



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 06:14 am GMT

PDB ID : 4FS1  
Title : Base pairing mechanism of N2,3-ethenoguanine with dTTP by human polymerase iota  
Authors : Zhao, L.  
Deposited on : 2012-06-26  
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

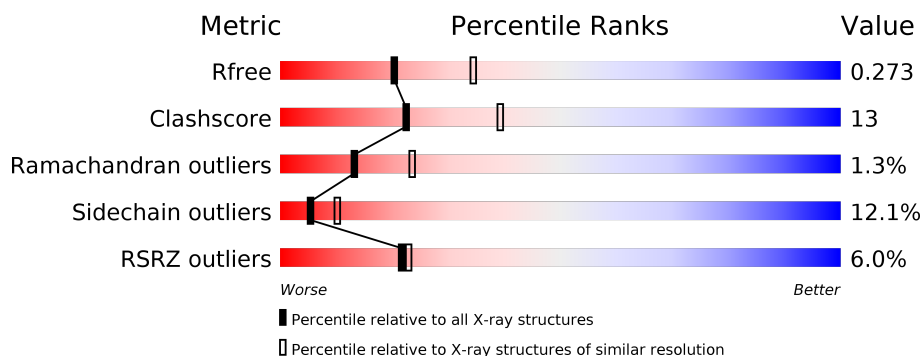
# 1 Overall quality at a glance

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))
$R_{free}$	111664	4155 (2.50-2.50)
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)
RSRZ outliers	108989	4058 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	B	18	
1	C	18	
2	A	420	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3370 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 DNA chain called DNA 5'-D(\*TP\*CP\*TP\*(EFG)P\*GP\*GP\*GP\*TP\*CP\*CP\*TP\*AP\*GP\*GP\*AP\*CP\*CP\*(DOC))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	7	Total	C	N	O	P	0	0	0
			142	67	29	39	7			
1	C	9	Total	C	F	N	O	P	0	0
			186	90	1	32	55	8		

- Molecule 2 is a protein called DNA polymerase iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	383	Total	C	N	O	S	0	0	0
			2944	1856	514	554	20			

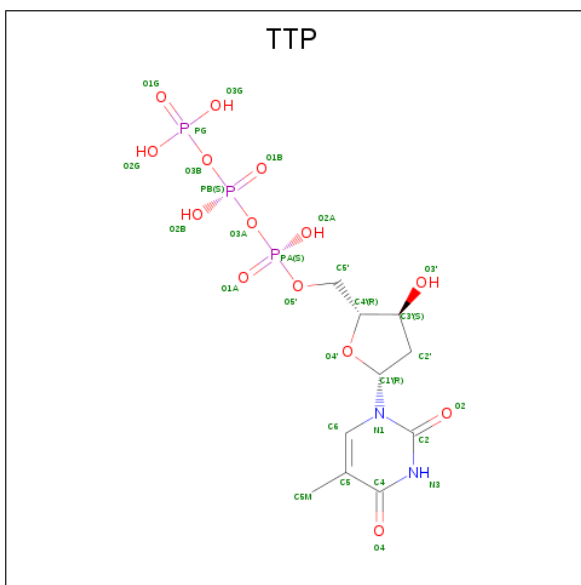
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		

- Molecule 5 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>2</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
5	A	1	Total	C	N	O	P	0	0
			29	10	2	14	3		

- Molecule 6 is water.

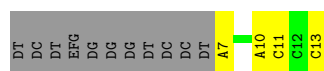
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total O 1 1	0	0
6	C	5	Total O 5 5	0	0
6	A	59	Total O 59 59	0	0

### 3 Residue-property plots


These plots are drawn for all protein, RNA and DNA 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 5'-D(\*TP\*CP\*TP\*(EFG)P\*GP\*GP\*GP\*TP\*CP\*CP\*TP\*AP\*GP\*GP\*AP\*CP\*CP\*(DOC))-3'

Chain B: 



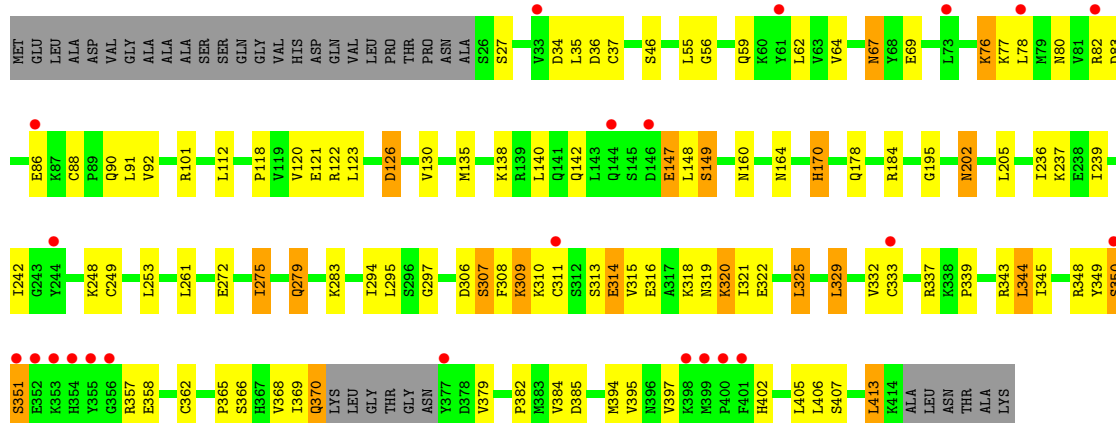
- Molecule 1: DNA 5'-D(\*TP\*CP\*TP\*(EFG)P\*GP\*GP\*GP\*TP\*CP\*CP\*TP\*AP\*GP\*GP\*AP\*CP\*CP\*(DOC))-3'

Chain C: 



- Molecule 2: DNA polymerase iota

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.47Å 97.47Å 203.54Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	84.41 – 2.50 29.30 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (84.41-2.50) 99.8 (29.30-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.23 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.215 , 0.270 0.219 , 0.273	Depositor DCC
$R_{free}$ test set	1179 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.4	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3370	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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: NA, MG, TTP, EFG, DOC

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	B	0.44	0/139	1.08	0/212
1	C	0.93	1/178 (0.6%)	1.10	3/271 (1.1%)
2	A	0.74	0/2984	0.89	5/4035 (0.1%)
All	All	0.74	1/3301 (0.0%)	0.92	8/4518 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	846	DC	O3'-P	-5.76	1.54	1.61

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	126	ASP	CB-CG-OD1	5.92	123.63	118.30
2	A	344	LEU	CA-CB-CG	5.85	128.75	115.30
2	A	101	ARG	NE-CZ-NH2	5.78	123.19	120.30
2	A	122	ARG	NE-CZ-NH1	-5.56	117.52	120.30
1	C	845	DC	P-O3'-C3'	5.53	126.34	119.70
1	C	842	DG	O5'-P-OP2	-5.46	100.79	105.70
2	A	34	ASP	CB-CG-OD2	5.34	123.11	118.30
1	C	846	DC	P-O3'-C3'	5.01	125.71	119.70

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	B	142	0	78	4	1
1	C	186	0	103	8	1
2	A	2944	0	2976	70	0
3	A	3	0	0	0	0
4	A	1	0	0	0	0
5	A	29	0	13	8	0
6	A	59	0	0	0	0
6	B	1	0	0	0	0
6	C	5	0	0	0	0
All	All	3370	0	3170	81	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 13.

All (81) 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:A:236:ILE:HD11	2:A:261:LEU:HD22	1.46	0.97
2:A:126:ASP:OD2	5:A:505:TTP:H5'	1.73	0.88
2:A:82:ARG:HG3	2:A:83:ASP:N	1.95	0.81
1:C:839:DT:H2''	1:C:840:EFG:OP2	1.81	0.81
2:A:366:SER:HA	2:A:369:ILE:HD12	1.64	0.79
1:C:839:DT:C2'	1:C:840:EFG:OP2	2.31	0.78
1:C:844:DT:H2''	1:C:845:DC:H5'	1.69	0.73
2:A:337:ARG:NH1	2:A:413:LEU:CD1	2.51	0.73
1:C:840:EFG:H10	1:C:840:EFG:O4'	1.89	0.72
2:A:345:ILE:HB	2:A:407:SER:HB3	1.71	0.72
5:A:505:TTP:O2B	5:A:505:TTP:H3'	1.89	0.71
2:A:370:GLN:OE1	2:A:370:GLN:C	2.29	0.71
2:A:202:ASN:C	2:A:202:ASN:HD22	1.96	0.69
2:A:309:LYS:O	2:A:311:CYS:N	2.28	0.66
2:A:82:ARG:HG3	2:A:83:ASP:H	1.60	0.66
2:A:348:ARG:NH2	2:A:358:GLU:OE1	2.28	0.65
2:A:120:VAL:HG22	2:A:130:VAL:HG22	1.77	0.65
2:A:365:PRO:HB2	2:A:368:VAL:HG23	1.78	0.64
2:A:236:ILE:HD11	2:A:261:LEU:CD2	2.23	0.63
1:B:13:DOC:H2'	5:A:505:TTP:C6	2.34	0.62
2:A:76:LYS:HE3	2:A:77:LYS:H	1.66	0.60
2:A:55:LEU:HD12	2:A:92:VAL:O	2.02	0.60

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:337:ARG:NH1	2:A:413:LEU:HD13	2.17	0.58
2:A:319:ASN:HA	2:A:322:GLU:OE1	2.03	0.58
2:A:164:ASN:H	2:A:170:HIS:HD2	1.50	0.58
2:A:202:ASN:HD21	2:A:205:LEU:H	1.53	0.57
2:A:253:LEU:HD11	2:A:272:GLU:HG2	1.85	0.57
2:A:202:ASN:ND2	2:A:205:LEU:H	2.03	0.56
2:A:36:ASP:HA	5:A:505:TTP:O2G	2.05	0.56
1:C:846:DC:H2''	1:C:847:DT:H5'	1.87	0.56
2:A:349:TYR:N	2:A:402:HIS:O	2.38	0.56
2:A:126:ASP:CG	5:A:505:TTP:H5'2	2.26	0.55
2:A:118:PRO:O	2:A:295:LEU:HD23	2.08	0.54
1:B:13:DOC:H2''	5:A:505:TTP:H5'1	1.89	0.53
1:B:10:DA:H2'	1:B:11:DC:C6	2.43	0.53
2:A:55:LEU:H	2:A:67:ASN:ND2	2.07	0.53
2:A:314:GLU:O	2:A:315:VAL:HB	2.09	0.53
2:A:325:LEU:HD13	2:A:329:LEU:HD13	1.91	0.51
2:A:350:SER:HA	2:A:351:SER:HB3	1.92	0.51
2:A:321:ILE:HD11	2:A:406:LEU:HD22	1.92	0.51
2:A:126:ASP:OD2	5:A:505:TTP:O2A	2.30	0.50
2:A:349:TYR:O	2:A:350:SER:HB2	2.11	0.50
2:A:295:LEU:HD13	2:A:297:GLY:N	2.27	0.49
2:A:337:ARG:NH1	2:A:413:LEU:HD11	2.27	0.49
2:A:202:ASN:C	2:A:202:ASN:ND2	2.66	0.48
2:A:320:LYS:HZ3	2:A:320:LYS:HA	1.79	0.48
2:A:332:VAL:HG11	2:A:339:PRO:HG3	1.95	0.48
1:C:841:DG:OP2	2:A:307:SER:HB2	2.14	0.48
2:A:35:LEU:N	2:A:35:LEU:HD23	2.29	0.47
2:A:135:MET:HA	2:A:138:LYS:HE3	1.96	0.47
2:A:147:GLU:C	2:A:149:SER:H	2.17	0.47
2:A:239:ILE:HB	2:A:242:ILE:HD12	1.95	0.47
2:A:275:ILE:O	2:A:275:ILE:HG13	2.13	0.47
2:A:343:ARG:HG2	2:A:345:ILE:HD11	1.98	0.46
2:A:318:LYS:HE3	2:A:384:VAL:HG11	1.97	0.45
2:A:348:ARG:NH1	2:A:394:MET:O	2.49	0.45
2:A:325:LEU:HD13	2:A:329:LEU:CD1	2.46	0.45
2:A:55:LEU:H	2:A:67:ASN:HD21	1.62	0.45
2:A:80:ASN:H	2:A:80:ASN:HD22	1.64	0.45
2:A:321:ILE:O	2:A:325:LEU:HB2	2.18	0.44
2:A:306:ASP:O	2:A:405:LEU:HD12	2.17	0.44
2:A:140:LEU:HD23	2:A:140:LEU:HA	1.88	0.44
2:A:80:ASN:H	2:A:80:ASN:ND2	2.15	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:112:LEU:HD23	2:A:112:LEU:C	2.39	0.43
2:A:55:LEU:HG	2:A:56:GLY:N	2.33	0.43
2:A:350:SER:HA	2:A:351:SER:CB	2.49	0.43
2:A:275:ILE:O	2:A:279:GLN:HB2	2.19	0.42
2:A:295:LEU:HD13	2:A:297:GLY:H	1.83	0.42
2:A:121:GLU:HB2	2:A:294:ILE:O	2.20	0.42
2:A:36:ASP:O	2:A:37:CYS:C	2.58	0.42
1:C:840:EFG:H8	2:A:59:GLN:OE1	2.19	0.42
2:A:64:VAL:HA	2:A:78:LEU:HD23	2.02	0.42
2:A:184:ARG:HD2	2:A:195:GLY:O	2.20	0.42
2:A:332:VAL:CG1	2:A:339:PRO:HD3	2.50	0.41
1:B:13:DOC:H2'	5:A:505:TTP:C5	2.55	0.41
2:A:337:ARG:CZ	2:A:413:LEU:HD13	2.50	0.41
2:A:164:ASN:H	2:A:170:HIS:CD2	2.34	0.41
2:A:237:LYS:O	2:A:237:LYS:HG2	2.21	0.41
1:C:839:DT:O4'	1:C:839:DT:O2	2.39	0.41
2:A:314:GLU:C	2:A:316:GLU:H	2.24	0.41
2:A:397:VAL:HG22	2:A:397:VAL:O	2.22	0.40

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)
1:B:7:DA:P	1:C:847:DT:O3'[12_564]	1.86	0.34

## 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
2	A	379/420 (90%)	345 (91%)	29 (8%)	5 (1%)	<b>13</b> <b>23</b>

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	310	LYS
2	A	351	SER
2	A	313	SER
2	A	350	SER
2	A	86	GLU

### 5.3.2 Protein sidechains ⓘ

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
2	A	330/376 (88%)	290 (88%)	40 (12%)	5 10

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	27	SER
2	A	46	SER
2	A	62	LEU
2	A	67	ASN
2	A	69	GLU
2	A	76	LYS
2	A	88	CYS
2	A	90	GLN
2	A	91	LEU
2	A	123	LEU
2	A	142	GLN
2	A	147	GLU
2	A	148	LEU
2	A	149	SER
2	A	160	ASN
2	A	170	HIS
2	A	178	GLN
2	A	202	ASN
2	A	248	LYS
2	A	249	CYS
2	A	275	ILE
2	A	279	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	A	283	LYS
2	A	307	SER
2	A	308	PHE
2	A	309	LYS
2	A	314	GLU
2	A	320	LYS
2	A	325	LEU
2	A	329	LEU
2	A	333	CYS
2	A	344	LEU
2	A	357	ARG
2	A	362	CYS
2	A	370	GLN
2	A	379	VAL
2	A	382	PRO
2	A	385	ASP
2	A	395	VAL
2	A	413	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	58	GLN
2	A	67	ASN
2	A	80	ASN
2	A	159	ASN
2	A	160	ASN
2	A	170	HIS
2	A	189	ASN
2	A	202	ASN
2	A	217	GLN
2	A	227	GLN
2	A	262	GLN
2	A	340	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are 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
1	DOC	B	13	1	13,19,20	1.09	2 (15%)	12,26,29	1.37	2 (16%)
1	EFG	C	840	1	18,28,29	1.93	5 (27%)	18,42,45	2.15	6 (33%)

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
1	DOC	B	13	1	-	0/3/18/19	0/2/2/2
1	EFG	C	840	1	-	0/3/25/26	0/4/4/4

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	840	EFG	C6-C5	-3.34	1.35	1.41
1	C	840	EFG	C5-C4	-3.00	1.33	1.40
1	B	13	DOC	O5'-C5'	-2.33	1.41	1.44
1	B	13	DOC	C2-N3	-2.28	1.33	1.38
1	C	840	EFG	C8-N9	-2.07	1.34	1.36
1	C	840	EFG	O4'-C1'	3.88	1.46	1.41
1	C	840	EFG	C6-N1	4.33	1.40	1.33

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	840	EFG	C4'-O4'-C1'	-5.65	103.93	109.83
1	C	840	EFG	O4'-C1'-C2'	-3.63	101.88	105.74
1	C	840	EFG	F-C2'-C1'	-2.68	103.78	109.28
1	B	13	DOC	C5-C4-N3	-2.40	118.88	121.69

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	13	DOC	C2'-C1'-N1	-2.31	108.14	112.48
1	C	840	EFG	C10-C9-N3	2.30	108.72	106.82
1	C	840	EFG	O5'-C5'-C4'	2.51	117.61	109.00
1	C	840	EFG	O4'-C4'-C5'	2.75	118.55	109.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	13	DOC	3	0
1	C	840	EFG	4	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 4 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	$\# Z  > 2$
5	TTP	A	505	3	22,30,30	0.64	0	27,47,47	1.92	4 (14%)

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	TTP	A	505	3	-	0/18/34/34	0/2/2/2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	505	TTP	PB-O3A-PA	-3.61	120.48	132.63
5	A	505	TTP	PB-O3B-PG	-3.61	120.49	132.63
5	A	505	TTP	C5-C6-N1	-2.59	119.35	122.15
5	A	505	TTP	C4-N3-C2	6.46	120.64	115.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	505	TTP	8	0

## 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	B	6/18 (33%)	-0.87	0 100 100	37, 48, 56, 60	0
1	C	8/18 (44%)	-0.24	1 (12%) 4 3	39, 43, 53, 115	0
2	A	383/420 (91%)	0.18	23 (6%) 22 23	26, 55, 102, 159	0
All	All	397/456 (87%)	0.16	24 (6%) 22 23	26, 55, 102, 159	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	355	TYR	15.2
2	A	353	LYS	7.8
2	A	352	GLU	6.5
2	A	351	SER	6.2
2	A	377	TYR	6.1
2	A	354	HIS	4.6
2	A	400	PRO	4.4
2	A	244	TYR	3.7
2	A	82	ARG	3.5
2	A	61	TYR	3.3
1	C	839	DT	3.1
2	A	350	SER	2.9
2	A	399	MET	2.8
2	A	73	LEU	2.7
2	A	144	GLN	2.7
2	A	78	LEU	2.7
2	A	86	GLU	2.6
2	A	356	GLY	2.4
2	A	311	CYS	2.3
2	A	33	VAL	2.2
2	A	398	LYS	2.2
2	A	401	PHE	2.2
2	A	333	CYS	2.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	A	146	ASP	2.1

## 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<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
1	EFG	C	840	25/26	0.96	0.12	48,58,88,91	0
1	DOC	B	13	18/19	0.99	0.10	28,36,43,47	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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
3	MG	A	501	1/1	0.62	0.36	61,61,61,61	0
5	TTP	A	505	29/29	0.68	0.29	30,46,70,80	8
4	NA	A	504	1/1	0.91	0.77	58,58,58,58	0
3	MG	A	502	1/1	0.96	0.08	45,45,45,45	0
3	MG	A	503	1/1	0.99	0.24	47,47,47,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.