

# Full wwPDB X-ray Structure Validation Report (i)

#### Sep 7, 2023 – 11:16 AM EDT

PDB ID	:	4F5O
Title	:	Open ternary complex of R283K DNA polymerase beta with a one metal bound
		dCTP analog
Authors	:	Freudenthal, B.D.; Beard, W.A.; Wilson, S.H.
Deposited on		
Resolution	:	2.00  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 A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

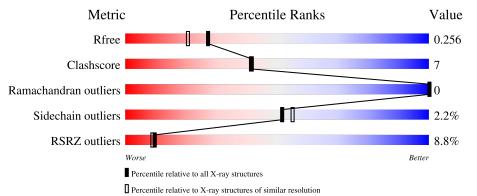
MolProbity	:	4.02b-467 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)		
EDS	:	2.35
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.00 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	А	335	9% 82%		13% • •				
2	Т	16	75%		25%				
3	Р	10	60%	40%					
4	D	5	60%	20%	20%				



## 2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 3519 atoms, of which 14 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 beta.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	321	Total 2559	C 1622	N 448	0 480	S 9	0	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	283	LYS	ARG	engineered mutation	UNP P06746

• Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*CP\*GP\*AP\*CP\*(8OG)P\*TP\*CP\*GP\* CP\*AP\*TP\*CP\*AP\*GP\*C)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	Т	16	Total	C	N 60	0	P 15	0	0	0
			322	153	60	94	10			

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

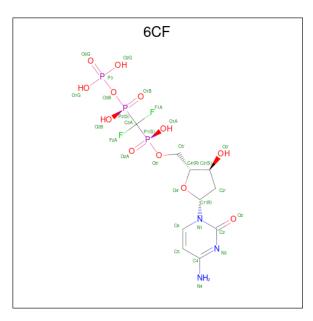
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
3	Р	10	Total 205	C 98	N 40	0 58	Р 9	0	0	0

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

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
4	D	5	Total 106	C 49	N 20	O 32	Р 5	0	0	0

• Molecule 5 is 2'-deoxy-5'-O-[(S)-{difluoro[(S)-hydroxy(phosphonooxy)phosphoryl]methyl}(h ydroxy)phosphoryl]cytidine (three-letter code: 6CF) (formula:  $C_{10}H_{16}F_2N_3O_{12}P_3$ ).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf		
5	Λ	1	Total	С	F	Η	Ν	Ο	Р	0	0
5	A	1	44	10	2	14	3	12	3	0	0

• Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	1	Total Mg 1 1	0	0

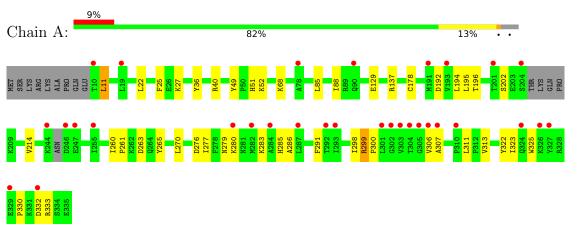
• Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	А	204	Total         O           204         204	0	0
7	Т	37	Total         O           37         37	0	0
7	Р	23	TotalO2323	0	0
7	D	18	Total O 18 18	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 polymerase beta

• Molecule 2: DNA (5'-D(\*CP\*CP\*GP\*AP\*CP\*(8OG)P\*TP\*CP\*GP\*CP\*AP\*TP\*CP\*AP\*GP \*C)-3')

Chain T:	75%		25%
C1 C2 C5 G6 G6 A11			
• Molecule	3: DNA (5'-D(*GP*CP*TP*GP*AP*7	ГР*GP*CP <sup>&gt;</sup>	*GP*A)-3')
Chain P:	60%	40	%
61 C2 64 A5 A10			
• Molecule	4: DNA (5'-D(P*GP*TP*CP*GP*G)-	3')	
Chain D:	60%	20%	20%



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	54.48Å 78.92Å 55.13Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $106.61^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	22.57 - 2.00	Depositor
Resolution (A)	22.57 - 2.00	EDS
% Data completeness	94.8 (22.57-2.00)	Depositor
(in resolution range)	$94.4\ (22.57-2.00)$	EDS
R <sub>merge</sub>	0.06	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.27 (at 1.99 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650), CNS	Depositor
B B.	0.203 , $0.267$	Depositor
$R, R_{free}$	0.198 , $0.256$	DCC
$R_{free}$ test set	3040 reflections $(10.05%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	30.9	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.34, $52.5$	EDS
L-test for twinning <sup>2</sup>	$< L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	0.027 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3519	wwPDB-VP
Average B, all atoms $(Å^2)$	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 6.68% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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,  $6\mathrm{CF},\,8\mathrm{OG}$ 

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	Boi	nd lengths	Bo	nd angles
IVIOI		RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.33	0/2611	0.49	0/3508
2	Т	0.65	0/333	1.19	3/508~(0.6%)
3	Р	0.68	0/230	1.20	0/354
4	D	1.04	1/118~(0.8%)	1.33	1/179~(0.6%)
All	All	0.45	1/3292~(0.0%)	0.72	4/4549~(0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	DG	OP3-P	-8.88	1.50	1.61

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
4	D	2	DT	O4'-C1'-N1	-7.45	102.79	108.00
2	Т	11	DA	O4'-C1'-N9	-5.62	104.07	108.00
2	Т	5	DC	O4'-C1'-C2'	-5.22	101.72	105.90
2	Т	2	DC	O4'-C1'-N1	5.05	111.53	108.00

All (4) bond angle outliers are listed below:

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	А	2559	0	2558	39	0
2	Т	322	0	180	0	0
3	Р	205	0	114	4	0
4	D	106	0	57	1	0
5	А	30	14	12	0	0
6	А	1	0	0	0	0
7	А	204	0	0	1	0
7	D	18	0	0	0	0
7	Р	23	0	0	0	0
7	Т	37	0	0	0	0
All	All	3505	14	2921	43	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 7.

All (43) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom 1	A + a	Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:A:279:ASN:O	1:A:283:LYS:HG3	1.47	1.15
1:A:279:ASN:O	1:A:283:LYS:CG	2.33	0.73
1:A:330:PRO:HA	1:A:333:ARG:HG3	1.71	0.72
1:A:36:TYR:CE1	1:A:40:ARG:HD2	2.26	0.69
1:A:68:LYS:HD3	4:D:1:DG:OP2	1.97	0.65
1:A:260:ILE:CG2	1:A:261:PRO:HD2	2.27	0.64
1:A:202:SER:OG	1:A:263:ASP:OD2	2.07	0.63
1:A:300:PRO:O	1:A:307:ALA:HA	2.00	0.62
1:A:36:TYR:CZ	1:A:40:ARG:HD2	2.37	0.59
3:P:1:DG:H2"	3:P:2:DC:H5'	1.86	0.58
1:A:311:LEU:HD13	1:A:322:TYR:CE1	2.37	0.58
1:A:270:LEU:HD23	7:A:672:HOH:O	2.04	0.58
1:A:36:TYR:CZ	1:A:40:ARG:CD	2.87	0.57
1:A:178:CYS:SG	1:A:194:LEU:HD13	2.44	0.57
1:A:260:ILE:HG23	1:A:261:PRO:HD2	1.87	0.57
1:A:279:ASN:HB3	1:A:283:LYS:HE3	1.87	0.56
1:A:286:ALA:HB1	1:A:291:PHE:HB2	1.88	0.55
1:A:277:ILE:HA	1:A:280:LYS:HD2	1.89	0.55
1:A:286:ALA:O	1:A:291:PHE:N	2.42	0.53
1:A:196:THR:HB	1:A:265:TYR:CD1	2.44	0.52
1:A:25:PHE:CZ	1:A:88:ILE:CG2	2.93	0.52
1:A:11:LEU:HD12	1:A:51:HIS:C	2.30	0.51
1:A:285:HIS:CD2	1:A:323:ILE:HB	2.46	0.51
1:A:298:ILE:HG23	1:A:298:ILE:O	2.13	0.48

Continued on next page...



		Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:A:195:LEU:HD21	1:A:214:VAL:HG21	1.95	0.47
1:A:49:TYR:CE1	1:A:51:HIS:HB2	2.50	0.47
1:A:311:LEU:HD13	1:A:322:TYR:CD1	2.50	0.47
1:A:22:LEU:CD2	1:A:85:LEU:HD13	2.46	0.46
1:A:306:VAL:HG12	1:A:306:VAL:O	2.16	0.46
1:A:25:PHE:CE2	1:A:88:ILE:CG2	2.98	0.46
1:A:25:PHE:CE2	1:A:88:ILE:HG21	2.52	0.45
1:A:276:ASP:O	1:A:280:LYS:HG3	2.17	0.44
3:P:1:DG:H2'	3:P:2:DC:C6	2.53	0.44
1:A:129:GLU:HG3	1:A:137:ARG:HE	1.83	0.44
1:A:299:ARG:HB2	1:A:307:ALA:HB1	1.99	0.43
1:A:277:ILE:O	1:A:280:LYS:HB2	2.19	0.42
1:A:25:PHE:CD2	1:A:88:ILE:HG21	2.54	0.42
1:A:27:LYS:HB2	1:A:36:TYR:CD1	2.54	0.42
1:A:52:LYS:HB2	1:A:52:LYS:HE2	1.72	0.42
1:A:260:ILE:HG22	1:A:261:PRO:HD2	2.01	0.42
3:P:4:DG:H2"	3:P:5:DA:OP2	2.20	0.41
1:A:286:ALA:HB1	1:A:291:PHE:CB	2.49	0.41
3:P:1:DG:H2"	3:P:2:DC:C5'	2.50	0.41

Continued from previous page...

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	А	318/335~(95%)	304 (96%)	14 (4%)	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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed Rotameric		Outliers	Percentiles	
1	А	275/295~(93%)	269~(98%)	6~(2%)	52 55	

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	А	11	LEU
1	А	192	ASP
1	А	299	ARG
1	А	313	VAL
1	А	325	TRP
1	А	332	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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	Bo	ond leng	ths	В	ond ang	les
Moi Type Chain	Unam	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2	
2	80G	Т	6	2	22,25,26	2.81	9 (40%)	30,37,40	2.23	9 (30%)



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.

Μ	ol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	2	80G	Т	6	2	-	0/7/21/22	0/3/3/3

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	Т	6	80G	O6-C6	8.68	1.40	1.23
2	Т	6	80G	O8-C8	6.21	1.34	1.23
2	Т	6	80G	C2-N2	3.29	1.42	1.34
2	Т	6	80G	C8-N9	-3.22	1.35	1.40
2	Т	6	80G	C4-N3	3.08	1.41	1.34
2	Т	6	80G	O3'-C3'	-2.54	1.38	1.43
2	Т	6	80G	C5'-C4'	-2.39	1.44	1.51
2	Т	6	80G	C2'-C3'	-2.20	1.47	1.52
2	Т	6	80G	C3'-C4'	-2.04	1.47	1.53

All (9) bond length outliers are listed below:

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	Т	6	80G	C5-C4-N3	-5.98	121.62	127.80
2	Т	6	80G	N9-C4-N3	4.87	131.38	125.81
2	Т	6	80G	O6-C6-C5	-3.76	118.63	127.24
2	Т	6	80G	C2-N3-C4	3.58	118.67	112.30
2	Т	6	80G	C1'-N9-C4	-3.55	121.04	126.54
2	Т	6	80G	O4'-C1'-N9	2.85	111.16	108.29
2	Т	6	80G	C2-N1-C6	-2.81	119.97	125.10
2	Т	6	80G	O5'-C5'-C4'	2.73	118.27	108.99
2	Т	6	80G	C5-C6-N1	2.55	119.91	112.31

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 2 ligands modelled in this entry, 1 is 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	Mol Type Chain Res		Link	Bond lengths			Bond angles			
IVIOI	туре	Ullalli	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
5	6CF	А	401	6	26,31,31	1.51	6 (23%)	36,50,50	1.01	2 (5%)

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	6CF	А	401	6	-	4/21/46/46	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
5	А	401	6CF	P3-O3G	3.53	1.61	1.50
5	А	401	6CF	P1-O5'	2.59	1.61	1.57
5	А	401	6CF	P1-01A	2.49	1.61	1.56
5	А	401	6CF	P2-O2B	2.42	1.61	1.56
5	А	401	6CF	P2-O3B	2.36	1.61	1.58
5	А	401	6CF	C2-N1	-2.36	1.34	1.40

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
5	А	401	$6\mathrm{CF}$	O5'-P1-O2A	-3.23	109.76	115.48
5	А	401	6CF	O5'-P1-C3A	2.16	109.02	102.06

There are no chirality outliers.

All (4) torsion outliers are listed below:

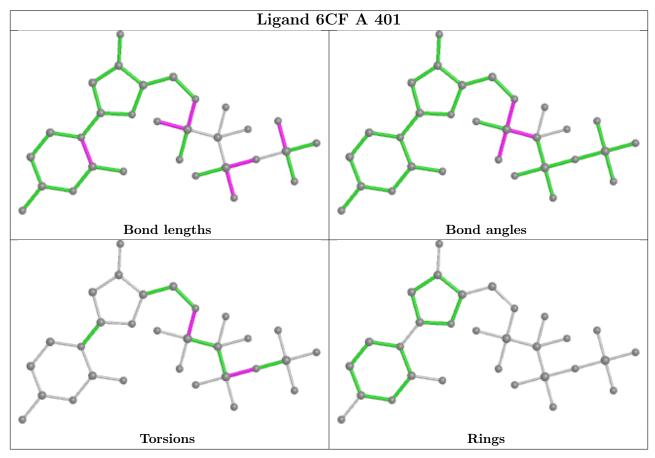


Mol	Chain	Res	Type	Atoms
5	А	401	6CF	C5'-O5'-P1-O2A
5	А	401	6CF	P3-O3B-P2-O2B
5	А	401	6CF	C5'-O5'-P1-O1A
5	А	401	6CF	P3-O3B-P2-O1B

There are no ring outliers.

No monomer is involved in short contacts.

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 then 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)

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,  $95^{th}$  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	$\langle RSRZ \rangle$	# RSRZ > 2	$OWAB(Å^2)$	Q<0.9
1	А	321/335~(95%)	0.40	31 (9%) 7 7	18, 35, 77, 103	0
2	Т	15/16~(93%)	-0.26	0 100 100	25, 38, 53, 55	0
3	Р	10/10~(100%)	-0.41	0 100 100	22, 34, 41, 43	0
4	D	5/5~(100%)	-0.59	0 100 100	25, 26, 37, 40	0
All	All	351/366~(95%)	0.34	31 (8%) 10 9	18, 35, 76, 103	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	246	ASP	6.3
1	А	304	THR	5.6
1	А	327	TYR	4.9
1	А	303	VAL	4.8
1	А	306	VAL	4.7
1	А	287	LEU	4.6
1	А	244	LYS	4.4
1	А	307	ALA	4.3
1	А	305	GLY	4.0
1	А	280	LYS	4.0
1	А	326	LYS	3.9
1	А	78	ALA	3.5
1	А	10	THR	3.3
1	А	324	GLN	3.2
1	А	193	VAL	2.6
1	А	301	LEU	2.5
1	А	255	ILE	2.5
1	А	284	ALA	2.5
1	А	332	ASP	2.5
1	А	293	ILE	2.5
1	А	90	GLN	2.3

Continued on next page...



Mol	Chain	Res	Type	RSRZ
1	А	19	LEU	2.3
1	А	204	SER	2.3
1	А	201	THR	2.3
1	А	247	GLU	2.3
1	А	329	GLU	2.2
1	А	282	MET	2.1
1	А	191	MET	2.1
1	А	310	PRO	2.1
1	А	292	THR	2.0
1	А	302	GLY	2.0

Continued from previous page...

### 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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q < 0.9
2	80G	Т	6	23/24	0.86	0.18	45,54,61,65	0

#### 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

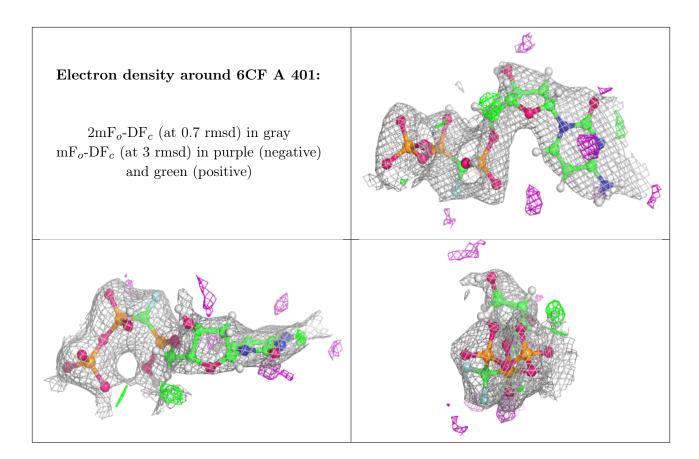
#### 6.4 Ligands (i)

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,  $95^{th}$  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	$B-factors(A^2)$	Q<0.9
6	MG	А	402	1/1	0.85	0.11	$50,\!50,\!50,\!50$	0
5	6CF	А	401	30/30	0.90	0.18	42,61,74,76	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.

