN1
C	495	HIS	-	expression tag	UNP Q9UHN1
C	496	HIS	-	expression tag	UNP Q9UHN1

- Molecule 3 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	T	25	508	241	89	153	25	0	0	0

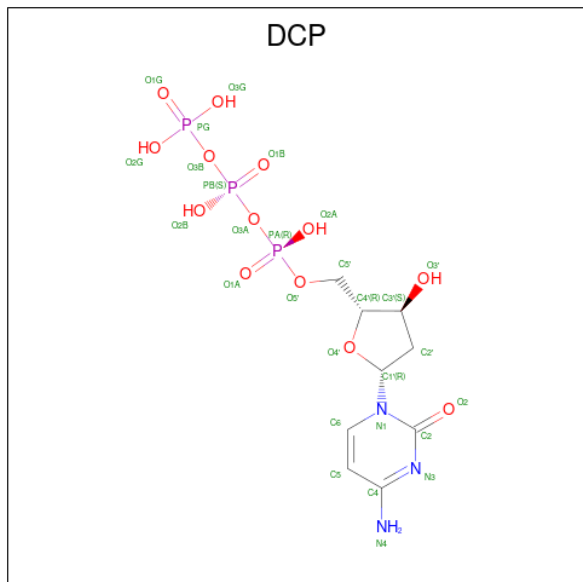
- Molecule 4 is a DNA chain called DNA (5'-D(P\*AP\*AP\*GP\*AP\*CP\*GP\*AP\*GP\*GP\*G P\*CP\*CP\*AP\*GP\*TP\*GP\*CP\*CP\*GP\*TP\*AP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	P	22	455	214	92	127	22	0	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- 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
			Total	C	N	O	P		
6	A	1	28	9	3	13	3	0	0

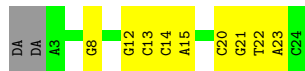






- Molecule 4: DNA (5'-D(P\*AP\*AP\*GP\*AP\*CP\*GP\*AP\*GP\*GP\*GP\*CP\*CP\*AP\*GP\*TP\*GP\*CP\*CP\*GP\*TP\*AP\*C)-3')

Chain P: 54% 38% 8%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	217.60Å 217.60Å 165.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.50 – 3.44 49.25 – 3.45	Depositor EDS
% Data completeness (in resolution range)	91.6 (45.50-3.44) 95.2 (49.25-3.45)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.42 (at 3.48Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.280 , 0.320 0.350 , 0.335	Depositor DCC
$R_{free}$ test set	2000 reflections (3.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	140.7	Xtrriage
Anisotropy	0.404	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 20.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	14651	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.39% 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: DOC, DCP, MG

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	0.26	1/8025 (0.0%)	0.48	6/10886 (0.1%)
2	B	0.21	0/3015	0.39	0/4074
2	C	0.23	0/2966	0.45	2/4009 (0.0%)
3	T	0.47	0/567	0.88	0/872
4	P	0.49	0/492	0.77	0/758
All	All	0.27	1/15065 (0.0%)	0.50	8/20599 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	611	PRO	N-CD	-10.82	1.32	1.47

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	611	PRO	N-CA-CB	-11.49	89.52	103.30
2	C	135	PRO	CA-N-CD	-8.60	99.46	111.50
2	C	96	HIS	C-N-CD	-8.37	102.18	120.60
1	A	611	PRO	N-CD-CG	-6.82	92.97	103.20
1	A	752	LEU	C-N-CD	-6.27	106.81	120.60

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	7823	0	7729	293	0
2	B	2942	0	2939	58	0
2	C	2893	0	2880	97	0
3	T	508	0	282	15	0
4	P	455	0	245	8	0
5	A	2	0	0	0	0
6	A	28	0	12	4	0
All	All	14651	0	14087	452	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 16.

The worst 5 of 452 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:610:PHE:CE1	1:A:652:ILE:HD13	1.26	1.70
1:A:610:PHE:CD1	1:A:611:PRO:HG3	1.19	1.67
1:A:610:PHE:CE1	1:A:611:PRO:HG3	1.12	1.64
1:A:610:PHE:CG	1:A:611:PRO:HD3	1.29	1.59
1:A:610:PHE:CD1	1:A:611:PRO:CG	1.80	1.58

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	A	969/1222 (79%)	846 (87%)	106 (11%)	17 (2%)	8	39
2	B	355/472 (75%)	326 (92%)	27 (8%)	2 (1%)	25	62
2	C	350/472 (74%)	327 (93%)	18 (5%)	5 (1%)	11	44
All	All	1674/2166 (77%)	1499 (90%)	151 (9%)	24 (1%)	11	44

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	752	LEU
1	A	1070	PRO
2	C	97	PRO
1	A	648	PRO
1	A	767	ALA

### 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	830/1029 (81%)	737 (89%)	93 (11%)	6 26
2	B	325/415 (78%)	308 (95%)	17 (5%)	23 55
2	C	318/415 (77%)	295 (93%)	23 (7%)	14 46
All	All	1473/1859 (79%)	1340 (91%)	133 (9%)	9 36

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	186	LEU
2	C	243	THR
2	C	402	LEU
1	A	761	ASN
1	A	752	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	298	HIS
1	A	354	ASN
1	A	497	GLN
1	A	1134	HIS
2	C	427	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](#)

1 non-standard protein/DNA/RNA residue is 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
4	DOC	P	24	4,3	16,19,20	0.43	0	20,26,29	0.44	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	DOC	P	24	4,3	-	0/7/18/19	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

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

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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
6	DCP	A	4003	5	25,29,29	1.67	6 (24%)	37,45,45	1.32	5 (13%)

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
6	DCP	A	4003	5	-	0/22/34/34	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	4003	DCP	C4-N4	3.69	1.42	1.33
6	A	4003	DCP	C2-N1	-3.51	1.32	1.40
6	A	4003	DCP	C1'-N1	-3.02	1.40	1.48
6	A	4003	DCP	C6-C5	2.57	1.41	1.35
6	A	4003	DCP	PB-O1B	2.23	1.58	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	4003	DCP	PB-O3A-PA	-2.94	122.72	132.83
6	A	4003	DCP	PB-O3B-PG	-2.80	123.20	132.83
6	A	4003	DCP	O3'-C3'-C4'	2.68	120.36	110.10
6	A	4003	DCP	O3'-C3'-C2'	2.68	120.47	110.90
6	A	4003	DCP	O2-C2-N3	-2.27	118.64	122.33

There are no chirality outliers.

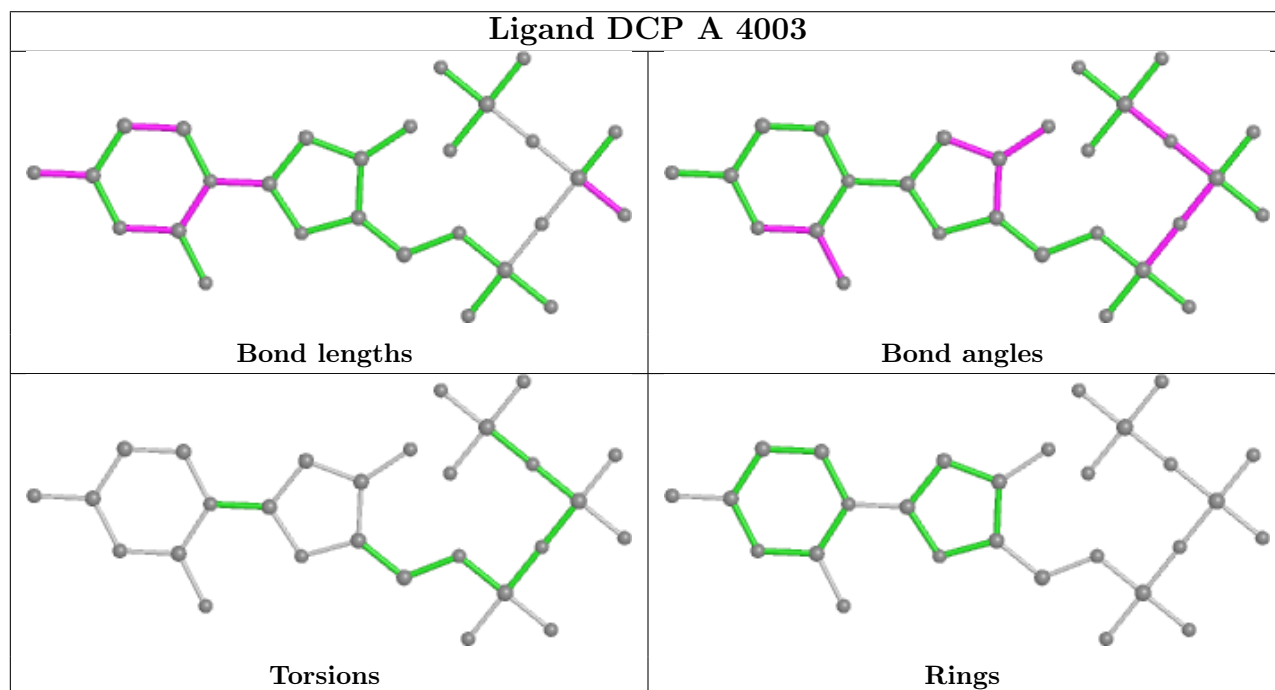
There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	4003	DCP	4	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

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

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

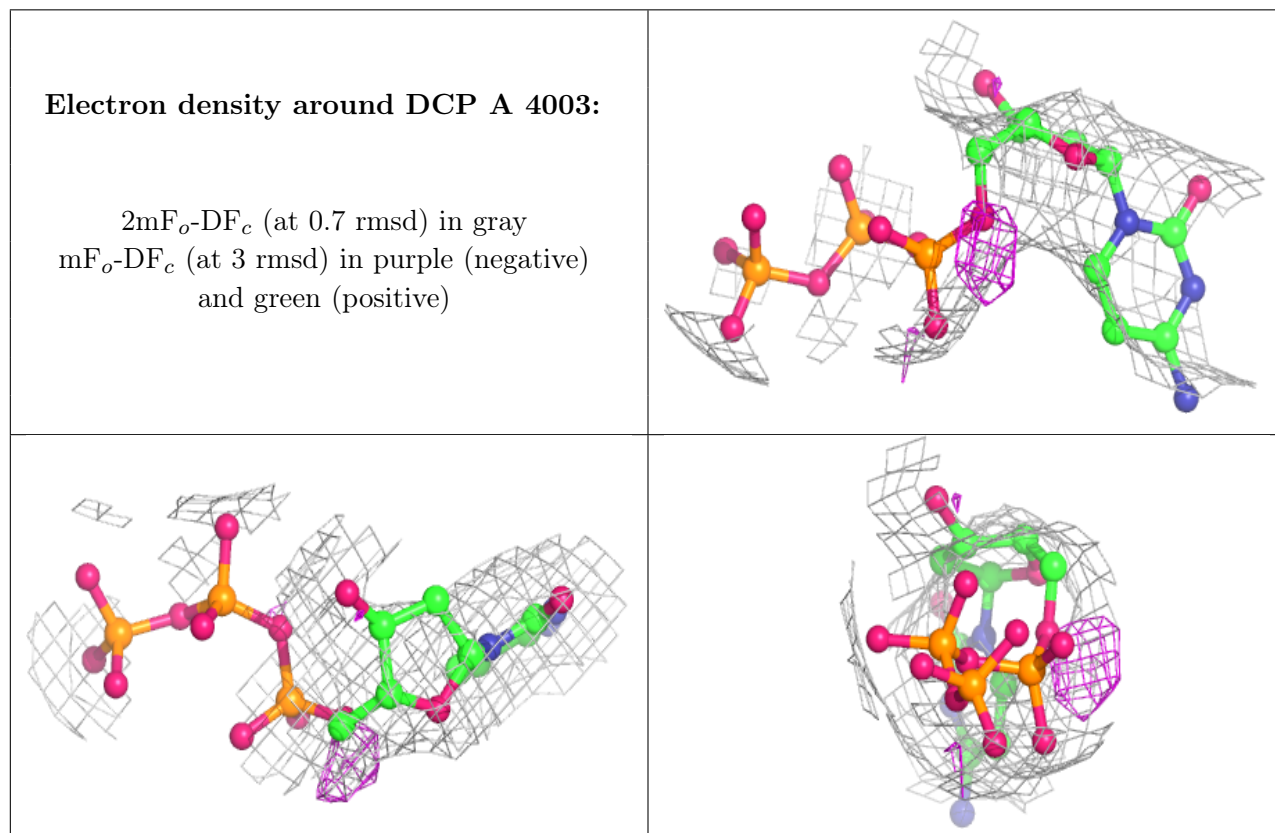
### 6.3 Carbohydrates [i](#)

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

### 6.4 Ligands [i](#)

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

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





## 6.5 Other polymers

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