



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:04 AM GMT

PDB ID : 3H4D  
Title : Ternary complex of human DNA polymerase iota with template U/T and incoming dGTP  
Authors : Jain, R.; Nair, D.T.; Johnson, R.E.; Prakash, L.; Prakash, S.; Aggarwal, A.K.  
Deposited on : 2009-04-18  
Resolution : 2.20 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  
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

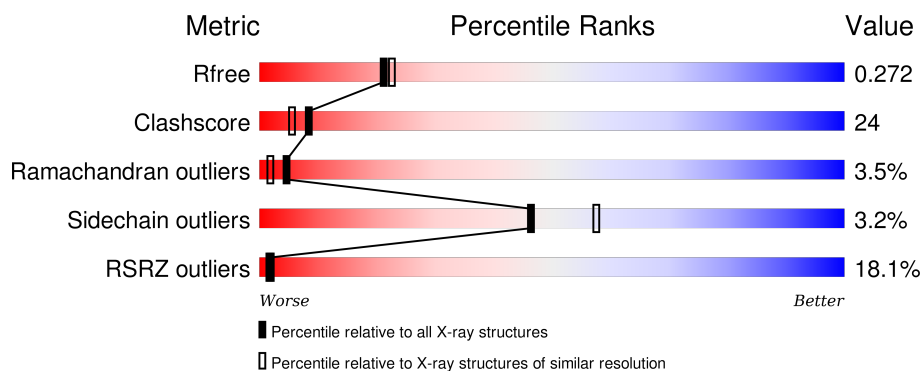
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	A	390	<div> <div>18%</div> <div>60%</div> <div>32%</div> <div>.</div> <div>.</div> </div>
2	P	7	<div> <div>14%</div> <div>57%</div> <div>29%</div> </div>
3	T	9	<div> <div>11%</div> <div>56%</div> <div>33%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	DGT	T	875	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3489 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 iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	2	0
			2844	1788	493	542	21			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	P	0	0	0
			139	67	29	37	6			

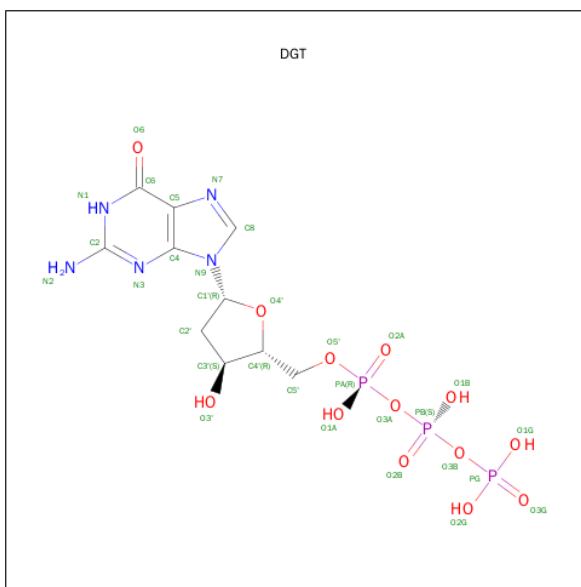
- Molecule 3 is a DNA chain called 5'-D(\*TP\*(BRU)P\*GP\*GP\*GP\*TP\*CP\*CP\*T).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	T	9	Total	Br	C	N	O	P	0	1	0
			201	2	96	31	63	9			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 5 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	T	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

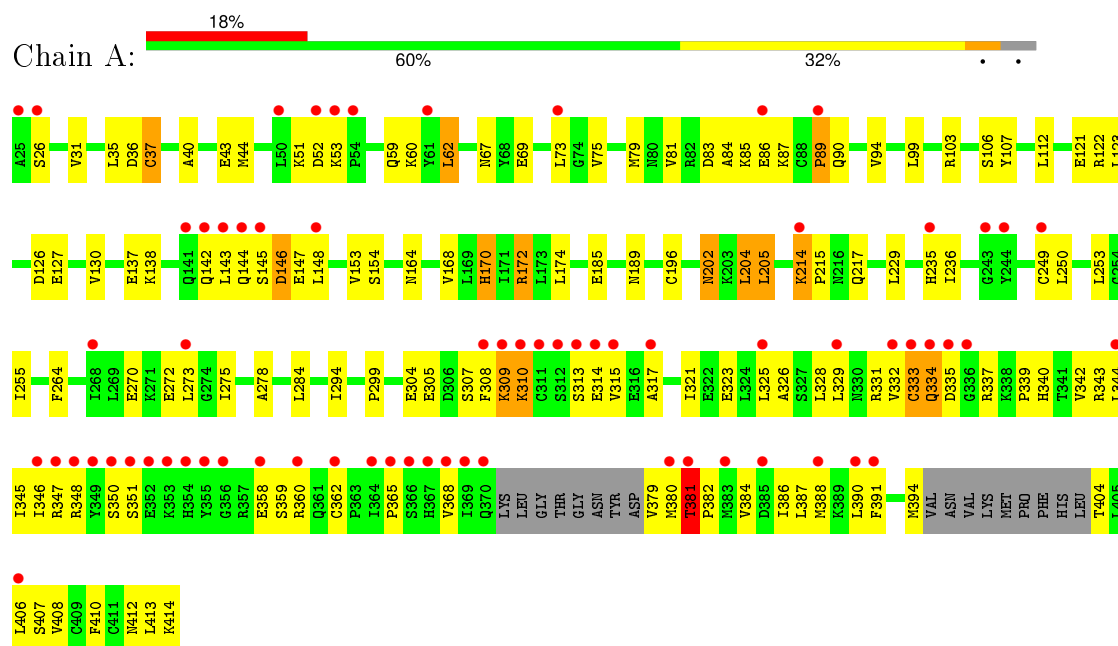
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	240	Total O 240 240	0	0
6	P	9	Total O 9 9	0	0
6	T	23	Total O 23 23	0	0

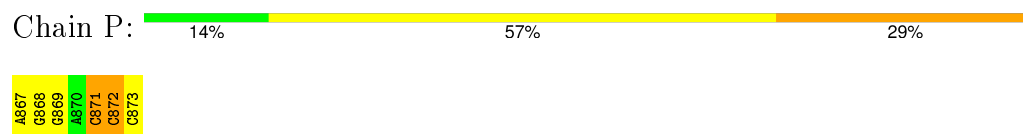
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

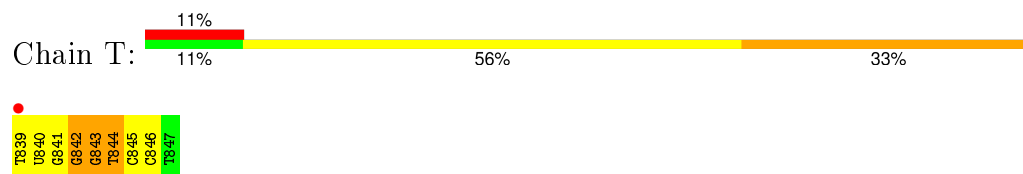
#### • Molecule 1: DNA polymerase iota



#### • Molecule 2: 5'-D(\*AP\*GP\*GP\*AP\*CP\*CP\*(DOC))



#### • Molecule 3: 5'-D(\*TP\*(BRU)P\*GP\*GP\*GP\*TP\*CP\*CP\*T)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.92Å 97.92Å 203.39Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.68 – 2.20 36.68 – 2.20	Depositor EDS
% Data completeness (in resolution range)	95.7 (36.68-2.20) 95.9 (36.68-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.55 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.254 , 0.270 0.255 , 0.272	Depositor DCC
$R_{free}$ test set	2306 reflections (8.71%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.2	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 58.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 30019 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3489	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.90% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DOC, MG, BRU, DGT

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	0.34	0/2892	0.61	2/3913 (0.1%)
2	P	0.39	0/136	0.76	0/208
3	T	0.39	0/178	0.88	0/271
All	All	0.35	0/3206	0.64	2/4392 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	P	0	2
3	T	0	3
All	All	0	5

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	214	LYS	N-CA-C	5.33	125.38	111.00
1	A	123	LEU	CB-CA-C	5.24	120.16	110.20

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	P	871	DC	Sidechain
2	P	872	DC	Sidechain
3	T	842	DG	Sidechain

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
3	T	843	DG	Sidechain
3	T	844	DT	Sidechain

## 5.2 Too-close contacts

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	2844	0	2815	120	0
2	P	139	0	79	7	0
3	T	201	0	111	23	0
4	A	2	0	0	0	0
5	T	31	0	11	10	0
6	A	240	0	0	5	0
6	P	9	0	0	3	0
6	T	23	0	0	3	0
All	All	3489	0	3016	150	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 24.

All (150) 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:126:ASP:OD2	5:T:875:DGT:H5'	1.42	1.16
3:T:842:DG:H2''	3:T:843:DG:H5''	1.39	1.04
1:A:344:LEU:HD11	1:A:387:LEU:HD22	1.39	1.00
1:A:44:MET:HE2	1:A:51:LYS:HA	1.50	0.93
1:A:137:GLU:HG3	6:A:578:HOH:O	1.69	0.93
3:T:841:DG:H2''	3:T:842:DG:H5''	1.50	0.92
1:A:164:ASN:H	1:A:170:HIS:HD2	1.24	0.85
3:T:842:DG:H2''	3:T:843:DG:C5'	2.09	0.82
3:T:840[B]:BRU:O4	5:T:875:DGT:O6	1.99	0.80
3:T:841:DG:C2'	3:T:842:DG:H5''	2.12	0.80
1:A:44:MET:CE	1:A:67:ASN:HD22	1.99	0.76
1:A:365:PRO:O	1:A:368:VAL:HG22	1.87	0.75
1:A:79:MET:HE3	1:A:84:ALA:HA	1.68	0.75

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:344:LEU:HD21	1:A:387:LEU:HB3	1.67	0.74
2:P:867:DA:H1'	2:P:868:DG:H5'	1.67	0.74
1:A:321:ILE:O	1:A:325:LEU:HD13	1.88	0.74
3:T:845:DC:H5''	6:T:176:HOH:O	1.86	0.74
3:T:840[B]:BRU:O4	5:T:875:DGT:C6	2.29	0.74
2:P:868:DG:N7	6:P:174:HOH:O	2.21	0.73
1:A:126:ASP:OD2	5:T:875:DGT:C5'	2.30	0.73
1:A:202:ASN:ND2	1:A:205:LEU:H	1.90	0.70
1:A:44:MET:HE1	1:A:67:ASN:HD22	1.57	0.69
3:T:839:DT:H2''	3:T:840[B]:BRU:O5'	1.93	0.67
3:T:841:DG:H2''	3:T:842:DG:C5'	2.25	0.67
1:A:106:SER:OG	1:A:122:ARG:NH2	2.29	0.66
2:P:869:DG:H1'	6:P:224:HOH:O	1.94	0.66
1:A:62:LEU:HD11	3:T:839:DT:O2	1.95	0.65
1:A:270:GLU:HG3	1:A:275:ILE:HA	1.78	0.65
1:A:325:LEU:HD11	1:A:387:LEU:CD1	2.29	0.63
1:A:331:ARG:HD2	6:A:457:HOH:O	1.97	0.63
1:A:347:ARG:HD2	1:A:404:THR:OG1	1.99	0.63
1:A:384:VAL:O	1:A:388:MET:HG2	1.99	0.63
1:A:202:ASN:HD21	1:A:205:LEU:H	1.47	0.62
1:A:315:VAL:C	1:A:317:ALA:H	2.02	0.62
1:A:51:LYS:O	1:A:53:LYS:N	2.33	0.62
1:A:335:ASP:OD2	1:A:337:ARG:HD3	2.00	0.61
1:A:43:GLU:HG3	1:A:94:VAL:HG21	1.83	0.61
1:A:164:ASN:H	1:A:170:HIS:CD2	2.13	0.61
1:A:51:LYS:C	1:A:53:LYS:H	2.04	0.60
1:A:386:ILE:O	1:A:390:LEU:HG	2.01	0.60
1:A:313:SER:O	1:A:314:GLU:HB2	2.01	0.59
1:A:388:MET:O	1:A:391:PHE:HB3	2.02	0.59
1:A:107:TYR:OH	1:A:299:PRO:HG3	2.03	0.59
1:A:79:MET:CE	1:A:84:ALA:HA	2.32	0.58
2:P:869:DG:N3	6:P:224:HOH:O	2.32	0.58
1:A:112:LEU:HD23	1:A:112:LEU:C	2.24	0.58
5:T:875:DGT:PB	5:T:875:DGT:H5'A	2.43	0.58
3:T:841:DG:H1'	3:T:842:DG:H5''	1.86	0.58
1:A:126:ASP:OD1	1:A:127:GLU:HG3	2.04	0.57
1:A:345:ILE:HG12	1:A:359:SER:HB3	1.86	0.57
1:A:412:ASN:OD1	1:A:414:LYS:HE3	2.04	0.57
1:A:325:LEU:CD1	1:A:408:VAL:HG21	2.35	0.57
3:T:845:DC:C5'	6:T:176:HOH:O	2.48	0.56
1:A:196:CYS:SG	1:A:214:LYS:O	2.63	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:343:ARG:C	1:A:344:LEU:HD12	2.26	0.56
1:A:365:PRO:HB2	1:A:368:VAL:HG13	1.87	0.56
1:A:249:CYS:SG	1:A:273:LEU:HD21	2.46	0.55
1:A:44:MET:HE3	1:A:67:ASN:HD22	1.72	0.55
1:A:62:LEU:HD11	3:T:839:DT:C2	2.42	0.54
1:A:381:THR:OG1	1:A:382:PRO:HD3	2.07	0.54
1:A:308:PHE:O	1:A:310:LYS:N	2.40	0.54
1:A:344:LEU:CD1	1:A:387:LEU:HD22	2.27	0.53
1:A:137:GLU:CG	6:A:578:HOH:O	2.44	0.53
1:A:153:VAL:HB	1:A:174:LEU:HD22	1.91	0.53
2:P:868:DG:H2''	2:P:869:DG:OP2	2.09	0.52
1:A:305:GLU:HG3	1:A:407:SER:HB3	1.91	0.52
1:A:144:GLN:NE2	6:A:523:HOH:O	2.41	0.52
1:A:185:GLU:OE2	1:A:189[A]:ASN:ND2	2.41	0.52
3:T:839:DT:H2'	3:T:840[A]:BRU:BR	2.65	0.52
1:A:202:ASN:C	1:A:202:ASN:HD22	2.13	0.52
1:A:60:LYS:HE2	1:A:307:SER:O	2.09	0.51
1:A:414:LYS:HD3	6:A:586:HOH:O	2.10	0.51
3:T:839:DT:H2''	3:T:840[A]:BRU:O5'	2.10	0.51
5:T:875:DGT:O2B	5:T:875:DGT:H5'A	2.11	0.51
1:A:79:MET:HE3	1:A:84:ALA:CA	2.40	0.51
1:A:59:GLN:OE1	3:T:840[B]:BRU:H6	2.11	0.51
1:A:236:ILE:HD12	1:A:250:LEU:CD1	2.41	0.51
1:A:325:LEU:O	1:A:329:LEU:HD13	2.11	0.50
1:A:344:LEU:O	1:A:359:SER:HA	2.11	0.50
1:A:214:LYS:HZ1	5:T:875:DGT:PG	2.34	0.50
1:A:346:ILE:HG22	1:A:406:LEU:HD23	1.94	0.50
1:A:204:LEU:HG	1:A:284:LEU:HD22	1.94	0.49
2:P:872:DC:H2''	2:P:873:DOC:OP2	2.12	0.49
1:A:215:PRO:O	1:A:217:GLN:HG3	2.13	0.49
1:A:235:HIS:CD2	1:A:236:ILE:HG12	2.47	0.49
1:A:304:GLU:HG3	1:A:328:LEU:CD2	2.42	0.49
3:T:841:DG:C1'	3:T:842:DG:H5''	2.42	0.49
1:A:365:PRO:HB2	1:A:368:VAL:HG22	1.95	0.48
1:A:347:ARG:HG2	1:A:348:ARG:N	2.28	0.48
1:A:387:LEU:HA	1:A:390:LEU:HD12	1.94	0.48
1:A:309:LYS:O	1:A:310:LYS:C	2.52	0.48
1:A:138:LYS:O	1:A:142:GLN:HG2	2.14	0.48
1:A:36:ASP:O	1:A:37:CYS:C	2.52	0.48
1:A:332:VAL:CG1	1:A:339:PRO:HD3	2.44	0.47
1:A:345:ILE:HA	1:A:358:GLU:O	2.14	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:325:LEU:HD11	1:A:408:VAL:HG21	1.97	0.47
1:A:69:GLU:O	1:A:73:LEU:HD13	2.15	0.47
1:A:304:GLU:HG3	1:A:328:LEU:HD21	1.96	0.47
1:A:344:LEU:HD11	1:A:387:LEU:CD2	2.28	0.46
1:A:379:VAL:O	1:A:382:PRO:HD2	2.15	0.46
1:A:85:LYS:O	1:A:89:PRO:HG3	2.15	0.46
1:A:85:LYS:C	1:A:87:LYS:H	2.18	0.46
5:T:875:DGT:O5'	5:T:875:DGT:H8	2.16	0.46
1:A:325:LEU:HD11	1:A:387:LEU:HD11	1.98	0.46
2:P:871:DC:H2''	2:P:872:DC:OP2	2.16	0.46
1:A:99:LEU:HD12	3:T:842:DG:H4'	1.99	0.45
1:A:103:ARG:CZ	1:A:331:ARG:NH1	2.79	0.45
1:A:323:GLU:O	1:A:326:ALA:HB3	2.17	0.45
1:A:148:LEU:HD23	1:A:148:LEU:O	2.17	0.45
1:A:315:VAL:C	1:A:317:ALA:N	2.70	0.45
1:A:75:VAL:HA	1:A:79:MET:HE2	1.99	0.44
1:A:339:PRO:HG3	1:A:410:PHE:HD2	1.82	0.44
3:T:845:DC:H2''	3:T:846:DC:OP2	2.17	0.44
1:A:144:GLN:O	1:A:146:ASP:N	2.42	0.44
1:A:121:GLU:HB2	1:A:294:ILE:O	2.18	0.44
1:A:202:ASN:OD1	1:A:205:LEU:HD22	2.17	0.44
1:A:270:GLU:HA	1:A:278:ALA:CB	2.48	0.44
1:A:153:VAL:HG22	1:A:154:SER:N	2.33	0.44
1:A:315:VAL:O	1:A:317:ALA:N	2.50	0.43
3:T:840[A]:BRU:O2	5:T:875:DGT:N2	2.48	0.43
1:A:325:LEU:HD11	1:A:387:LEU:HD12	1.98	0.43
3:T:840[B]:BRU:O2	3:T:840[B]:BRU:H2'	2.19	0.43
1:A:196:CYS:HA	1:A:217:GLN:O	2.19	0.43
1:A:360:ARG:HG2	1:A:394:MET:CG	2.48	0.43
5:T:875:DGT:O1B	5:T:875:DGT:H3'	2.18	0.43
1:A:81:VAL:O	1:A:84:ALA:HB3	2.19	0.43
3:T:845:DC:P	6:T:176:HOH:O	2.77	0.42
1:A:168:VAL:HG22	1:A:172:ARG:HD2	2.01	0.42
1:A:236:ILE:HD12	1:A:250:LEU:HD12	2.00	0.42
1:A:40:ALA:O	1:A:44:MET:HG3	2.19	0.42
1:A:381:THR:CB	1:A:382:PRO:HD3	2.49	0.42
1:A:73:LEU:HD12	1:A:73:LEU:N	2.34	0.42
1:A:143:LEU:HA	1:A:147:GLU:OE2	2.18	0.42
1:A:86:GLU:HG3	1:A:86:GLU:O	2.20	0.42
1:A:340:HIS:ND1	1:A:414:LYS:HD2	2.34	0.42
1:A:255:ILE:HD11	1:A:264:PHE:CG	2.55	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:342:VAL:HG21	1:A:387:LEU:HD21	2.02	0.42
1:A:380:MET:O	1:A:381:THR:C	2.58	0.42
1:A:83:ASP:O	1:A:87:LYS:HB2	2.20	0.42
1:A:31:VAL:CG1	1:A:130:VAL:HB	2.50	0.42
1:A:51:LYS:C	1:A:53:LYS:N	2.71	0.41
1:A:112:LEU:O	1:A:112:LEU:HD23	2.20	0.41
1:A:253:LEU:HD21	1:A:272:GLU:OE2	2.20	0.41
1:A:35:LEU:N	1:A:35:LEU:HD12	2.35	0.41
1:A:137:GLU:OE2	1:A:172:ARG:NE	2.53	0.41
1:A:365:PRO:HB2	1:A:368:VAL:CG1	2.51	0.41
1:A:333:CYS:O	1:A:334:GLN:C	2.58	0.41
3:T:843:DG:H2'	3:T:844:DT:H72	2.02	0.41
1:A:343:ARG:HG2	1:A:344:LEU:N	2.35	0.40
1:A:337:ARG:HH21	1:A:413:LEU:HB2	1.86	0.40

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	369/390 (95%)	322 (87%)	34 (9%)	13 (4%)	<b>4</b> <b>2</b>

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	ASP
1	A	309	LYS
1	A	350	SER
1	A	26	SER
1	A	37	CYS
1	A	145	SER
1	A	146	ASP

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	310	LYS
1	A	334	GLN
1	A	333	CYS
1	A	351	SER
1	A	89	PRO
1	A	381	THR

### 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
1	A	313/354 (88%)	303 (97%)	10 (3%)	46 57

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	62	LEU
1	A	90	GLN
1	A	170	HIS
1	A	172	ARG
1	A	202	ASN
1	A	204	LEU
1	A	205	LEU
1	A	229	LEU
1	A	362	CYS
1	A	381	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	58	GLN
1	A	67	ASN
1	A	90	GLN
1	A	170	HIS
1	A	202	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	235	HIS
1	A	262	GLN
1	A	334	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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
2	DOC	P	873	3,2	11,19,20	0.75	0	14,26,29	1.14	1 (7%)
3	BRU	T	840[A]	3	13,21,22	1.87	3 (23%)	16,30,33	4.04	2 (12%)
3	BRU	T	840[B]	3	13,21,22	1.86	3 (23%)	16,30,33	4.16	2 (12%)

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	DOC	P	873	3,2	-	0/3/18/19	0/2/2/2
3	BRU	T	840[A]	3	-	0/3/21/22	0/2/2/2
3	BRU	T	840[B]	3	-	0/3/21/22	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	840[A]	BRU	C6-N1	3.04	1.39	1.35
3	T	840[B]	BRU	C6-N1	3.08	1.39	1.35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	840[A]	BRU	C4-N3	3.52	1.39	1.33
3	T	840[B]	BRU	C4-N3	3.67	1.39	1.33
3	T	840[B]	BRU	C4-C5	4.43	1.44	1.38
3	T	840[A]	BRU	C4-C5	4.67	1.44	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	840[B]	BRU	C5-C4-N3	-8.46	114.96	124.00
3	T	840[A]	BRU	C5-C4-N3	-8.43	115.00	124.00
2	P	873	DOC	C2-N3-C4	2.98	119.82	115.61
3	T	840[A]	BRU	C4-N3-C2	13.66	127.06	115.25
3	T	840[B]	BRU	C4-N3-C2	14.14	127.47	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	873	DOC	1	0
3	T	840[A]	BRU	3	0
3	T	840[B]	BRU	5	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates 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		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	DGT	T	875	-	25,33,33	1.98	8 (32%)	35,52,52	2.31	12 (34%)

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	DGT	T	875	-	-	0/18/34/34	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	T	875	DGT	O4'-C4'	-4.14	1.35	1.45
5	T	875	DGT	PA-O1A	-3.16	1.41	1.54
5	T	875	DGT	PB-O1B	-2.22	1.45	1.54
5	T	875	DGT	C6-C5	2.25	1.45	1.41
5	T	875	DGT	PB-O2B	2.27	1.59	1.51
5	T	875	DGT	PG-O3G	2.77	1.60	1.51
5	T	875	DGT	C6-N1	3.99	1.40	1.33
5	T	875	DGT	C2-N1	4.23	1.42	1.35

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	T	875	DGT	C5-C6-N1	-6.83	114.26	123.59
5	T	875	DGT	C6-C5-C4	-4.93	115.00	120.90
5	T	875	DGT	N3-C2-N1	-3.62	121.93	127.44
5	T	875	DGT	O1B-PB-O2B	-2.06	101.38	112.53
5	T	875	DGT	PA-O3A-PB	-2.03	127.02	132.73
5	T	875	DGT	O3A-PA-O5'	2.17	108.70	102.94
5	T	875	DGT	O2G-PG-O3G	2.20	117.65	110.58
5	T	875	DGT	O2G-PG-O1G	2.22	115.84	107.38
5	T	875	DGT	O5'-C5'-C4'	2.30	117.60	109.12
5	T	875	DGT	C2'-C1'-N9	2.42	120.05	114.16
5	T	875	DGT	C4'-O4'-C1'	2.81	116.56	109.47
5	T	875	DGT	C6-N1-C2	6.02	124.29	115.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	T	875	DGT	10	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 [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	373/390 (95%)	1.01	69 (18%) <b>2</b> <b>1</b>	18, 44, 96, 113	0
2	P	6/7 (85%)	1.25	0 <b>100</b> <b>100</b>	45, 54, 60, 64	0
3	T	8/9 (88%)	0.86	1 (12%) <b>5</b> <b>5</b>	32, 37, 52, 82	0
All	All	387/406 (95%)	1.01	70 (18%) <b>2</b> <b>2</b>	18, 45, 95, 113	0

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	355	TYR	15.7
1	A	354	HIS	13.6
1	A	25	ALA	9.2
1	A	349	TYR	7.5
1	A	336	GLY	7.3
1	A	368	VAL	7.2
1	A	312	SER	7.2
1	A	244	TYR	7.2
1	A	350	SER	7.0
1	A	311	CYS	6.5
1	A	353	LYS	6.5
1	A	352	GLU	6.3
1	A	313	SER	4.9
1	A	351	SER	4.8
1	A	26	SER	4.8
1	A	333	CYS	4.8
1	A	348	ARG	4.8
1	A	388	MET	4.5
1	A	366	SER	4.1
1	A	391	PHE	4.0
1	A	145	SER	4.0
1	A	356	GLY	4.0
1	A	385	ASP	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	310	LYS	3.6
1	A	315	VAL	3.6
3	T	839	DT	3.5
1	A	249	CYS	3.5
1	A	334	GLN	3.4
1	A	308	PHE	3.4
1	A	364	ILE	3.3
1	A	329	LEU	3.2
1	A	390	LEU	3.2
1	A	335	ASP	3.1
1	A	370	GLN	3.1
1	A	317	ALA	3.1
1	A	50	LEU	3.1
1	A	383	MET	3.0
1	A	314	GLU	3.0
1	A	406	LEU	3.0
1	A	347	ARG	3.0
1	A	61	TYR	2.9
1	A	141	GLN	2.9
1	A	243	GLY	2.9
1	A	332	VAL	2.8
1	A	358	GLU	2.8
1	A	214	LYS	2.8
1	A	54	PRO	2.8
1	A	142	GLN	2.7
1	A	365	PRO	2.7
1	A	144	GLN	2.7
1	A	235	HIS	2.5
1	A	52	ASP	2.5
1	A	360	ARG	2.5
1	A	367	HIS	2.5
1	A	369	ILE	2.5
1	A	148	LEU	2.5
1	A	344	LEU	2.4
1	A	73	LEU	2.4
1	A	309	LYS	2.4
1	A	86	GLU	2.3
1	A	380	MET	2.3
1	A	381	THR	2.3
1	A	268	ILE	2.3
1	A	143	LEU	2.3
1	A	273	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	362	CYS	2.2
1	A	89	PRO	2.1
1	A	53	LYS	2.1
1	A	346	ILE	2.1
1	A	325	LEU	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	DOC	P	873	18/19	0.96	0.18	-	30,36,40,41	0
3	BRU	T	840[B]	20/21	0.85	0.19	-	48,54,71,80	20
3	BRU	T	840[A]	20/21	0.85	0.19	-	49,58,73,98	20

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	DGT	T	875	31/31	0.80	0.23	0.98	40,46,50,51	31
4	MG	A	871	1/1	0.66	0.12	-0.77	64,64,64,64	0
4	MG	A	872	1/1	0.89	0.34	-	58,58,58,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.