

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

### Feb 19, 2024 – 10:33 AM EST

PDB ID : 4JV2

Title : Ternary complex of gamma-OHPDG adduct modified dna with dna (-1 primer)

polymerase iv and incoming datp

Authors: Banerjee, S.; Shanmugam, G.; Stone, M.P.

Deposited on : 2013-03-25

Resolution : 2.74 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

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)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

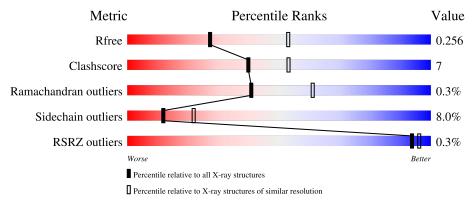
Validation Pipeline (wwPDB-VP) : 2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 2.74 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	Similar resolution		
Metric	$(\# \mathrm{Entries})$	$(\#  ext{Entries},  ext{resolution range}( ext{Å}))$		
$R_{free}$	130704	1271 (2.76-2.72)		
Clashscore	141614	1322 (2.76-2.72)		
Ramachandran outliers	138981	1297 (2.76-2.72)		
Sidechain outliers	138945	1298 (2.76-2.72)		
RSRZ outliers	127900	1243 (2.76-2.72)		

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	y of chain		
1	A	347	79%		18%	• •
2	В	15	53%	40%		7%
3	С	13	46%	46%		8%



# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3392 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.

$\mathbf{Mol}$	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace			
1	A	341	Total 2744	C 1760	N 472	O 505	S 7	0	0	0	

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	expression tag	UNP Q97W02
A	-4	HIS	-	expression tag	UNP Q97W02
A	-3	HIS	-	expression tag	UNP Q97W02
A	-2	HIS	-	expression tag	UNP Q97W02
A	-1	HIS	-	expression tag	UNP Q97W02
A	0	HIS	-	expression tag	UNP Q97W02

• Molecule 2 is a DNA chain called DNA (5'-D(P\*TP\*(HN1)P\*GP\*AP\*AP\*TP\*CP\*CP\*TP \*TP\*CP\*CP\*CP\*C)-3').

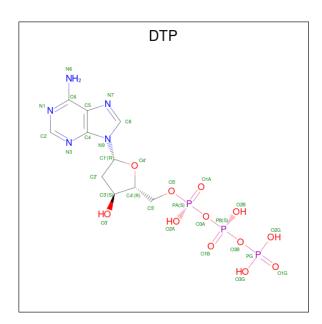
$\mathbf{Mol}$	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	В	15	Total 303	C 146	N 49	O 93	P 15	0	0	0

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

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	C	13	Total	С	N	О	Р	0	0	0
3		10	273	129	57	75	12	0	U	U

• Molecule 4 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
4	A	1	Total		_	0	P	0	0
			30	10	5	12	3		

• Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

• Molecule 6 is water.

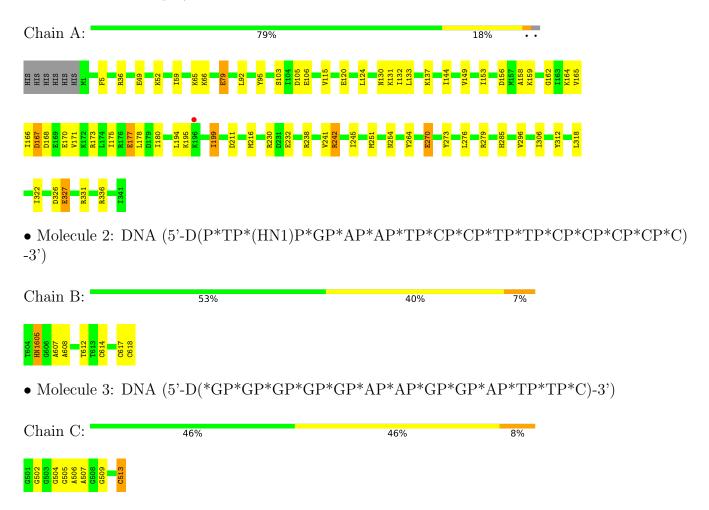
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	28	Total O 28 28	0	0
6	В	7	Total O 7 7	0	0
6	С	5	Total O 5 5	0	0



# 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 IV





# 4 Data and refinement statistics (i)

Property	Value	Source	
Space group	P 21 21 2	Depositor	
Cell constants	94.58Å 103.31Å 52.78Å	Donositon	
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 90.00° 90.00°	Depositor	
Resolution (Å)	42.09 - 2.74	Depositor	
Resolution (A)	42.09 - 2.74	EDS	
% Data completeness	99.0 (42.09-2.74)	Depositor	
(in resolution range)	94.4 (42.09-2.74)	EDS	
$R_{merge}$	0.09	Depositor	
$R_{sym}$	(Not available)	Depositor	
$< I/\sigma(I) > 1$	1.54 (at 2.73Å)	Xtriage	
Refinement program	PHENIX 1.7.2_869	Depositor	
D D.	0.203 , 0.265	Depositor	
$R, R_{free}$	0.196 , $0.256$	DCC	
$R_{free}$ test set	1401 reflections (10.01%)	wwPDB-VP	
Wilson B-factor (Å <sup>2</sup> )	56.6	Xtriage	
Anisotropy	0.240	Xtriage	
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.33, 30.3	EDS	
L-test for twinning <sup>2</sup>	$ < L > = 0.48, < L^2> = 0.31$	Xtriage	
Estimated twinning fraction	No twinning to report.	Xtriage	
$F_o, F_c$ correlation	0.94	EDS	
Total number of atoms	3392	wwPDB-VP	
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP	

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.92% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $<L^2>$  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: DTP, HN1, CA

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		
IVIOI		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.46	0/2783	0.56	0/3736	
2	В	0.83	0/306	1.56	$6/465 \ (1.3\%)$	
3	С	0.86	0/308	1.63	6/476 (1.3%)	
All	All	0.55	0/3397	0.88	$12/4677 \ (0.3\%)$	

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
3	С	513	DC	O4'-C1'-N1	12.37	116.66	108.00
3	С	502	DG	O4'-C1'-N9	8.26	113.78	108.00
3	С	504	DG	O4'-C1'-N9	8.15	113.71	108.00
2	В	607	DA	C1'-O4'-C4'	-6.42	103.68	110.10
2	В	617	DC	O4'-C1'-N1	6.21	112.34	108.00
2	В	618	DC	O4'-C1'-N1	5.83	112.08	108.00
2	В	612	DT	N3-C4-O4	5.82	123.39	119.90
2	В	607	DA	O4'-C1'-C2'	-5.59	101.43	105.90
3	С	509	DG	O4'-C1'-N9	5.49	111.84	108.00
3	С	507	DA	O4'-C1'-C2'	-5.42	101.56	105.90
3	С	509	DG	P-O5'-C5'	-5.41	112.24	120.90
2	В	614	DC	C1'-O4'-C4'	-5.25	104.85	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	2744	0	2889	37	0
2	В	303	0	172	6	0
3	С	273	0	147	3	0
4	A	30	0	12	4	0
5	A	2	0	0	0	0
6	A	28	0	0	3	0
6	В	7	0	0	0	0
6	С	5	0	0	0	0
All	All	3392	0	3220	46	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 (46) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
	1100111 _	$\operatorname{distance} (\text{\AA})$	overlap (Å)
2:B:605:HN1:C1'	2:B:605:HN1:N9	1.73	1.51
2:B:605:HN1:H	3:C:513:DC:C4	2.03	0.92
2:B:605:HN1:C1'	2:B:605:HN1:C4	2.57	0.83
2:B:605:HN1:N9	2:B:605:HN1:C2'	2.53	0.68
1:A:242:ARG:HD2	1:A:245:ILE:HD11	1.77	0.66
1:A:270:GLU:OE2	1:A:312:TYR:OH	2.15	0.65
4:A:401:DTP:O2A	6:A:501:HOH:O	2.14	0.63
1:A:5:PHE:CZ	1:A:106:GLU:HG2	2.34	0.63
1:A:158:ALA:HB2	1:A:164:LYS:HB2	1.82	0.61
2:B:605:HN1:H	3:C:513:DC:N3	2.14	0.61
3:C:505:DG:H2"	3:C:506:DA:H5"	1.84	0.59
1:A:242:ARG:HH11	1:A:242:ARG:HA	1.68	0.58
1:A:166:ILE:HG23	1:A:170:GLU:HB3	1.87	0.56
1:A:285:HIS:HD2	6:A:505:HOH:O	1.89	0.54
1:A:242:ARG:NH1	2:B:608:DA:OP1	2.41	0.54
4:A:401:DTP:H8	4:A:401:DTP:O5'	2.09	0.53
1:A:36:ARG:NH2	1:A:254:ASN:OD1	2.36	0.51
1:A:79:GLU:H	1:A:79:GLU:CD	2.15	0.51
1:A:105:ASP:OD1	4:A:401:DTP:H5'1	2.10	0.50
1:A:106:GLU:OE2	6:A:501:HOH:O	2.18	0.50
1:A:5:PHE:HZ	1:A:106:GLU:HG2	1.75	0.49
1:A:92:LEU:HD21	1:A:132:ILE:HD11	1.95	0.49
1:A:49:GLU:HA	1:A:52:LYS:HE3	1.96	0.47

Continued on next page...



Continued from previous page...

A 4 1	A4 0	Interatomic	Clash
Atom-1	Atom-2	${\rm distance}\ ({\rm \AA})$	overlap (Å)
1:A:180:ILE:HD13	1:A:194:LEU:HD13	1.97	0.47
1:A:318:LEU:O	1:A:322:ILE:HG13	2.16	0.46
1:A:171:VAL:O	1:A:175:ILE:HG13	2.15	0.46
4:A:401:DTP:O2B	4:A:401:DTP:H5'2	2.16	0.46
1:A:173:ARG:O	1:A:177:GLU:HB2	2.16	0.46
1:A:95:TYR:HD2	1:A:124:LEU:HD11	1.82	0.44
1:A:156:ASP:HA	1:A:159:LYS:HE3	2.00	0.43
1:A:115:VAL:HG22	1:A:120:GLU:HB2	2.00	0.43
1:A:199:ILE:H	1:A:199:ILE:HG12	1.63	0.43
1:A:251:MET:HA	1:A:264:TYR:CE1	2.52	0.43
1:A:273:TYR:OH	1:A:306:ILE:O	2.31	0.43
1:A:173:ARG:NH1	1:A:177:GLU:HG2	2.34	0.42
1:A:276:LEU:O	1:A:279:ARG:HB2	2.18	0.42
1:A:242:ARG:HA	1:A:242:ARG:NH1	2.34	0.42
1:A:327:GLU:H	1:A:327:GLU:HG2	1.51	0.42
1:A:144:ILE:HB	1:A:165:VAL:HG22	2.02	0.42
1:A:103:SER:OG	1:A:106:GLU:HB2	2.19	0.42
1:A:273:TYR:HA	1:A:276:LEU:HD12	2.01	0.42
1:A:167:ASP:OD1	1:A:168:ASP:N	2.53	0.41
1:A:66:LYS:HB2	1:A:66:LYS:NZ	2.35	0.41
1:A:158:ALA:O	1:A:162:GLY:HA3	2.20	0.41
1:A:245:ILE:HD13	1:A:245:ILE:HA	1.85	0.41
1:A:149:VAL:O	1:A:153:ILE:HG13	2.21	0.41

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	Percentile	s
1	A	339/347 (98%)	320 (94%)	18 (5%)	1 (0%)	41 61	

### All (1) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	167	ASP

### 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	Analysed Rotameric		Percentiles	
1	A	300/306 (98%)	276 (92%)	24 (8%)	12 22	

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ILE
1	A	65	LYS
1	A	79	GLU
1	A	130	ASN
1	A	131	LYS
1	A	133	LEU
1	A	137	LYS
1	A	177	GLU
1	A	178	LEU
1	A	195	LYS
1	A	199	ILE
1	A	211	ASP
1	A	216	MET
1	A	230	ARG
1	A	232	GLU
1	A	238	ARG
1	A	241	VAL
1	A	242	ARG
1	A	270	GLU
1	A	296	VAL
1	A	326	ASP
1	A	327	GLU
1	A	331	ARG
1	A	336	ARG

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		in Res Link		Bond lengths			Bond angles			
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	HN1	В	605	2	18,29,37	4.29	4 (22%)	24,43,55	3.47	11 (45%)

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
2	HN1	В	605	2	-	3/3/31/43	0/4/4/4

#### All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(\mathbf{\mathring{A}})$	$Ideal(\AA)$
2	В	605	HN1	O5'-C5'	-12.25	1.14	1.44
2	В	605	HN1	O6-C6	9.60	1.41	1.22
2	В	605	HN1	C1'-N9	8.05	1.73	1.49
2	В	605	HN1	C3-N2	-2.17	1.42	1.47

#### All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	605	HN1	O5'-C5'-C4'	9.50	141.31	108.99
2	В	605	HN1	N1-C2-N2	-7.94	116.44	124.27

Continued on next page...



$\alpha \cdots$	, r	•	
Continued	trom	mromonie	maaa
-	110116	DICULUUS	Duuc
	J	1	1

Mol	Chain	$\operatorname{Res}$	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(^{o})$	$\operatorname{Ideal}({}^{o})$
2	В	605	HN1	C2'-C1'-N9	-5.33	101.99	114.27
2	В	605	HN1	O6-C6-C5	-4.56	116.12	124.19
2	В	605	HN1	C5-C6-N1	4.06	119.87	113.91
2	В	605	HN1	C3-C-C1	3.51	118.18	110.34
2	В	605	HN1	O11-C1-C	3.41	118.68	110.92
2	В	605	HN1	C-C3-N2	2.86	121.68	113.29
2	В	605	HN1	N3-C2-N1	2.77	120.52	115.25
2	В	605	HN1	C8-N7-C5	2.71	108.16	102.99
2	В	605	HN1	C4'-O4'-C1'	-2.41	103.63	109.45

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	В	605	HN1	O4'-C4'-C5'-O5'
2	В	605	HN1	C3'-C4'-C5'-O5'
2	В	605	HN1	C4'-C5'-O5'-P

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	$\operatorname{Res}$	Type	Clashes	Symm-Clashes
2	В	605	HN1	5	0

### 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

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		
MIOI			nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
4	DTP	A	401	5	26,32,32	1.18	3 (11%)	30,50,50	1.48	5 (16%)

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.

$\mathbf{Mol}$	$\mathbf{Type}$	Chain	Res	Link	Chirals	Torsions	Rings
4	DTP	A	401	5	-	4/18/34/34	0/3/3/3

### All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\operatorname{\AA})$	$\mathbf{Ideal}( exttt{\AA})$
4	A	401	DTP	C6-N6	2.93	1.44	1.34
4	A	401	DTP	C5'-C4'	-2.31	1.44	1.51
4	A	401	DTP	O3'-C3'	-2.23	1.38	1.43

### All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
4	A	401	DTP	N3-C2-N1	-4.27	122.01	128.68
4	A	401	DTP	O3G-PG-O3B	2.99	114.68	104.64
4	A	401	DTP	C4-C5-N7	-2.96	106.31	109.40
4	A	401	DTP	O5'-C5'-C4'	2.96	119.17	108.99
4	A	401	DTP	PB-O3B-PG	-2.12	125.57	132.83

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	DTP	PB-O3A-PA-O5'
4	A	401	DTP	PG-O3B-PB-O3A
4	A	401	DTP	PA-O3A-PB-O2B
4	A	401	DTP	PG-O3B-PB-O1B

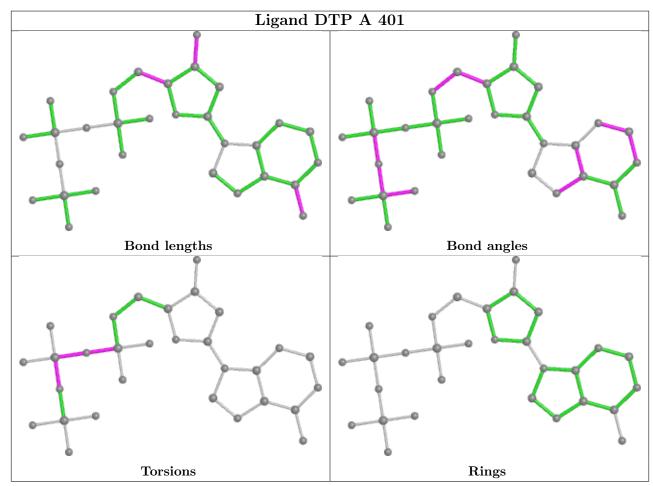
There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	401	DTP	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 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	<RSRZ $>$	# RSRZ > 2	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	A	341/347 (98%)	-0.05	1 (0%) 94 96	40, 54, 74, 88	0
2	В	14/15 (93%)	-0.55	0 100 100	46, 55, 65, 65	0
3	С	13/13 (100%)	-0.62	0 100 100	42, 47, 68, 72	0
All	All	368/375 (98%)	-0.09	1 (0%) 94 96	40, 54, 74, 88	0

#### All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	196	LYS	2.0

### 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	HN1	В	605	26/34	0.92	0.17	40,55,65,70	0

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

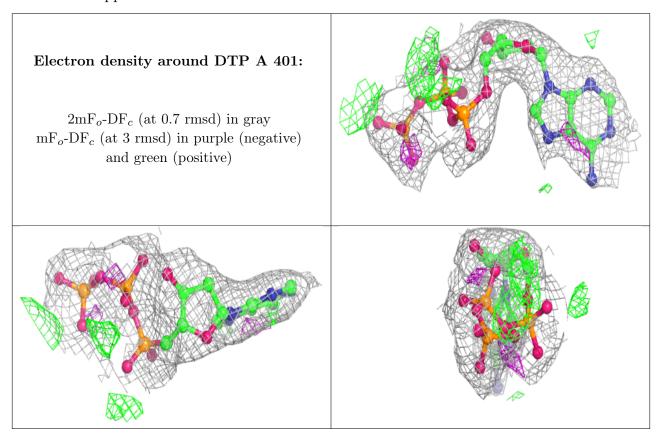
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,



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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
5	CA	A	402	1/1	0.79	0.19	50,50,50,50	0
5	CA	A	403	1/1	0.90	0.31	63,63,63,63	0
4	DTP	A	401	30/30	0.95	0.14	43,49,61,67	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.

