



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 19, 2016 – 06:22 PM EDT

PDB ID : 5HIO  
Title : Crystal structure of the DNA polymerase lambda binary complex  
Authors : Burak, M.J.; Guja, K.E.; Garcia-Diaz, M.  
Deposited on : 2016-03-01  
Resolution : 2.08 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
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) : rb-20027939

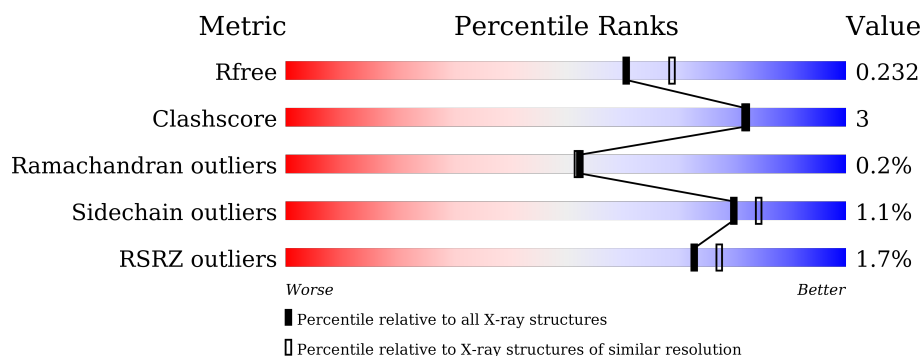
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.08 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	4546 (2.10-2.06)
Clashscore	102246	5101 (2.10-2.06)
Ramachandran outliers	100387	5048 (2.10-2.06)
Sidechain outliers	100360	5049 (2.10-2.06)
RSRZ outliers	91569	4556 (2.10-2.06)

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	11	<div> <div>82%</div> <div>18%</div> </div>
1	F	11	<div> <div>82%</div> <div>18%</div> </div>
1	J	11	<div> <div>100%</div> </div>
1	N	11	<div> <div>82%</div> <div>18%</div> </div>
2	C	6	<div> <div>100%</div> </div>
2	G	6	<div> <div>100%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	K	6	<div><div></div>100%</div>
2	O	6	<div><div></div>67%<div></div>33%</div>
3	D	4	<div><div></div>75%<div></div>25%</div>
3	H	4	<div><div></div>25%<div></div>50%<div></div>25%</div>
3	L	4	<div><div></div>75%<div></div>25%</div>
3	P	4	<div><div></div>75%<div></div>25%</div>
4	A	334	<div><div></div>%<div></div>92%<div></div>5% .</div>
4	E	334	<div><div></div>4%<div></div>87%<div></div>9% . .</div>
4	I	334	<div><div></div>2%<div></div>90%<div></div>8% .</div>
4	M	334	<div><div></div>%<div></div>93%<div></div>5% .</div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13086 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(\*CP\*GP\*GP\*CP\*(8OG)P\*GP\*TP\*AP\*CP\*TP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	11	Total	C	N	O	P	0	0	0
			226	107	43	66	10			
1	F	11	Total	C	N	O	P	0	0	0
			226	107	43	66	10			
1	J	11	Total	C	N	O	P	0	0	0
			226	107	43	66	10			
1	N	11	Total	C	N	O	P	0	0	0
			226	107	43	66	10			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	P	0	0	0
			119	58	23	33	5			
2	G	6	Total	C	N	O	P	0	0	0
			119	58	23	33	5			
2	K	6	Total	C	N	O	P	0	0	0
			119	58	23	33	5			
2	O	6	Total	C	N	O	P	0	0	0
			119	58	23	33	5			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*GP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	H	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	L	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	P	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			

- Molecule 4 is a protein called DNA polymerase lambda.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	327	Total	C	N	O	S	0	1	0
			2563	1607	467	477	12			
4	E	325	Total	C	N	O	S	0	1	0
			2504	1572	455	465	12			
4	I	326	Total	C	N	O	S	0	0	0
			2531	1591	462	466	12			
4	M	327	Total	C	N	O	S	0	0	0
			2557	1605	468	472	12			

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	I	1	Total	Na	0	0
			1	1		
5	A	2	Total	Na	0	0
			2	2		
5	M	2	Total	Na	0	0
			2	2		
5	E	1	Total	Na	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	35	Total	O	0	0
			35	35		
6	C	20	Total	O	0	0
			20	20		
6	D	4	Total	O	0	0
			4	4		
6	F	22	Total	O	0	0
			22	22		
6	G	9	Total	O	0	0
			9	9		
6	H	10	Total	O	0	0
			10	10		
6	J	23	Total	O	0	0
			23	23		
6	K	19	Total	O	0	0
			19	19		
6	L	7	Total	O	0	0
			7	7		

*Continued on next page...*


*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	N	38	Total 38	O 38	0	0
6	O	22	Total 22	O 22	0	0
6	P	10	Total 10	O 10	0	0
6	A	285	Total 285	O 285	0	0
6	E	156	Total 156	O 156	0	0
6	I	191	Total 191	O 191	0	0
6	M	362	Total 362	O 362	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(\*CP\*GP\*GP\*CP\*(8OG)P\*GP\*TP\*AP\*CP\*TP\*G)-3')

Chain B:  82% 18%



- Molecule 1: DNA (5'-D(\*CP\*GP\*GP\*CP\*(8OG)P\*GP\*TP\*AP\*CP\*TP\*G)-3')

Chain F:  82% 18%




- Molecule 1: DNA (5'-D(\*CP\*GP\*GP\*CP\*(8OG)P\*GP\*TP\*AP\*CP\*TP\*G)-3')

Chain J:  100%

There are no outlier residues recorded for this chain.

- Molecule 1: DNA (5'-D(\*CP\*GP\*GP\*CP\*(8OG)P\*GP\*TP\*AP\*CP\*TP\*G)-3')

Chain N:  82% 18%



- Molecule 2: DNA (5'-D(\*CP\*AP\*GP\*TP\*AP\*C)-3')

Chain C:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: DNA (5'-D(\*CP\*AP\*GP\*TP\*AP\*C)-3')

Chain G:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: DNA (5'-D(\*CP\*AP\*GP\*TP\*AP\*C)-3')

Chain K:  100%

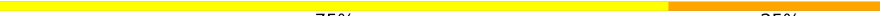
There are no outlier residues recorded for this chain.

- Molecule 2: DNA (5'-D(\*CP\*AP\*GP\*TP\*AP\*C)-3')

Chain O:  67% 33%



- Molecule 3: DNA (5'-D(P\*GP\*CP\*CP\*G)-3')

Chain D:  75% 25%



- Molecule 3: DNA (5'-D(P\*GP\*CP\*CP\*G)-3')

Chain H:  25% 50% 25%




- Molecule 3: DNA (5'-D(P\*GP\*CP\*CP\*G)-3')

Chain L:  75% 25%




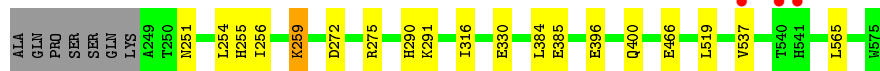
- Molecule 3: DNA (5'-D(P\*GP\*CP\*CP\*G)-3')

Chain P:  75% 25%




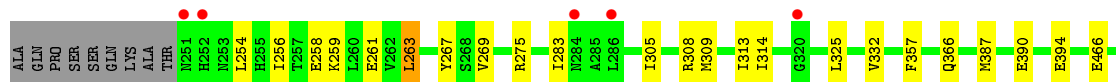
- Molecule 4: DNA polymerase lambda

Chain A:  % 92% 5% .

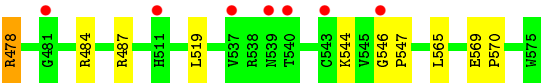


- Molecule 4: DNA polymerase lambda

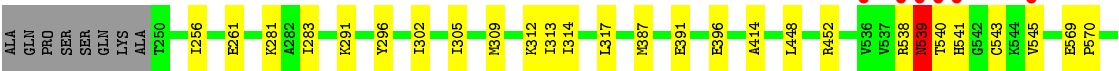
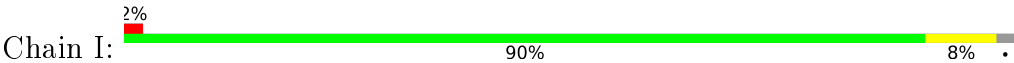
Chain E:  4% 87% 9% . .



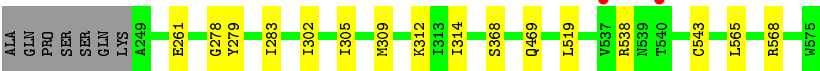
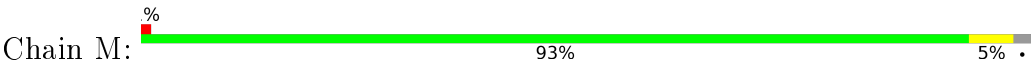




● Molecule 4: DNA polymerase lambda



● Molecule 4: DNA polymerase lambda



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	191.42Å 98.98Å 105.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.09 – 2.08 92.08 – 2.08	Depositor EDS
% Data completeness (in resolution range)	98.1 (45.09-2.08) 93.8 (92.08-2.08)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 2.08Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.191 , 0.235 0.189 , 0.232	Depositor DCC
$R_{free}$ test set	2462 reflections (2.18%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.1	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 51.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13086	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.84% 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, 8OG

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	0.86	0/226	0.98	0/345
1	F	0.77	0/226	0.84	0/345
1	J	0.78	0/226	0.88	0/345
1	N	0.97	0/226	0.99	0/345
2	C	0.87	0/133	0.83	0/203
2	G	0.81	0/133	0.89	0/203
2	K	0.86	0/133	0.88	0/203
2	O	0.94	0/133	0.86	0/203
3	D	1.26	1/92 (1.1%)	0.77	0/138
3	H	1.22	1/92 (1.1%)	0.66	0/138
3	L	1.16	1/92 (1.1%)	0.68	0/138
3	P	1.41	1/92 (1.1%)	0.74	0/138
4	A	0.41	0/2615	0.54	0/3532
4	E	0.35	0/2557	0.47	0/3455
4	I	0.35	0/2582	0.48	0/3487
4	M	0.44	0/2609	0.54	0/3521
All	All	0.52	4/12167 (0.0%)	0.58	0/16739

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	1	DG	OP3-P	-10.78	1.48	1.61
3	H	1	DG	OP3-P	-10.56	1.48	1.61
3	D	1	DG	OP3-P	-10.17	1.49	1.61
3	L	1	DG	OP3-P	-9.73	1.49	1.61

There are no bond angle outliers.

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	B	226	0	125	3	0
1	F	226	0	125	1	0
1	J	226	0	125	0	0
1	N	226	0	125	1	0
2	C	119	0	69	0	0
2	G	119	0	69	0	0
2	K	119	0	69	0	0
2	O	119	0	69	2	0
3	D	83	0	45	2	0
3	H	83	0	45	2	0
3	L	83	0	45	0	0
3	P	83	0	45	6	0
4	A	2563	0	2529	11	0
4	E	2504	0	2447	21	0
4	I	2531	0	2497	14	0
4	M	2557	0	2543	10	0
5	A	2	0	0	0	0
5	E	1	0	0	0	0
5	I	1	0	0	0	0
5	M	2	0	0	0	0
6	A	285	0	0	1	0
6	B	35	0	0	1	0
6	C	20	0	0	0	0
6	D	4	0	0	0	0
6	E	156	0	0	2	1
6	F	22	0	0	0	0
6	G	9	0	0	0	0
6	H	10	0	0	0	0
6	I	191	0	0	2	0
6	J	23	0	0	0	0
6	K	19	0	0	0	0
6	L	7	0	0	0	0
6	M	362	0	0	0	1
6	N	38	0	0	0	0
6	O	22	0	0	2	0
6	P	10	0	0	2	0
All	All	13086	0	10972	69	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 3.

All (69) 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:O:2:DA:N3	6:O:101:HOH:O	2.11	0.82
4:M:519:LEU:HD13	4:M:565:LEU:HD11	1.78	0.64
1:B:2:DG:N7	6:B:101:HOH:O	2.30	0.64
2:O:3:DG:OP1	6:O:102:HOH:O	2.16	0.62
4:M:261:GLU:HG3	4:M:283:ILE:HD13	1.81	0.61
1:N:1:DC:H2'	1:N:2:DG:C8	2.36	0.60
4:I:539:ASN:N	4:I:539:ASN:OD1	2.24	0.58
3:H:3:DC:H2''	3:H:4:DG:C8	2.40	0.57
4:E:519:LEU:HD22	4:E:565:LEU:HD11	1.86	0.56
4:A:251:ASN:HD21	4:A:254:LEU:HA	1.71	0.56
4:I:312:LYS:NZ	6:I:702:HOH:O	2.38	0.56
4:I:569:GLU:HG3	4:I:570:PRO:HD2	1.88	0.56
4:E:254:LEU:O	4:E:258:GLU:HB2	2.07	0.55
4:I:256:ILE:HD13	4:I:313:ILE:HG23	1.89	0.54
4:E:466:GLU:OE2	6:E:701:HOH:O	2.17	0.53
4:I:281:LYS:NZ	6:I:704:HOH:O	2.40	0.53
4:I:302:ILE:HB	4:I:305:ILE:HD12	1.91	0.53
4:E:357:PHE:CZ	4:E:366:GLN:HB2	2.44	0.52
4:I:540:THR:N	4:I:541:HIS:O	2.42	0.52
4:A:251:ASN:ND2	4:A:254:LEU:HA	2.24	0.52
3:H:1:DG:OP3	4:E:308:ARG:NH2	2.40	0.52
4:A:519:LEU:HD22	4:A:565:LEU:HD11	1.92	0.51
4:E:259:LYS:HB3	4:E:325:LEU:HD11	1.92	0.51
1:B:1:DC:H2'	1:B:2:DG:C8	2.46	0.50
1:F:10:DT:H2''	1:F:11:DG:H5''	1.93	0.50
4:E:387:MET:HE1	4:E:487:ARG:HG3	1.94	0.50
4:E:305:ILE:HG23	4:E:309:MET:HB3	1.93	0.49
4:M:302:ILE:HB	4:M:305:ILE:HD12	1.95	0.49
4:I:261:GLU:HG3	4:I:283:ILE:HD13	1.95	0.48
4:I:538:ARG:HA	4:I:543:CYS:O	2.15	0.47
4:E:565:LEU:O	6:E:702:HOH:O	2.21	0.46
3:P:2:DC:OP1	4:M:309:MET:HB2	2.16	0.46
4:E:478:ARG:NH2	4:E:484:ARG:O	2.41	0.46
3:P:3:DC:H2''	3:P:4:DG:C8	2.52	0.45
4:A:256:ILE:HA	4:A:316:ILE:HG21	1.98	0.45
1:B:1:DC:H2''	1:B:2:DG:H5'	1.98	0.45
4:A:396:GLU:OE2	4:A:400:GLN:NE2	2.49	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:I:296:TYR:HB2	4:I:314:ILE:HG21	1.98	0.45
4:A:275:ARG:NH2	6:A:718:HOH:O	2.50	0.44
4:A:466[B]:GLU:OE1	4:M:469:GLN:HA	2.17	0.44
4:E:390:GLU:O	4:E:394:GLU:HG3	2.17	0.44
4:E:314:ILE:HA	4:E:314:ILE:HD12	1.86	0.44
4:E:263:LEU:HD13	4:E:325:LEU:HD23	1.99	0.44
3:D:1:DG:H2''	3:D:2:DC:H5'	2.00	0.44
3:D:3:DC:H2''	3:D:4:DG:C8	2.53	0.44
4:E:261:GLU:HG3	4:E:283:ILE:HD13	1.98	0.44
4:E:569:GLU:HG3	4:E:570:PRO:HD2	2.00	0.43
4:I:448:LEU:O	4:I:452:ARG:HB2	2.17	0.43
4:M:305:ILE:HG23	4:M:309:MET:HB3	2.00	0.43
4:E:267:TYR:CZ	4:E:275:ARG:HD3	2.53	0.43
4:A:384:LEU:HD12	4:A:384:LEU:HA	1.70	0.43
3:P:4:DG:N3	6:P:102:HOH:O	2.37	0.43
4:A:290:HIS:CD2	4:A:291:LYS:HB2	2.53	0.43
4:E:357:PHE:CE1	4:E:366:GLN:HB2	2.54	0.43
4:I:387:MET:HE2	4:I:391:GLU:HB3	2.00	0.42
4:E:256:ILE:HD13	4:E:313:ILE:HG23	2.01	0.42
3:P:1:DG:H1'	4:M:278:GLY:HA3	2.01	0.42
4:E:269:VAL:HG21	4:E:332:VAL:HG13	2.02	0.42
3:P:1:DG:H2''	3:P:2:DC:H5'	2.02	0.42
4:M:279:TYR:OH	4:M:312:LYS:HE2	2.20	0.41
4:M:538:ARG:HA	4:M:543:CYS:O	2.20	0.41
4:I:309:MET:O	4:I:313:ILE:HG13	2.20	0.41
3:P:1:DG:OP2	6:P:101:HOH:O	2.21	0.41
4:E:387:MET:HE3	4:E:387:MET:HB2	1.97	0.41
4:I:396:GLU:HG3	4:I:414:ALA:HB2	2.02	0.41
4:M:314:ILE:HA	4:M:314:ILE:HD13	1.86	0.41
4:A:272:ASP:OD2	4:A:275:ARG:NH1	2.54	0.40
4:A:255:HIS:CE1	4:A:259:LYS:HE3	2.57	0.40
4:E:546:GLY:HA2	4:E:547:PRO:HD3	1.83	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)
6:E:805:HOH:O	6:M:852:HOH:O[2_565]	2.12	0.08

## 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	
4	A	326/334 (98%)	320 (98%)	6 (2%)	0	100	100
4	E	324/334 (97%)	313 (97%)	10 (3%)	1 (0%)	46	44
4	I	324/334 (97%)	316 (98%)	7 (2%)	1 (0%)	46	44
4	M	325/334 (97%)	319 (98%)	6 (2%)	0	100	100
All	All	1299/1336 (97%)	1268 (98%)	29 (2%)	2 (0%)	52	52

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	I	539	ASN
4	E	544	LYS

### 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	
4	A	269/280 (96%)	265 (98%)	4 (2%)	72	77
4	E	256/280 (91%)	254 (99%)	2 (1%)	86	90
4	I	262/280 (94%)	258 (98%)	4 (2%)	72	77
4	M	270/280 (96%)	268 (99%)	2 (1%)	88	92
All	All	1057/1120 (94%)	1045 (99%)	12 (1%)	80	84

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	259	LYS
4	A	330	GLU
4	A	385	GLU
4	A	537	VAL
4	E	263	LEU
4	E	478	ARG
4	I	291	LYS
4	I	317	LEU
4	I	539	ASN
4	I	545	VAL
4	M	368	SER
4	M	568	ARG

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 molecules 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	8OG	B	5	1	16,25,26	5.16	10 (62%)	20,37,40	2.12	7 (35%)
1	8OG	F	5	1	16,25,26	5.19	10 (62%)	20,37,40	2.31	8 (40%)
1	8OG	J	5	1	16,25,26	5.20	10 (62%)	20,37,40	1.79	3 (15%)
1	8OG	N	5	1	16,25,26	4.94	10 (62%)	20,37,40	1.83	4 (20%)

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	8OG	B	5	1	-	0/3/21/22	0/3/3/3
1	8OG	F	5	1	-	0/3/21/22	0/3/3/3
1	8OG	J	5	1	-	0/3/21/22	0/3/3/3
1	8OG	N	5	1	-	0/3/21/22	0/3/3/3

All (40) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	5	8OG	C3'-C4'	-8.05	1.30	1.53
1	F	5	8OG	C3'-C4'	-7.64	1.31	1.53
1	J	5	8OG	C3'-C4'	-7.36	1.32	1.53
1	N	5	8OG	C3'-C4'	-7.20	1.32	1.53
1	F	5	8OG	O4'-C1'	-5.08	1.30	1.42
1	B	5	8OG	O4'-C1'	-4.67	1.31	1.42
1	J	5	8OG	O4'-C1'	-4.60	1.31	1.42
1	N	5	8OG	O4'-C1'	-4.44	1.32	1.42
1	B	5	8OG	O3'-C3'	2.36	1.48	1.43
1	F	5	8OG	O3'-C3'	2.69	1.49	1.43
1	N	5	8OG	O3'-C3'	2.76	1.49	1.43
1	J	5	8OG	O3'-C3'	2.95	1.49	1.43
1	N	5	8OG	C2-N2	3.11	1.40	1.34
1	J	5	8OG	C2-N2	3.40	1.41	1.34
1	B	5	8OG	C6-C5	3.66	1.48	1.41
1	F	5	8OG	C2-N2	3.71	1.41	1.34
1	B	5	8OG	C2-N2	3.74	1.42	1.34
1	F	5	8OG	C6-C5	3.91	1.49	1.41
1	N	5	8OG	C4-N3	4.19	1.42	1.35
1	J	5	8OG	C6-C5	4.20	1.49	1.41
1	N	5	8OG	C6-C5	4.21	1.49	1.41
1	J	5	8OG	C6-N1	4.35	1.40	1.33
1	B	5	8OG	C6-N1	4.45	1.41	1.33
1	N	5	8OG	C2-N1	4.96	1.44	1.35
1	J	5	8OG	C2-N1	5.03	1.44	1.35
1	N	5	8OG	C6-N1	5.14	1.42	1.33
1	F	5	8OG	C6-N1	5.19	1.42	1.33
1	B	5	8OG	C2-N1	5.20	1.45	1.35
1	F	5	8OG	C4-N3	5.22	1.44	1.35
1	F	5	8OG	C2-N1	5.34	1.45	1.35
1	J	5	8OG	C4-N3	5.34	1.44	1.35
1	B	5	8OG	C4-N3	5.41	1.44	1.35
1	N	5	8OG	O4'-C4'	7.31	1.61	1.45
1	F	5	8OG	O4'-C4'	7.59	1.62	1.45

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	5	8OG	O4'-C4'	7.81	1.62	1.45
1	B	5	8OG	O4'-C4'	7.86	1.63	1.45
1	N	5	8OG	C8-N7	12.48	1.47	1.34
1	B	5	8OG	C8-N7	12.78	1.48	1.34
1	F	5	8OG	C8-N7	12.82	1.48	1.34
1	J	5	8OG	C8-N7	13.41	1.48	1.34

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	N	5	8OG	N3-C2-N1	-5.18	120.51	127.56
1	B	5	8OG	N3-C2-N1	-5.11	120.61	127.56
1	F	5	8OG	N3-C2-N1	-4.99	120.76	127.56
1	J	5	8OG	N3-C2-N1	-4.85	120.95	127.56
1	F	5	8OG	O4'-C1'-N9	-4.36	104.45	108.16
1	F	5	8OG	C5-C6-N1	-4.27	117.94	123.52
1	J	5	8OG	C5-C6-N1	-3.46	119.00	123.52
1	B	5	8OG	C5-C6-N1	-3.26	119.26	123.52
1	B	5	8OG	C8-N9-C1'	-3.10	120.92	125.83
1	N	5	8OG	C5-C6-N1	-3.00	119.61	123.52
1	F	5	8OG	C8-N9-C1'	-2.19	122.36	125.83
1	B	5	8OG	O3'-C3'-C4'	-2.08	101.54	110.10
1	B	5	8OG	C4'-O4'-C1'	-2.03	104.21	109.42
1	F	5	8OG	C4'-O4'-C1'	-2.01	104.27	109.42
1	F	5	8OG	N2-C2-N1	2.05	120.58	117.20
1	F	5	8OG	C1'-N9-C4	2.65	130.40	127.30
1	N	5	8OG	C6-N1-C2	2.76	119.12	115.88
1	N	5	8OG	N2-C2-N1	3.04	122.21	117.20
1	J	5	8OG	C6-N1-C2	3.42	119.89	115.88
1	B	5	8OG	C6-N1-C2	3.44	119.91	115.88
1	B	5	8OG	C1'-N9-C4	3.50	131.41	127.30
1	F	5	8OG	C6-N1-C2	3.80	120.33	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 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	10/11 (90%)	-0.42	0 100 100	30, 41, 53, 54	0
1	F	10/11 (90%)	-0.34	0 100 100	37, 54, 76, 78	0
1	J	10/11 (90%)	-0.36	0 100 100	35, 44, 63, 75	0
1	N	10/11 (90%)	-0.55	0 100 100	25, 33, 47, 50	0
2	C	6/6 (100%)	-0.46	0 100 100	29, 32, 52, 62	0
2	G	6/6 (100%)	-0.52	0 100 100	37, 38, 66, 70	0
2	K	6/6 (100%)	-0.61	0 100 100	33, 36, 53, 58	0
2	O	6/6 (100%)	-0.38	0 100 100	25, 27, 46, 52	0
3	D	4/4 (100%)	-0.47	0 100 100	38, 41, 48, 50	0
3	H	4/4 (100%)	-0.49	0 100 100	46, 46, 54, 65	0
3	L	4/4 (100%)	-0.69	0 100 100	46, 46, 54, 67	0
3	P	4/4 (100%)	-0.46	0 100 100	36, 37, 44, 52	0
4	A	327/334 (97%)	-0.11	3 (0%) 85 88	20, 36, 70, 101	0
4	E	325/334 (97%)	0.03	12 (3%) 45 54	30, 49, 80, 128	0
4	I	326/334 (97%)	-0.19	6 (1%) 71 76	31, 47, 76, 140	0
4	M	327/334 (97%)	-0.19	2 (0%) 90 92	20, 31, 63, 94	0
All	All	1385/1420 (97%)	-0.14	23 (1%) 73 77	20, 41, 75, 140	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	540	THR	6.6
4	I	540	THR	6.2
4	E	539	ASN	4.1
4	I	539	ASN	4.0
4	E	286	LEU	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	I	541	HIS	4.0
4	I	538	ARG	3.2
4	E	543	CYS	3.0
4	M	537	VAL	3.0
4	E	537	VAL	2.9
4	E	511[A]	HIS	2.7
4	A	537	VAL	2.7
4	A	541	HIS	2.6
4	I	536	VAL	2.4
4	M	540	THR	2.4
4	I	545	VAL	2.3
4	E	251	ASN	2.3
4	E	481	GLY	2.2
4	E	284	ASN	2.2
4	E	546	GLY	2.1
4	A	540	THR	2.1
4	E	320	GLY	2.1
4	E	252	HIS	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
1	8OG	B	5	23/24	0.95	0.13	-	28,32,42,46	0
1	8OG	J	5	23/24	0.96	0.12	-	32,37,41,43	0
1	8OG	N	5	23/24	0.98	0.12	-	23,28,33,36	0
1	8OG	F	5	23/24	0.92	0.16	-	34,44,55,62	0

## 6.3 Carbohydrates ⓘ

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	NA	M	602	1/1	0.98	0.10	-0.05	34,34,34,34	0
5	NA	M	601	1/1	0.98	0.10	-0.86	24,24,24,24	0
5	NA	A	602	1/1	0.95	0.08	-1.45	46,46,46,46	0
5	NA	E	601	1/1	0.94	0.07	-1.93	39,39,39,39	0
5	NA	I	601	1/1	0.97	0.06	-3.20	33,33,33,33	0
5	NA	A	601	1/1	0.99	0.07	-3.38	35,35,35,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.