



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 08:56 PM GMT

PDB ID : 3V6H
Title : Replication of N2,3-Ethenoguanine by DNA Polymerases
Authors : Zhao, L.
Deposited on : 2011-12-19
Resolution : 2.30 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

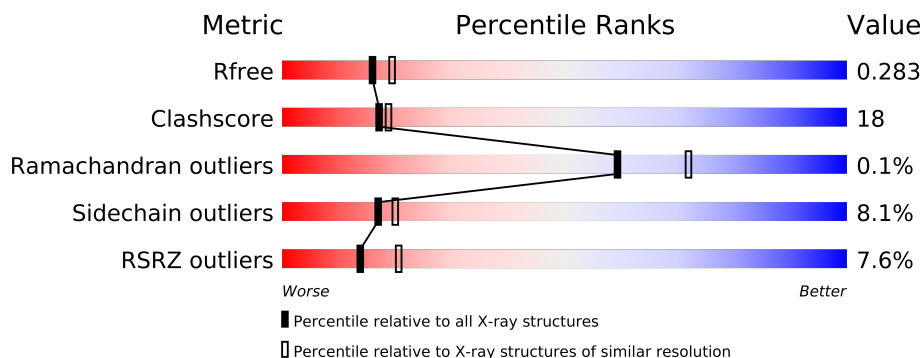
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.30 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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	C	18	
1	T	18	
2	D	13	
2	P	13	
3	A	348	
3	B	348	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	CA	A	404	-	X
6	MG	A	403	-	X
6	MG	B	402	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6867 atoms, of which 0 are hydrogen and 0 are deuterium.

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*AP*CP*(EFG)P*GP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*C)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	T	13	Total	C	F	N	O	P	0	0	0
			263	126	1	44	79	13			
1	C	17	Total	C	F	N	O	P	0	0	0
			339	164	1	57	101	16			

- Molecule 2 is a DNA chain called DNA (5'-D(*GP*G*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*(DOC))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	11	Total	C	N	O	P	0	0	0
			228	109	47	62	10			
2	D	12	Total	C	N	O	P	0	0	0
			250	119	52	68	11			

- Molecule 3 is a protein called DNA polymerase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	343	Total	C	N	O	S	0	0	0
			2762	1771	476	508	7			
3	B	342	Total	C	N	O	S	0	0	0
			2752	1765	473	507	7			

There are 12 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

Continued on next page...

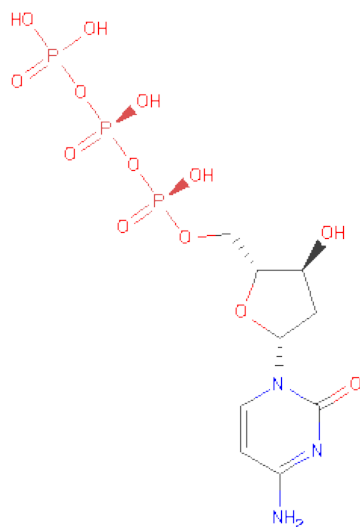
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	HIS	-	EXPRESSION TAG	UNP Q97W02
B	-4	HIS	-	EXPRESSION TAG	UNP Q97W02
B	-3	HIS	-	EXPRESSION TAG	UNP Q97W02
B	-2	HIS	-	EXPRESSION TAG	UNP Q97W02
B	-1	HIS	-	EXPRESSION TAG	UNP Q97W02
B	0	HIS	-	EXPRESSION TAG	UNP Q97W02

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	A	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0

- Molecule 5 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: C₉H₁₆N₃O₁₃P₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O P 28 9 3 13 3	0	0
5	B	1	Total C N O P 28 9 3 13 3	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total 2	Mg 2	0	0
6	A	2	Total 2	Mg 2	0	0

- Molecule 7 is water.

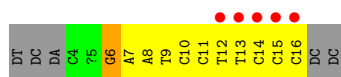
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	T	8	Total 8	O 8	0	0
7	P	4	Total 4	O 4	0	0
7	C	15	Total 15	O 15	0	0
7	D	7	Total 7	O 7	0	0
7	A	115	Total 115	O 115	0	0
7	B	61	Total 61	O 61	0	0

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.

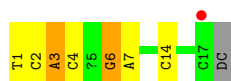
- Molecule 1: DNA (5'-D(*TP*CP*AP*CP*(EFG)P*GP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*C)-3')

Chain T: 



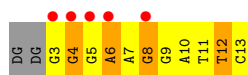
- Molecule 1: DNA (5'-D(*TP*CP*AP*CP*(EFG)P*GP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*C)-3')

Chain C: 



- Molecule 2: DNA (5'-D(*GP*G*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*(DOC))-3')

Chain P: 



- Molecule 2: DNA (5'-D(*GP*G*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*(DOC))-3')

Chain D: 



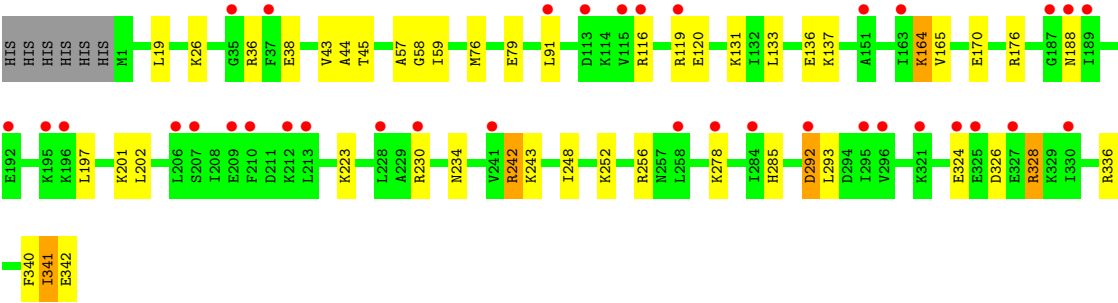
- Molecule 3: DNA polymerase IV

Chain A: 



- Molecule 3: DNA polymerase IV

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	52.53Å 111.27Å 98.94Å 90.00° 102.68° 90.00°	Depositor
Resolution (Å)	27.80 – 2.30 29.80 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.0 (27.80-2.30) 98.1 (29.80-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.72 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.215 , 0.267 0.231 , 0.283	Depositor DCC
R_{free} test set	2446 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	51.5	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.0	EDS
Estimated twinning fraction	0.030 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 48486 reflections (0.004%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6867	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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: DOC, MG, CA, DCP, EFG

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	C	0.75	2/348 (0.6%)	1.09	3/530 (0.6%)
1	T	0.36	0/263	0.98	2/399 (0.5%)
2	D	0.55	1/262 (0.4%)	1.05	3/405 (0.7%)
2	P	1.23	3/237 (1.3%)	1.28	5/366 (1.4%)
3	A	0.42	0/2802	0.60	0/3763
3	B	0.36	0/2791	0.57	0/3748
All	All	0.48	6/6703 (0.1%)	0.71	13/9211 (0.1%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	8	DG	O3'-P	-12.50	1.46	1.61
2	P	4	DG	O3'-P	9.41	1.72	1.61
1	C	1	DT	O3'-P	-6.91	1.52	1.61
1	C	3	DA	O3'-P	-6.88	1.52	1.61
2	D	11	DT	O3'-P	6.27	1.68	1.61
2	P	6	DA	O3'-P	-5.03	1.55	1.61

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	14	DC	P-O3'-C3'	7.40	128.58	119.70
2	D	11	DT	P-O3'-C3'	7.25	128.40	119.70
1	C	6	DG	C1'-O4'-C4'	-6.91	103.19	110.10
2	D	3	DG	P-O3'-C3'	6.20	127.14	119.70
2	P	12	DT	O5'-P-OP1	-5.90	100.39	105.70
1	T	10	DC	P-O3'-C3'	5.84	126.71	119.70
2	P	4	DG	P-O3'-C3'	-5.79	112.75	119.70
2	P	10	DA	P-O3'-C3'	5.58	126.40	119.70
2	P	9	DG	OP1-P-O3'	5.55	117.41	105.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	12	DT	O5'-P-OP1	-5.38	100.85	105.70
2	P	4	DG	O3'-P-O5'	5.37	114.19	104.00
1	C	1	DT	P-O3'-C3'	-5.33	113.30	119.70
1	T	6	DG	P-O3'-C3'	5.28	126.03	119.70

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	339	0	0	6	0
1	T	263	0	0	29	0
2	D	250	0	0	11	0
2	P	228	0	0	42	0
3	A	2762	0	9	20	0
3	B	2752	0	0	17	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	A	28	0	12	0	0
5	B	28	0	12	2	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	A	115	0	0	9	0
7	B	61	0	0	8	0
7	C	15	0	0	0	0
7	D	7	0	0	1	0
7	P	4	0	0	0	0
7	T	8	0	0	0	0
All	All	6867	0	33	112	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 18.

All (112) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:T:11:DC:C2'	1:T:12:DT:C6	2.30	1.14
2:P:7:DA:C4	2:P:8:DG:C8	2.35	1.14
1:T:11:DC:C3'	1:T:12:DT:C7	2.32	1.06
2:P:12:DT:C2'	2:P:13:DOC:C5	2.37	1.02
2:P:12:DT:C2'	2:P:13:DOC:C6	2.42	0.97
2:P:7:DA:C2	2:P:8:DG:C4	2.53	0.97
1:T:11:DC:C6	1:T:11:DC:C5'	2.48	0.96
2:P:7:DA:N3	2:P:8:DG:C8	2.35	0.94
2:P:5:DG:C5'	2:P:5:DG:C8	2.50	0.94
2:P:7:DA:C5	2:P:8:DG:N7	2.36	0.93
2:P:7:DA:C6	2:P:8:DG:C5	2.56	0.92
2:P:7:DA:N1	2:P:8:DG:C5	2.39	0.89
3:B:164:LYS:CE	3:B:165:VAL:O	2.21	0.89
1:T:7:DA:C2	2:P:12:DT:N3	2.43	0.86
1:T:12:DT:O2	1:T:13:DT:C2	2.30	0.84
3:A:36:ARG:NH2	3:A:254:ASN:ND2	2.26	0.84
2:P:7:DA:C2	2:P:8:DG:N9	2.47	0.82
1:T:11:DC:C2'	1:T:12:DT:C5	2.62	0.81
1:C:2:DC:O2	1:C:2:DC:C2'	2.30	0.79
1:T:15:DC:C1'	1:T:16:DC:O5'	2.30	0.79
1:T:14:DC:C2'	1:T:15:DC:O5'	2.31	0.79
1:T:11:DC:C2'	1:T:12:DT:OP1	2.30	0.79
2:D:2:DG:C2'	2:D:3:DG:C2'	2.61	0.79
2:P:7:DA:C6	2:P:8:DG:N7	2.51	0.79
3:B:201:LYS:CE	7:B:501:HOH:O	2.31	0.79
1:C:7:DA:C2	2:D:12:DT:N3	2.52	0.77
2:P:7:DA:C2	2:P:8:DG:C8	2.72	0.76
2:P:11:DT:C6	2:P:12:DT:C7	2.69	0.76
3:A:31:CYS:SG	7:A:581:HOH:O	2.43	0.75
3:B:341:ILE:CA	7:B:505:HOH:O	2.32	0.75
2:P:3:DG:C2'	2:P:4:DG:O5'	2.35	0.75
3:B:292:ASP:OD2	3:B:328:ARG:NH1	2.20	0.74
2:P:13:DOC:C2'	5:B:401:DCP:H1'	2.19	0.73
2:P:7:DA:C4	2:P:8:DG:N7	2.57	0.72
1:T:11:DC:C2'	1:T:12:DT:C7	2.69	0.71
2:P:7:DA:C2	2:P:8:DG:C5	2.79	0.69
2:P:4:DG:N7	2:P:5:DG:C2	2.62	0.68
1:T:14:DC:C1'	1:T:15:DC:C5'	2.71	0.68
3:A:67:ILE:CG1	7:A:609:HOH:O	2.40	0.68
3:A:342:GLU:C	7:A:538:HOH:O	2.32	0.67
1:T:13:DT:C6	1:T:14:DC:N4	2.62	0.67
2:D:6:DA:C2'	2:D:7:DA:C8	2.77	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:59:ILE:N	7:B:561:HOH:O	2.29	0.65
3:A:0:HIS:ND1	3:A:0:HIS:O	2.30	0.65
3:B:292:ASP:OD1	3:B:292:ASP:N	2.30	0.65
1:T:15:DC:C2'	1:T:16:DC:OP2	2.46	0.64
7:D:204:HOH:O	3:A:285:HIS:CE1	2.51	0.64
3:A:278:LYS:N	7:A:568:HOH:O	2.31	0.63
2:P:7:DA:C6	2:P:8:DG:C6	2.87	0.62
1:T:6:DG:N7	2:P:13:DOC:O2	2.33	0.62
3:A:319:LEU:CD1	7:A:580:HOH:O	2.47	0.62
2:P:7:DA:N1	2:P:8:DG:C4	2.66	0.61
2:P:4:DG:C8	2:P:5:DG:C6	2.89	0.61
2:D:13:DOC:C5'	7:A:614:HOH:O	2.48	0.60
1:T:11:DC:O2	2:P:8:DG:N2	2.34	0.60
3:A:2:ILE:N	3:A:2:ILE:CD1	2.65	0.60
1:T:14:DC:N4	2:P:5:DG:N2	2.53	0.57
1:T:9:DT:OP1	3:B:243:LYS:N	2.38	0.57
3:A:341:ILE:O	3:A:342:GLU:C	2.43	0.56
2:P:4:DG:C2'	2:P:5:DG:O5'	2.53	0.56
2:P:12:DT:C2'	2:P:13:DOC:C5'	2.84	0.56
1:C:7:DA:C2	2:D:12:DT:C2	2.95	0.55
2:P:7:DA:N3	2:P:8:DG:N9	2.50	0.55
2:D:2:DG:C4	2:D:3:DG:C8	2.95	0.54
1:T:14:DC:C1'	1:T:15:DC:P	2.95	0.54
2:D:2:DG:C4	2:D:3:DG:N7	2.76	0.54
3:B:58:GLY:N	7:B:561:HOH:O	2.40	0.54
3:B:44:ALA:O	5:B:401:DCP:H2'1	2.09	0.53
2:D:6:DA:C4'	2:D:7:DA:OP1	2.56	0.53
1:T:14:DC:C2'	1:T:15:DC:C5'	2.87	0.53
1:T:7:DA:C2	2:P:12:DT:C2	2.97	0.53
2:D:2:DG:O3'	2:D:3:DG:C3'	2.56	0.53
1:T:13:DT:O2	2:P:7:DA:C2	2.62	0.52
2:P:4:DG:C8	2:P:5:DG:C5	2.98	0.52
3:B:340:PHE:N	3:B:340:PHE:CD1	2.77	0.52
1:T:8:DA:C2	2:P:11:DT:N3	2.78	0.51
3:B:43:VAL:CB	7:B:561:HOH:O	2.58	0.51
2:P:11:DT:C2'	2:P:12:DT:C6	2.94	0.51
2:P:7:DA:C2'	2:P:8:DG:O5'	2.59	0.51
3:B:57:ALA:N	7:B:546:HOH:O	2.43	0.50
1:T:14:DC:C1'	1:T:15:DC:OP1	2.59	0.50
1:C:3:DA:N6	1:C:4:DC:N4	2.60	0.50
1:T:13:DT:C5	1:T:14:DC:N4	2.81	0.49
3:A:242:ARG:NH1	7:A:597:HOH:O	2.46	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:T:11:DC:C6	1:T:11:DC:C4'	2.95	0.49
1:T:12:DT:OP1	1:T:12:DT:O4'	2.30	0.49
1:T:15:DC:C2'	1:T:16:DC:O5'	2.60	0.49
1:C:7:DA:C2	2:D:12:DT:O2	2.66	0.48
3:B:136:GLU:O	3:B:137:LYS:CB	2.62	0.48
2:P:7:DA:N6	2:P:8:DG:C6	2.81	0.48
3:A:79:GLU:N	3:A:79:GLU:CD	2.68	0.47
3:B:242:ARG:NH1	3:B:242:ARG:CG	2.77	0.47
3:A:1:MET:SD	3:A:234:ASN:ND2	2.88	0.47
2:P:3:DG:N2	2:P:4:DG:C2	2.82	0.47
1:T:15:DC:C1'	1:T:16:DC:C5'	2.93	0.47
2:P:7:DA:N1	2:P:8:DG:C6	2.82	0.46
3:A:52:LYS:CE	7:A:561:HOH:O	2.62	0.46
3:A:1:MET:C	3:A:2:ILE:CD1	2.84	0.46
2:P:4:DG:N7	2:P:5:DG:N1	2.64	0.46
3:A:233:TYR:CD1	3:A:233:TYR:C	2.89	0.46
3:B:341:ILE:C	7:B:505:HOH:O	2.54	0.45
1:T:14:DC:C4	2:P:5:DG:N2	2.85	0.44
3:B:285:HIS:CD2	7:B:523:HOH:O	2.71	0.43
2:P:7:DA:N6	2:P:8:DG:O6	2.53	0.42
2:P:5:DG:C2'	2:P:6:DA:O4'	2.68	0.42
3:B:326:ASP:OD1	3:B:328:ARG:CG	2.67	0.42
3:A:173:ARG:NH1	3:A:177:GLU:OE1	2.52	0.42
2:P:13:DOC:C6	2:P:13:DOC:C5'	2.98	0.42
2:D:5:DG:C2'	2:D:6:DA:C8	3.02	0.42
3:A:270:GLU:OE2	3:A:312:TYR:OH	2.39	0.41
1:C:6:DG:N2	3:A:332:ARG:NH1	2.69	0.40
3:A:285:HIS:CD2	7:A:550:HOH:O	2.74	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	
3	A	341/348 (98%)	329 (96%)	12 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	B	340/348 (98%)	326 (96%)	13 (4%)	1 (0%)	50	60
All	All	681/696 (98%)	655 (96%)	25 (4%)	1 (0%)	59	72

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	341	ILE

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	
3	A	302/307 (98%)	286 (95%)	16 (5%)	32	41
3	B	301/307 (98%)	268 (89%)	33 (11%)	9	10
All	All	603/614 (98%)	554 (92%)	49 (8%)	17	20

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	0	HIS
3	A	2	ILE
3	A	19	LEU
3	A	56	LYS
3	A	62	VAL
3	A	97	GLU
3	A	105	ASP
3	A	116	ARG
3	A	192	GLU
3	A	195	LYS
3	A	242	ARG
3	A	253	ARG
3	A	267	ARG
3	A	323	LEU
3	A	336	ARG
3	A	342	GLU
3	B	19	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	26	LYS
3	B	36	ARG
3	B	38	GLU
3	B	45	THR
3	B	76	MET
3	B	79	GLU
3	B	91	LEU
3	B	116	ARG
3	B	119	ARG
3	B	120	GLU
3	B	131	LYS
3	B	133	LEU
3	B	164	LYS
3	B	170	GLU
3	B	176	ARG
3	B	188	ASN
3	B	197	LEU
3	B	202	LEU
3	B	223	LYS
3	B	230	ARG
3	B	234	ASN
3	B	242	ARG
3	B	248	ILE
3	B	252	LYS
3	B	256	ARG
3	B	278	LYS
3	B	292	ASP
3	B	293	LEU
3	B	324	GLU
3	B	328	ARG
3	B	336	ARG
3	B	342	GLU

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

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

4 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	EFG	C	5	1	26,28,29	2.17	9 (34%)	32,42,45	4.60	12 (37%)
2	DOC	D	13	1,2	17,19,20	1.17	1 (5%)	20,26,29	2.40	5 (25%)
2	DOC	P	13	2	17,19,20	0.94	1 (5%)	20,26,29	3.14	6 (30%)
1	EFG	T	5	1	26,28,29	2.15	10 (38%)	32,42,45	4.49	7 (21%)

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	EFG	C	5	1	-	0/8/25/26	0/1/4/4
2	DOC	D	13	1,2	-	0/5/18/19	0/2/2/2
2	DOC	P	13	2	-	0/5/18/19	0/2/2/2
1	EFG	T	5	1	-	0/8/25/26	0/1/4/4

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	5	EFG	C4-N3	-5.10	1.34	1.39
1	C	5	EFG	C2-N3	-4.89	1.33	1.39
1	T	5	EFG	C4-N3	-4.81	1.34	1.39
1	T	5	EFG	C2-N3	-4.39	1.33	1.39
1	C	5	EFG	O4'-C1'	3.73	1.47	1.41
1	T	5	EFG	O4'-C1'	3.69	1.47	1.41
1	T	5	EFG	C6-C5	-3.68	1.35	1.41
1	C	5	EFG	C6-C5	-3.43	1.35	1.41
1	T	5	EFG	C9-N3	-2.79	1.33	1.37
2	D	13	DOC	P-OP1	2.72	1.49	1.46
1	C	5	EFG	C9-N3	-2.69	1.33	1.37

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	5	EFG	C4-N9	-2.63	1.33	1.38
1	C	5	EFG	C5-C4	-2.55	1.34	1.40
1	T	5	EFG	C5-C4	-2.52	1.34	1.40
1	C	5	EFG	C2-N2	-2.45	1.33	1.36
1	T	5	EFG	C2-N2	-2.21	1.33	1.36
2	P	13	DOC	C6-N1	-2.13	1.32	1.35
1	C	5	EFG	C8-N9	-2.13	1.33	1.36
1	T	5	EFG	C4-N9	-2.08	1.34	1.38
1	T	5	EFG	P-OP1	2.02	1.49	1.46
1	T	5	EFG	C6-N1	2.02	1.40	1.37

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	5	EFG	C6-C5-N7	-21.04	131.31	134.14
1	C	5	EFG	C6-C5-N7	-19.30	131.54	134.14
1	C	5	EFG	C10-N2-C2	13.76	108.13	104.92
1	T	5	EFG	C10-N2-C2	11.18	107.53	104.92
2	P	13	DOC	C4'-O4'-C1'	10.62	114.47	110.05
2	D	13	DOC	C4'-O4'-C1'	6.01	112.55	110.05
2	P	13	DOC	C6-C5-C4	5.27	119.66	117.47
2	D	13	DOC	C6-C5-C4	5.15	119.61	117.47
1	C	5	EFG	O4'-C1'-N9	-4.63	104.13	108.44
1	T	5	EFG	N2-C2-N3	-4.17	108.82	111.58
2	D	13	DOC	C2'-C1'-N1	4.16	120.97	112.66
1	C	5	EFG	N2-C2-N3	-4.11	108.86	111.58
1	T	5	EFG	C6-N1-C2	-4.01	117.44	120.28
2	P	13	DOC	C2-N3-C4	3.94	121.28	115.57
2	P	13	DOC	C3'-C2'-C1'	3.05	106.18	102.80
1	C	5	EFG	C6-N1-C2	-2.91	118.22	120.28
1	T	5	EFG	O4'-C1'-N9	2.88	111.12	108.44
2	D	13	DOC	O4'-C4'-C5'	2.83	113.62	109.67
1	C	5	EFG	C5-C4-N3	-2.70	118.28	124.96
1	C	5	EFG	C3'-C2'-C1'	2.58	106.62	103.30
1	C	5	EFG	C2'-C3'-C4'	2.54	106.02	102.10
1	T	5	EFG	C5-C4-N3	-2.53	118.69	124.96
1	C	5	EFG	F-C2'-C3'	-2.53	102.56	109.39
1	C	5	EFG	C4-C5-N7	-2.51	107.38	109.52
2	P	13	DOC	C2-N1-C1'	2.40	122.05	119.25
2	D	13	DOC	C2-N3-C4	2.38	119.02	115.57
1	C	5	EFG	C2'-C1'-N9	2.30	117.80	113.79
1	C	5	EFG	F-C2'-C1'	2.23	114.89	109.61

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	5	EFG	C4-C5-N7	-2.14	107.69	109.52
2	P	13	DOC	N4-C4-N3	2.04	120.51	116.59

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 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	DCP	A	401	6	29,29,29	1.12	1 (3%)	42,45,45	1.87	6 (14%)
5	DCP	B	401	6	29,29,29	1.12	1 (3%)	42,45,45	1.71	7 (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.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DCP	A	401	6	-	0/19/34/34	0/2/2/2
5	DCP	B	401	6	-	0/19/34/34	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	401	DCP	C2-N1	5.07	1.43	1.38
5	B	401	DCP	C2-N1	4.60	1.43	1.38

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	401	DCP	C6-C5-C4	8.32	120.92	117.47
5	B	401	DCP	C6-C5-C4	5.26	119.66	117.47
5	B	401	DCP	PB-O3B-PG	-4.13	119.57	131.68
5	B	401	DCP	O4'-C1'-N1	3.99	115.19	107.68
5	A	401	DCP	O4'-C1'-N1	3.59	114.42	107.68
5	B	401	DCP	C2-N3-C4	3.52	120.66	115.57
5	A	401	DCP	PB-O3B-PG	-3.04	122.78	131.68
5	A	401	DCP	C2-N3-C4	2.87	119.73	115.57
5	A	401	DCP	O3B-PB-O3A	-2.84	95.89	101.66
5	B	401	DCP	O5'-C5'-C4'	2.79	119.17	108.94
5	B	401	DCP	O4'-C4'-C5'	2.76	119.22	109.36
5	A	401	DCP	PB-O3A-PA	-2.60	124.07	131.68
5	B	401	DCP	O3G-PG-O2G	2.12	115.86	107.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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 ²)	Q<0.9
1	C	17/18 (94%)	0.26	1 (5%) 22 30	47, 62, 135, 137	2 (11%)
1	T	13/18 (72%)	1.46	5 (38%) 1 1	61, 86, 137, 139	4 (30%)
2	D	12/13 (92%)	0.55	2 (16%) 2 4	61, 70, 126, 133	2 (16%)
2	P	11/13 (84%)	2.50	5 (45%) 1 1	50, 89, 123, 127	4 (36%)
3	A	343/348 (98%)	0.33	9 (2%) 53 63	28, 53, 82, 103	3 (0%)
3	B	342/348 (98%)	0.82	35 (10%) 7 12	43, 79, 114, 137	8 (2%)
All	All	738/758 (97%)	0.61	57 (7%) 14 20	28, 64, 112, 139	23 (3%)

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	4	DG	7.8
3	B	213	LEU	7.4
3	B	116	ARG	7.2
1	T	16	DC	5.6
3	B	207	SER	5.0
1	T	13	DT	4.9
2	P	3	DG	4.8
2	P	8	DG	4.5
2	P	6	DA	4.4
3	B	295	ILE	4.0
3	A	0	HIS	3.9
3	B	296	VAL	3.8
2	D	2	DG	3.8
2	P	5	DG	3.7
3	B	189	ILE	3.6
3	B	327	GLU	3.4
3	B	241	VAL	3.3
3	B	258	LEU	3.2
3	B	292	ASP	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	B	119	ARG	3.1
3	B	209	GLU	3.1
3	B	212	LYS	3.1
3	B	115	VAL	3.1
3	B	188	ASN	2.9
3	B	278	LYS	2.9
3	B	195	LYS	2.8
1	T	12	DT	2.8
3	A	210	PHE	2.7
3	A	274	TYR	2.7
3	B	210	PHE	2.7
3	B	324	GLU	2.6
3	B	321	LYS	2.6
3	B	113	ASP	2.5
3	B	325	GLU	2.5
2	D	3	DG	2.5
3	B	163	ILE	2.5
3	B	196	LYS	2.5
3	A	318	LEU	2.5
1	T	15	DC	2.5
3	B	37	PHE	2.4
3	B	228	LEU	2.4
3	A	253	ARG	2.4
3	A	234	ASN	2.4
3	B	206	LEU	2.4
3	A	286	VAL	2.4
3	B	330	ILE	2.4
1	C	17	DC	2.3
3	B	187	GLY	2.3
3	B	192	GLU	2.3
3	B	230	ARG	2.2
3	B	35	GLY	2.2
3	B	151	ALA	2.1
3	A	4	LEU	2.1
3	A	312	TYR	2.1
3	B	284	ILE	2.0
1	T	14	DC	2.0
3	B	91	LEU	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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	DOC	D	13	18/19	0.25	4.58	51,57,77,78	3
2	DOC	P	13	18/19	0.23	2.00	85,87,89,91	1
1	EFG	T	5	25/26	0.14	-0.46	72,79,100,107	0
1	EFG	C	5	25/26	0.11	-1.05	44,50,56,59	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	MG	A	403	1/1	0.27	20.69	29,29,29,29	0
6	MG	B	402	1/1	0.23	2.92	40,40,40,40	0
4	CA	A	404	1/1	0.18	2.23	117,117,117,117	0
4	CA	B	403	1/1	0.21	1.91	77,77,77,77	0
4	CA	D	101	1/1	0.19	1.33	73,73,73,73	0
6	MG	B	404	1/1	0.21	0.99	52,52,52,52	0
6	MG	A	402	1/1	0.18	0.78	19,19,19,19	0
5	DCP	B	401	28/28	0.14	-0.06	59,67,74,79	0
5	DCP	A	401	28/28	0.12	-0.98	31,46,52,53	0

6.5 Other polymers ⓘ

There are no such residues in this entry.