



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:38 PM GMT

PDB ID : 4M8O  
Title : TERNARY COMPLEX OF DNA POLYMERASE EPSILON WITH AN IN-COMING dATP  
Authors : Sauer-Eriksson, A.E.; Hogg, M.; Osterman, P.; Johansson, E.  
Deposited on : 2013-08-13  
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

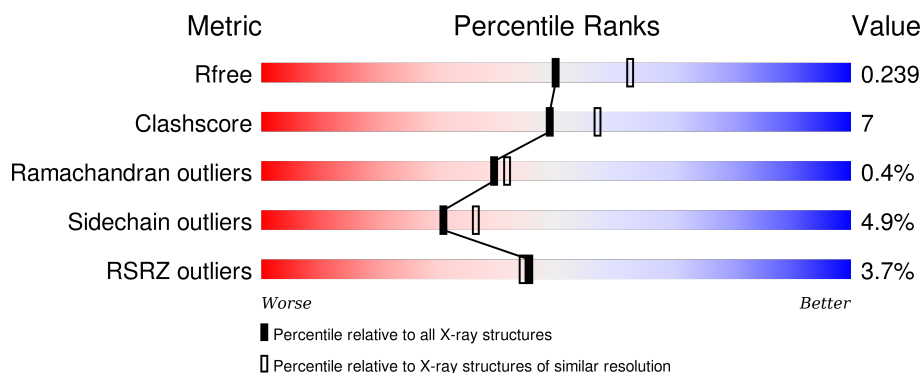
# 1 Overall quality at a glance

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	1228	 3% 78% 12% • 8%
2	P	11	 73% 27%
3	T	16	 69% 25% 6%

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
3	C38	T	3	-	-	X	-
3	5IU	T	4[A]	-	-	X	-

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 10235 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 epsilon catalytic subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1126	Total	C	N	O	S	0	0	0
			9173	5870	1539	1719	45			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	290	ALA	ASP	ENGINEERED MUTATION	UNP P21951
A	292	ALA	GLU	ENGINEERED MUTATION	UNP P21951

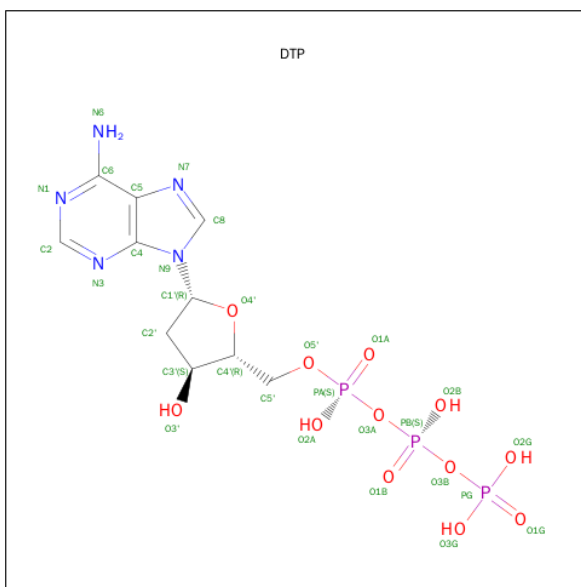
- Molecule 2 is a DNA chain called PRIMER DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	11	Total	C	I	N	O	P	0	0
			224	103	6	38	66	11		

- Molecule 3 is a DNA chain called TEMPLATE DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	T	15	Total	C	I	N	O	P	0	1
			331	152	8	56	99	16		

- Molecule 4 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula:  $C_{10}H_{16}N_5O_{12}P_3$ ).



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

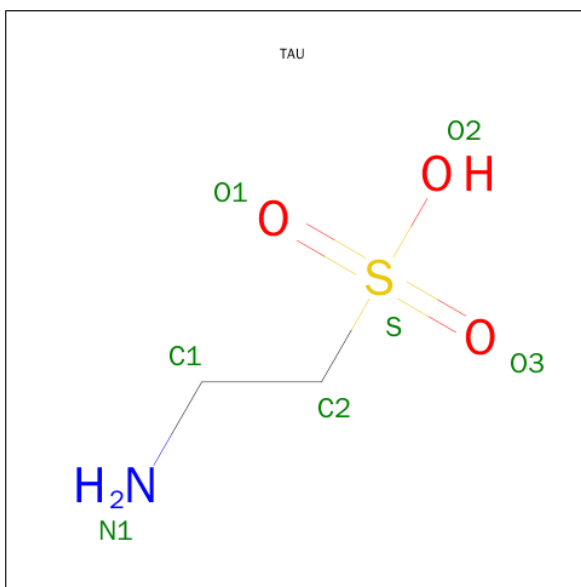
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mg 1 1	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Zn 1 1	0	0

- Molecule 7 is 2-AMINOETHANESULFONIC ACID (three-letter code: TAU) (formula:  $\text{C}_2\text{H}_7\text{NO}_3\text{S}$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	S	0	0
			7	2	1	3	1		
7	A	1	Total	C	N	O	S	0	0
			7	2	1	3	1		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	P	2	Total	Cl	0	0
			2	2		

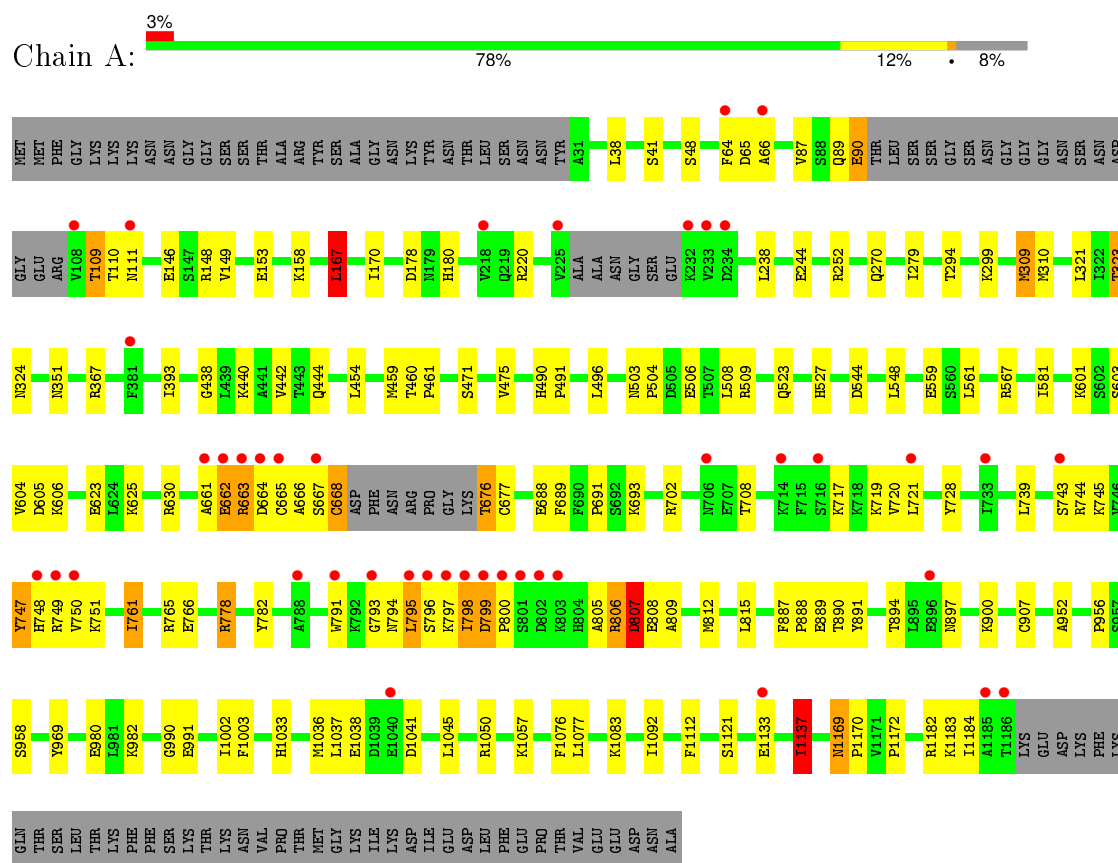
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	407	Total	O	0	0
			407	407		
9	P	22	Total	O	0	0
			22	22		
9	T	30	Total	O	0	0
			30	30		

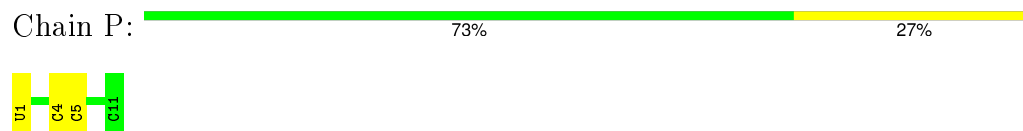
### 3 Residue-property plots

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 epsilon catalytic subunit A

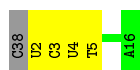


#### • Molecule 2: PRIMER DNA



#### • Molecule 3: TEMPLATE DNA





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.50Å 68.92Å 149.85Å 90.00° 109.50° 90.00°	Depositor
Resolution (Å)	47.05 – 2.20 47.00 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.05-2.20) 99.9 (47.00-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.180 , 0.237 0.187 , 0.239	Depositor DCC
$R_{free}$ test set	3667 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.8	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 45.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 72721 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10235	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: TAU, ZN, CL, DOC, MG, 5IU, C38, DTP

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	0.84	4/9381 (0.0%)	0.87	8/12676 (0.1%)
2	P	0.55	0/95	0.92	0/140
3	T	1.06	1/190 (0.5%)	1.15	1/285 (0.4%)
All	All	0.85	5/9666 (0.1%)	0.88	9/13101 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	506	GLU	CD-OE1	5.74	1.31	1.25
3	T	5	DT	P-OP2	-5.31	1.40	1.49
1	A	506	GLU	CG-CD	5.26	1.59	1.51
1	A	1170	PRO	N-CD	5.22	1.55	1.47
1	A	1172	PRO	N-CD	5.03	1.54	1.47

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	747	TYR	N-CA-C	6.78	129.30	111.00
3	T	5	DT	O5'-P-OP2	-5.89	100.39	105.70
1	A	778	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	309	MET	CA-CB-CG	5.60	122.81	113.30
1	A	109	THR	CB-CA-C	-5.46	96.87	111.60
1	A	367	ARG	NE-CZ-NH1	5.44	123.02	120.30
1	A	509	ARG	NE-CZ-NH1	5.25	122.92	120.30
1	A	1169	ASN	C-N-CD	5.08	139.07	128.40
1	A	167	LEU	CA-CB-CG	5.02	126.85	115.30

There are no chirality outliers.

There are no planarity outliers.

## 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	9173	0	9089	117	0
2	P	224	0	113	3	0
3	T	331	0	165	17	0
4	A	30	0	12	0	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
7	A	14	0	14	1	0
8	P	2	0	0	0	0
9	A	407	0	0	4	0
9	P	22	0	0	1	0
9	T	30	0	0	5	0
All	All	10235	0	9393	137	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 (137) 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)
3:T:4[B]:5IU:O2	9:T:120:HOH:O	1.63	1.16
1:A:178:ASP:OD1	1:A:664:ASP:HB3	1.50	1.07
1:A:110:THR:HG22	1:A:111:ASN:H	1.25	1.01
1:A:64:PHE:CB	1:A:270:GLN:HG2	1.90	1.00
1:A:603:SER:OG	1:A:606:LYS:HE2	1.61	1.00
1:A:663:ARG:HD3	1:A:664:ASP:CG	1.83	0.98
1:A:64:PHE:CD1	1:A:270:GLN:HB3	1.98	0.97
1:A:64:PHE:HB2	1:A:270:GLN:CG	1.95	0.96
1:A:64:PHE:HB2	1:A:270:GLN:HG2	0.99	0.96
1:A:806:ARG:HD3	1:A:809:ALA:HB2	1.44	0.95
1:A:310:MET:CE	1:A:321:LEU:HD11	2.01	0.91
1:A:806:ARG:HG2	1:A:809:ALA:H	1.35	0.89
3:T:4[A]:5IU:H1'	9:T:120:HOH:O	1.71	0.89
3:T:4[A]:5IU:H6	9:T:120:HOH:O	1.74	0.88
1:A:806:ARG:HA	1:A:808:GLU:H	1.35	0.88
1:A:798:ILE:C	1:A:800:PRO:HD2	1.98	0.83

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:797:LYS:CG	1:A:798:ILE:HD12	2.09	0.81
1:A:64:PHE:CD1	1:A:270:GLN:CB	2.63	0.81
1:A:310:MET:HE2	1:A:321:LEU:HD11	1.64	0.79
1:A:797:LYS:HG3	1:A:798:ILE:HD12	1.64	0.78
1:A:310:MET:HE3	1:A:321:LEU:HD11	1.67	0.77
1:A:799:ASP:N	1:A:800:PRO:CD	2.47	0.76
3:T:3:C38:H2'2	3:T:4[B]:5IU:C6	2.17	0.75
1:A:662:GLU:O	1:A:663:ARG:HG2	1.87	0.73
1:A:688:GLU:OE1	1:A:751:LYS:HE3	1.90	0.72
1:A:1002:ILE:HD12	1:A:1003:PHE:N	2.05	0.72
1:A:110:THR:HG22	1:A:111:ASN:N	2.02	0.72
1:A:799:ASP:N	1:A:800:PRO:HD2	2.04	0.71
1:A:796:SER:HA	1:A:800:PRO:HG3	1.73	0.70
1:A:797:LYS:O	1:A:798:ILE:HG13	1.91	0.70
1:A:663:ARG:HD3	1:A:664:ASP:OD1	1.91	0.69
1:A:665:CYS:HB3	1:A:668:CYS:HB2	1.74	0.69
3:T:4[B]:5IU:H1'	9:T:120:HOH:O	1.93	0.68
1:A:791:TRP:CE2	1:A:812:MET:HG3	2.29	0.68
1:A:795:LEU:O	1:A:800:PRO:HB3	1.94	0.67
2:P:5:C38:I	9:P:214:HOH:O	2.81	0.67
1:A:794:ASN:O	1:A:799:ASP:HB2	1.94	0.67
1:A:797:LYS:HG2	1:A:798:ILE:HD12	1.78	0.64
1:A:170:ILE:C	1:A:170:ILE:HD12	2.18	0.64
1:A:310:MET:CE	1:A:321:LEU:HD21	2.28	0.63
1:A:806:ARG:HA	1:A:808:GLU:N	2.11	0.62
1:A:806:ARG:HG2	1:A:809:ALA:N	2.11	0.62
1:A:1112:PHE:CE2	1:A:1137:ILE:HD12	2.34	0.62
1:A:744:ARG:O	1:A:748:HIS:HB3	2.00	0.62
1:A:1112:PHE:CZ	1:A:1137:ILE:HD12	2.35	0.61
1:A:146:GLU:O	1:A:149:VAL:HG23	2.01	0.61
3:T:2:5IU:H3'	3:T:2:5IU:H6	1.85	0.59
1:A:743:SER:O	1:A:747:TYR:HB2	2.03	0.58
1:A:294:THR:CA	1:A:309:MET:HE1	2.34	0.58
3:T:3:C38:H2'1	3:T:4[A]:5IU:H5''	1.86	0.57
1:A:460:THR:HB	1:A:461:PRO:HD3	1.86	0.57
1:A:310:MET:HE1	1:A:321:LEU:HD21	1.86	0.57
1:A:815:LEU:HD11	9:A:1597:HOH:O	2.04	0.57
1:A:294:THR:HA	1:A:309:MET:HE1	1.86	0.56
3:T:3:C38:H2'2	3:T:4[B]:5IU:H6	1.88	0.56
1:A:661:ALA:O	1:A:662:GLU:HB2	2.06	0.55
3:T:3:C38:O1P	3:T:3:C38:I	2.94	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:791:TRP:CD1	1:A:812:MET:HB3	2.42	0.55
1:A:603:SER:HG	1:A:606:LYS:HE2	1.69	0.55
1:A:744:ARG:O	1:A:748:HIS:HA	2.06	0.54
3:T:3:C38:I	3:T:4[B]:5IU:I5	3.66	0.54
1:A:623:GLU:OE1	1:A:630:ARG:NH2	2.40	0.54
1:A:806:ARG:CD	1:A:809:ALA:HB2	2.30	0.53
1:A:805:ALA:O	1:A:807:ASP:HB3	2.09	0.53
3:T:4[A]:5IU:C2'	3:T:4[A]:5IU:O2	2.54	0.53
1:A:523:GLN:O	1:A:527:HIS:HD2	1.92	0.53
1:A:454:LEU:HD23	1:A:459:MET:HG2	1.89	0.53
1:A:666:ALA:N	1:A:667:SER:HA	2.23	0.53
1:A:299:LYS:HE2	1:A:1077:LEU:O	2.09	0.53
1:A:548:LEU:HD23	1:A:689:PHE:HB3	1.92	0.52
1:A:1050:ARG:NH1	9:A:1527:HOH:O	2.43	0.52
1:A:440:LYS:NZ	1:A:444:GLN:OE1	2.42	0.52
1:A:688:GLU:HB3	1:A:751:LYS:HG2	1.92	0.51
1:A:676:THR:OG1	1:A:677:CYS:N	2.43	0.51
1:A:323:THR:HG21	1:A:351:ASN:OD1	2.10	0.51
1:A:969:TYR:CZ	1:A:982:LYS:HG3	2.46	0.51
1:A:310:MET:HE2	1:A:321:LEU:CD1	2.37	0.50
3:T:2:5IU:H6	3:T:2:5IU:C3'	2.41	0.50
1:A:294:THR:CG2	1:A:309:MET:CE	2.91	0.49
1:A:761:ILE:CD1	1:A:761:ILE:N	2.75	0.49
1:A:604:VAL:HG23	1:A:605:ASP:N	2.27	0.49
1:A:980:GLU:HG3	7:A:1305:TAU:H1C2	1.94	0.49
1:A:65:ASP:OD1	1:A:66:ALA:N	2.45	0.49
3:T:3:C38:H2'1	3:T:4[A]:5IU:C5'	2.42	0.48
1:A:64:PHE:HD1	1:A:270:GLN:CB	2.22	0.48
3:T:4[A]:5IU:C6	9:T:120:HOH:O	2.45	0.48
1:A:663:ARG:CD	1:A:664:ASP:CG	2.71	0.48
1:A:310:MET:CE	1:A:321:LEU:CD1	2.84	0.48
1:A:744:ARG:O	1:A:748:HIS:CA	2.62	0.48
1:A:324:ASN:OD1	1:A:324:ASN:C	2.52	0.47
1:A:990:GLY:O	1:A:991:GLU:HB2	2.14	0.47
1:A:1037:LEU:HD21	1:A:1045:LEU:HD12	1.96	0.47
1:A:887:PHE:CG	1:A:888:PRO:HD2	2.50	0.47
2:P:1:5IU:H5'	2:P:1:5IU:H6	1.96	0.46
1:A:805:ALA:C	1:A:807:ASP:HB3	2.36	0.46
1:A:490:HIS:HB3	1:A:491:PRO:HD3	1.98	0.46
1:A:294:THR:CG2	1:A:309:MET:HE2	2.45	0.46
1:A:148:ARG:CZ	1:A:238:LEU:HD11	2.45	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:180:HIS:NE2	1:A:766:GLU:OE2	2.49	0.46
1:A:567:ARG:HG3	1:A:952:ALA:HB2	1.97	0.46
1:A:980:GLU:HG2	1:A:982:LYS:HE2	1.97	0.45
1:A:956:PRO:HD2	9:A:1692:HOH:O	2.16	0.45
1:A:508:LEU:HD12	9:A:1693:HOH:O	2.16	0.45
1:A:677:CYS:SG	1:A:765:ARG:HD3	2.57	0.45
3:T:3:C38:H2'1	3:T:4[B]:5IU:H5''	1.98	0.45
1:A:244:GLU:OE1	1:A:252:ARG:NH2	2.49	0.45
1:A:691:PRO:HD3	1:A:750:VAL:O	2.17	0.45
1:A:294:THR:N	1:A:309:MET:CE	2.80	0.44
1:A:178:ASP:OD1	1:A:664:ASP:CB	2.42	0.44
1:A:294:THR:N	1:A:309:MET:HE1	2.33	0.43
1:A:310:MET:HE1	1:A:471:SER:CB	2.49	0.43
1:A:805:ALA:C	1:A:807:ASP:CB	2.87	0.43
1:A:744:ARG:O	1:A:748:HIS:CB	2.65	0.43
1:A:791:TRP:CZ2	1:A:812:MET:HG3	2.53	0.43
1:A:471:SER:O	1:A:475:VAL:HG13	2.18	0.43
1:A:153:GLU:HG3	1:A:167:LEU:HD21	2.01	0.43
1:A:309:MET:HB3	1:A:309:MET:HE2	1.94	0.42
1:A:665:CYS:CB	1:A:668:CYS:HB2	2.45	0.42
1:A:1182:ARG:O	1:A:1183:LYS:C	2.57	0.42
1:A:559:GLU:OE1	1:A:561:LEU:HD21	2.19	0.42
1:A:663:ARG:CD	1:A:664:ASP:OD1	2.65	0.42
1:A:889:GLU:HB3	1:A:890:THR:H	1.65	0.42
1:A:581:ILE:HG21	1:A:625:LYS:HB2	2.02	0.42
2:P:4:C38:H2'2	2:P:5:C38:I	2.90	0.41
1:A:623:GLU:HG3	1:A:630:ARG:NH2	2.36	0.41
1:A:438:GLY:O	1:A:442:VAL:HG23	2.21	0.41
1:A:89:GLN:O	1:A:90:GLU:HB2	2.20	0.41
1:A:794:ASN:HD22	1:A:797:LYS:HE2	1.85	0.41
1:A:793:GLY:O	1:A:797:LYS:N	2.52	0.41
3:T:4[A]:5IU:C3'	3:T:4[A]:5IU:O2	2.70	0.40
1:A:891:TYR:CD1	1:A:891:TYR:N	2.89	0.40
1:A:294:THR:HG23	1:A:309:MET:HE2	2.03	0.40
1:A:167:LEU:C	1:A:167:LEU:HD23	2.42	0.40
1:A:739:LEU:HD23	1:A:739:LEU:HA	1.91	0.40
1:A:503:ASN:HB2	1:A:504:PRO:CD	2.52	0.40
3:T:4[A]:5IU:H3'	3:T:4[A]:5IU:O2	2.22	0.40
1:A:778:ARG:HD3	1:A:782:TYR:CE2	2.56	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	1118/1228 (91%)	1066 (95%)	47 (4%)	5 (0%)	39 42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	798	ILE
1	A	807	ASP
1	A	799	ASP
1	A	1169	ASN
1	A	1137	ILE

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	1017/1102 (92%)	967 (95%)	50 (5%)	31 36

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	LEU
1	A	41	SER
1	A	48	SER
1	A	87	VAL
1	A	90	GLU
1	A	109	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	158	LYS
1	A	167	LEU
1	A	220	ARG
1	A	279	ILE
1	A	323	THR
1	A	393	ILE
1	A	496	LEU
1	A	544	ASP
1	A	601	LYS
1	A	662	GLU
1	A	663	ARG
1	A	668	CYS
1	A	676	THR
1	A	693	LYS
1	A	702	ARG
1	A	708	THR
1	A	717	LYS
1	A	719	LYS
1	A	720	VAL
1	A	721	LEU
1	A	728	TYR
1	A	745	LYS
1	A	749	ARG
1	A	761	ILE
1	A	795	LEU
1	A	806	ARG
1	A	807	ASP
1	A	894	THR
1	A	897	ASN
1	A	900	LYS
1	A	907	CYS
1	A	958	SER
1	A	1033	HIS
1	A	1036	MET
1	A	1038	GLU
1	A	1041	ASP
1	A	1057	LYS
1	A	1076	PHE
1	A	1083	LYS
1	A	1092	ILE
1	A	1121	SER
1	A	1133	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1137	ILE
1	A	1184	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	ASN
1	A	83	HIS
1	A	224	ASN
1	A	360	GLN
1	A	434	GLN
1	A	523	GLN
1	A	527	HIS
1	A	713	ASN
1	A	748	HIS
1	A	794	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

15 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	5IU	P	1	3,2	12,21,22	0.66	0	14,30,33	1.84	1 (7%)
2	5IU	P	10	3,2	12,21,22	1.14	1 (8%)	14,30,33	2.66	1 (7%)
2	DOC	P	11	3,2	11,19,20	1.02	1 (9%)	14,26,29	1.92	4 (28%)
2	C38	P	4	3,2	12,21,22	0.70	0	16,30,33	1.39	2 (12%)
2	C38	P	5	3,2	12,21,22	0.88	0	16,30,33	1.43	4 (25%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	C38	P	7	3,2	12,21,22	0.85	0	16,30,33	1.55	5 (31%)
2	5IU	P	9	3,2	12,21,22	0.88	0	14,30,33	2.59	2 (14%)
3	C38	T	11	3,2	12,21,22	0.95	0	16,30,33	1.25	3 (18%)
3	5IU	T	14	3,2	12,21,22	1.06	1 (8%)	14,30,33	2.51	2 (14%)
3	5IU	T	15	3,2	12,21,22	0.63	0	14,30,33	1.94	1 (7%)
3	5IU	T	2	3	12,21,22	0.65	0	14,30,33	2.38	3 (21%)
3	C38	T	3	3	12,21,22	0.68	0	16,30,33	1.45	4 (25%)
3	5IU	T	4[A]	3	12,21,22	1.21	1 (8%)	14,30,33	2.76	6 (42%)
3	5IU	T	4[B]	3	12,21,22	0.71	0	14,30,33	2.21	4 (28%)
3	C38	T	9	3,2	12,21,22	0.65	0	16,30,33	1.43	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	5IU	P	1	3,2	-	0/3/21/22	0/2/2/2
2	5IU	P	10	3,2	-	0/3/21/22	0/2/2/2
2	DOC	P	11	3,2	-	0/3/18/19	0/2/2/2
2	C38	P	4	3,2	-	0/3/21/22	0/2/2/2
2	C38	P	5	3,2	-	0/3/21/22	0/2/2/2
2	C38	P	7	3,2	-	0/3/21/22	0/2/2/2
2	5IU	P	9	3,2	-	0/3/21/22	0/2/2/2
3	C38	T	11	3,2	-	0/3/21/22	0/2/2/2
3	5IU	T	14	3,2	-	0/3/21/22	0/2/2/2
3	5IU	T	15	3,2	-	0/3/21/22	0/2/2/2
3	5IU	T	2	3	-	0/3/21/22	0/2/2/2
3	C38	T	3	3	-	0/3/21/22	0/2/2/2
3	5IU	T	4[A]	3	-	0/3/21/22	0/2/2/2
3	5IU	T	4[B]	3	-	0/3/21/22	0/2/2/2
3	C38	T	9	3,2	-	0/3/21/22	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	11	DOC	C4-N3	-2.02	1.31	1.35
3	T	14	5IU	O4'-C1'	2.10	1.47	1.42
3	T	4[A]	5IU	C6-N1	2.44	1.38	1.35
2	P	10	5IU	C6-N1	2.75	1.39	1.35

All (44) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	9	C38	O4'-C1'-C2'	-3.52	99.25	106.27
3	T	4[B]	5IU	C2'-C1'-N1	-3.21	106.34	114.16
3	T	9	C38	O4'-C1'-N1	-3.16	102.25	107.72
2	P	5	C38	C5-C4-N4	-3.01	120.79	122.67
2	P	7	C38	O4'-C1'-C2'	-2.73	100.83	106.27
3	T	3	C38	C6-C5-I	-2.57	112.57	118.99
3	T	2	5IU	C2'-C1'-N1	-2.40	108.33	114.16
2	P	11	DOC	C6-N1-C2	-2.32	117.53	121.28
3	T	11	C38	O3'-C3'-C4'	-2.13	101.45	110.05
3	T	3	C38	O4'-C1'-N1	2.01	111.19	107.72
3	T	3	C38	O3'-C3'-C2'	2.05	117.52	110.74
3	T	4[A]	5IU	O4'-C4'-C5'	2.06	116.67	109.32
2	P	11	DOC	O4'-C4'-C5'	2.14	112.70	109.54
3	T	2	5IU	O4'-C1'-N1	2.20	111.53	107.72
2	P	7	C38	O3'-C3'-C2'	2.21	118.04	110.74
2	P	4	C38	O4'-C1'-C2'	2.22	110.70	106.27
3	T	11	C38	O3'-C3'-C2'	2.23	118.13	110.74
2	P	7	C38	C2'-C1'-N1	2.25	119.64	114.16
2	P	7	C38	C6-C5-C4	2.33	119.64	115.74
2	P	5	C38	N4-C4-N3	2.34	120.34	116.95
3	T	4[A]	5IU	C3'-C2'-C1'	2.44	108.27	102.40
2	P	7	C38	N4-C4-N3	2.45	120.50	116.95
2	P	5	C38	C6-C5-C4	2.45	119.84	115.74
3	T	3	C38	C3'-C2'-C1'	2.67	108.82	102.40
2	P	5	C38	C2'-C3'-C4'	2.74	108.46	102.77
3	T	4[B]	5IU	O4'-C1'-N1	2.76	112.49	107.72
3	T	11	C38	C6-C5-C4	2.77	120.37	115.74
3	T	4[B]	5IU	C6-C5-I5	2.85	126.11	118.99
2	P	4	C38	O4'-C1'-N1	2.90	112.73	107.72
3	T	4[A]	5IU	O3'-C3'-C2'	2.93	120.45	110.74
3	T	14	5IU	O4'-C1'-N1	3.16	113.18	107.72
2	P	9	5IU	C2'-C1'-N1	3.18	121.88	114.16
2	P	11	DOC	C2-N3-C4	3.33	120.31	115.61
2	P	11	DOC	C3'-C2'-C1'	3.54	106.66	102.71
3	T	4[A]	5IU	C6-C5-I5	3.83	128.54	118.99
3	T	4[A]	5IU	O4'-C1'-N1	4.39	115.32	107.72
3	T	15	5IU	C4-N3-C2	6.02	120.45	115.25
3	T	4[B]	5IU	C4-N3-C2	6.22	120.62	115.25
2	P	1	5IU	C4-N3-C2	6.35	120.73	115.25
3	T	4[A]	5IU	C4-N3-C2	6.84	121.16	115.25
3	T	2	5IU	C4-N3-C2	7.88	122.06	115.25
3	T	14	5IU	C4-N3-C2	8.11	122.26	115.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	9	5IU	C4-N3-C2	8.76	122.81	115.25
2	P	10	5IU	C4-N3-C2	9.51	123.47	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	1	5IU	1	0
2	P	4	C38	1	0
2	P	5	C38	2	0
3	T	2	5IU	2	0
3	T	3	C38	7	0
3	T	4[A]	5IU	8	0
3	T	4[B]	5IU	6	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	$\# Z  > 2$
4	DTP	A	1301	5	24,32,32	1.38	4 (16%)	32,50,50	1.80	4 (12%)
7	TAU	A	1304	-	5,6,6	3.03	1 (20%)	6,8,8	3.51	3 (50%)
7	TAU	A	1305	-	5,6,6	2.59	1 (20%)	6,8,8	2.58	2 (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
4	DTP	A	1301	5	-	0/18/34/34	0/3/3/3
7	TAU	A	1304	-	-	0/4/4/4	0/0/0/0
7	TAU	A	1305	-	-	0/4/4/4	0/0/0/0

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1301	DTP	PB-O2B	-2.72	1.43	1.54
4	A	1301	DTP	PA-O2A	-2.41	1.44	1.54
4	A	1301	DTP	C8-N7	2.01	1.38	1.34
4	A	1301	DTP	C5-C4	2.88	1.47	1.40
7	A	1305	TAU	O2-S	5.68	1.61	1.46
7	A	1304	TAU	O2-S	6.69	1.63	1.46

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1301	DTP	N3-C2-N1	-6.29	124.08	128.89
4	A	1301	DTP	C2'-C1'-N9	-4.36	103.56	114.16
7	A	1305	TAU	O2-S-O3	-2.91	104.83	111.61
7	A	1304	TAU	O3-S-O1	-2.52	104.31	113.48
4	A	1301	DTP	O3G-PG-O1G	2.85	119.76	110.58
7	A	1304	TAU	O1-S-C2	3.29	109.71	106.91
4	A	1301	DTP	O4'-C1'-N9	4.14	114.89	107.72
7	A	1305	TAU	O1-S-C2	5.22	111.36	106.91
7	A	1304	TAU	O3-S-C2	7.51	113.31	106.91

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1305	TAU	1	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	A	1126/1228 (91%)	-0.01	42 (3%) 45 44	15, 33, 73, 146	0
2	P	4/11 (36%)	-0.44	0 100 100	20, 31, 46, 48	0
3	T	8/16 (50%)	-0.45	0 100 100	16, 19, 34, 41	0
All	All	1138/1255 (90%)	-0.01	42 (3%) 45 44	15, 33, 73, 146	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	663	ARG	12.2
1	A	664	ASP	10.7
1	A	798	ILE	7.6
1	A	800	PRO	6.9
1	A	1186	THR	6.1
1	A	661	ALA	5.8
1	A	225	VAL	5.7
1	A	665	CYS	5.6
1	A	799	ASP	5.2
1	A	802	ASP	4.5
1	A	64	PHE	4.2
1	A	233	VAL	3.9
1	A	108	VAL	3.9
1	A	743	SER	3.7
1	A	1185	ALA	3.7
1	A	801	SER	3.4
1	A	796	SER	3.3
1	A	662	GLU	3.2
1	A	803	LYS	3.2
1	A	66	ALA	3.1
1	A	218	VAL	3.0
1	A	795	LEU	2.9
1	A	716	SER	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	748	HIS	2.6
1	A	750	VAL	2.6
1	A	714	LYS	2.6
1	A	232	LYS	2.6
1	A	896	GLU	2.5
1	A	793	GLY	2.5
1	A	667	SER	2.4
1	A	749	ARG	2.4
1	A	381	PHE	2.4
1	A	791	TRP	2.3
1	A	788	ALA	2.3
1	A	1040	GLU	2.3
1	A	111	ASN	2.2
1	A	706	ASN	2.2
1	A	234	ASP	2.2
1	A	1133	GLU	2.1
1	A	721	LEU	2.0
1	A	733	ILE	2.0
1	A	797	LYS	2.0

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

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
3	5IU	T	2	20/21	0.60	0.36	-	98,115,123,126	1
3	C38	T	11	20/21	0.98	0.12	-	22,26,30,32	1
2	C38	P	4	20/21	0.94	0.10	-	29,47,65,72	1
2	5IU	P	9	20/21	0.99	0.15	-	16,18,23,23	1
3	5IU	T	15	20/21	0.96	0.10	-	32,39,43,52	1
2	C38	P	5	20/21	0.95	0.10	-	25,44,56,66	1
2	DOC	P	11	18/19	0.98	0.16	-	17,19,24,24	0
3	5IU	T	4[B]	20/21	0.92	0.17	-	26,37,56,73	20
2	5IU	P	10	20/21	0.98	0.16	-	13,18,20,25	1
2	C38	P	7	20/21	0.98	0.12	-	19,24,29,35	1
3	5IU	T	14	20/21	0.98	0.11	-	30,35,38,39	1
3	C38	T	9	20/21	0.98	0.11	-	16,21,25,28	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	5IU	T	4[A]	20/21	0.92	0.17	-	26,37,58,87	20
2	5IU	P	1	20/21	0.91	0.19	-	43,56,140,144	1
3	C38	T	3	20/21	0.81	0.14	-	37,68,94,110	1

### 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( $\text{\AA}^2$ )	Q<0.9
4	DTP	A	1301	30/30	0.99	0.13	-0.81	13,14,23,24	0
6	ZN	A	1303	1/1	0.99	0.07	-1.16	41,41,41,41	0
5	MG	A	1302	1/1	0.96	0.10	-1.87	21,21,21,21	0
8	CL	P	101	1/1	1.00	0.25	-	9,9,9,9	0
7	TAU	A	1305	7/7	0.87	0.30	-	50,59,67,68	7
7	TAU	A	1304	7/7	0.84	0.28	-	66,79,82,88	0
8	CL	P	102	1/1	0.99	0.29	-	24,24,24,24	1

### 6.5 Other polymers [i](#)

There are no such residues in this entry.