



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 28, 2023 – 03:58 AM EDT

PDB ID : 3KHH  
Title : Dpo4 extension ternary complex with a C base opposite the 2-aminofluorene-guanine [AF]G lesion  
Authors : Rechkoblit, O.; Malinina, L.; Patel, D.J.  
Deposited on : 2009-10-30  
Resolution : 2.70 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35

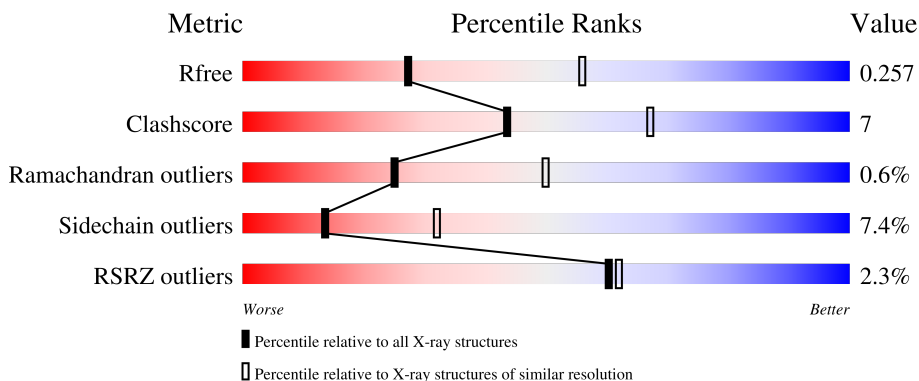
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.



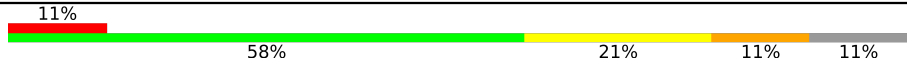
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	341	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 80%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 19%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 0%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 0%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 10px;">3%      80%      19%      .</p>
1	B	341	<div style="display: flex; align-items: center;"> <div style="width: 0%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 82%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 0%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 0%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 10px;">%      82%      16%      .</p>
2	D	13	<div style="display: flex; align-items: center;"> <div style="width: 54%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 46%; height: 10px; background-color: yellow; margin-right: 5px;"></div> </div> <p style="margin-left: 10px;">54%      46%</p>
2	H	13	<div style="display: flex; align-items: center;"> <div style="width: 31%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 54%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 15%; height: 10px; background-color: orange; margin-right: 5px;"></div> </div> <p style="margin-left: 10px;">31%      54%      15%</p>
3	E	19	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 26%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 42%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 10px;">5%      26%      42%      16%      16%</p>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	J	19	 <p>A horizontal bar chart representing the quality of chain J. The bar is divided into five segments: a red segment (11%), a green segment (58%), a yellow segment (21%), an orange segment (11%), and a grey segment (11%).</p>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6862 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 IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	341	2740	1757	472	505	6	0	0	0
1	B	341	2740	1757	472	505	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q97W02
B	1001	GLY	-	expression tag	UNP Q97W02

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	13	269	129	51	77	12	0	0	0
2	H	13	272	129	51	79	13	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	814	DOC	C	engineered mutation	PDB 3KHH
H	1814	DOC	C	engineered mutation	PDB 3KHH

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

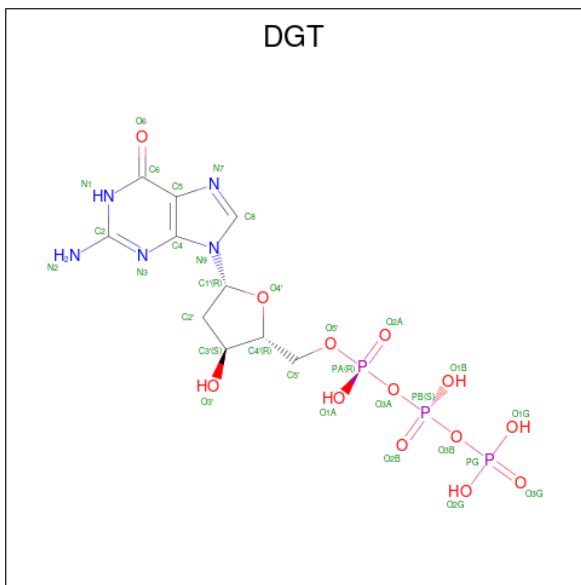
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	16	316	152	58	91	15	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	J	17	337	162	63	96	16	0	0	0

- Molecule 4 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>).

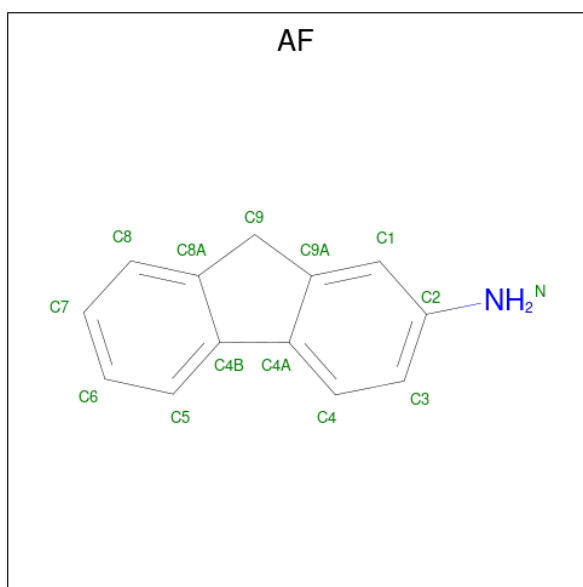


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	31	10	5	13	3	0	0
4	B	1	31	10	5	13	3	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	3	3	3	0	0
5	B	3	3	3	0	0

- Molecule 6 is 2-AMINOFLUORENE (three-letter code: AF) (formula: C<sub>13</sub>H<sub>11</sub>N).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total	C	N	0	0
			14	13	1		
6	J	1	Total	C	N	0	0
			14	13	1		

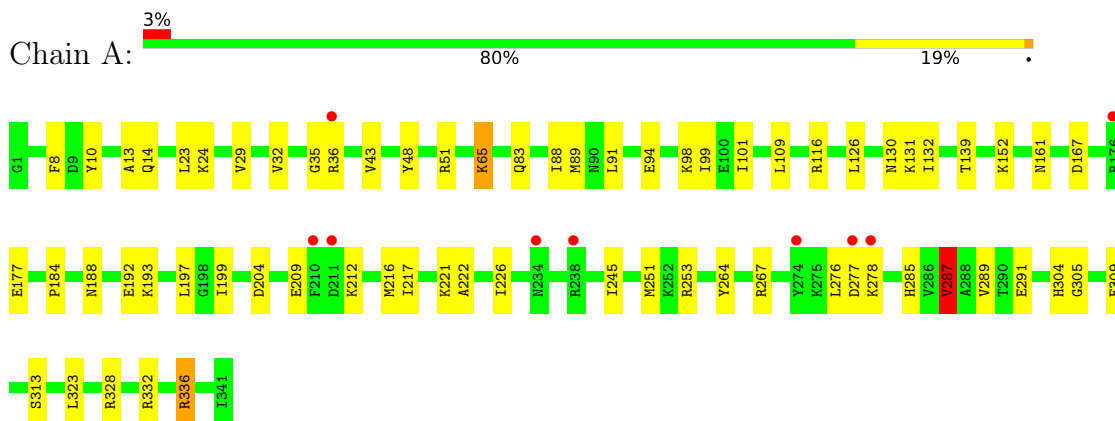
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	33	Total	O	0	0
			33	33		
7	D	2	Total	O	0	0
			2	2		
7	E	7	Total	O	0	0
			7	7		
7	B	38	Total	O	0	0
			38	38		
7	H	9	Total	O	0	0
			9	9		
7	J	3	Total	O	0	0
			3	3		

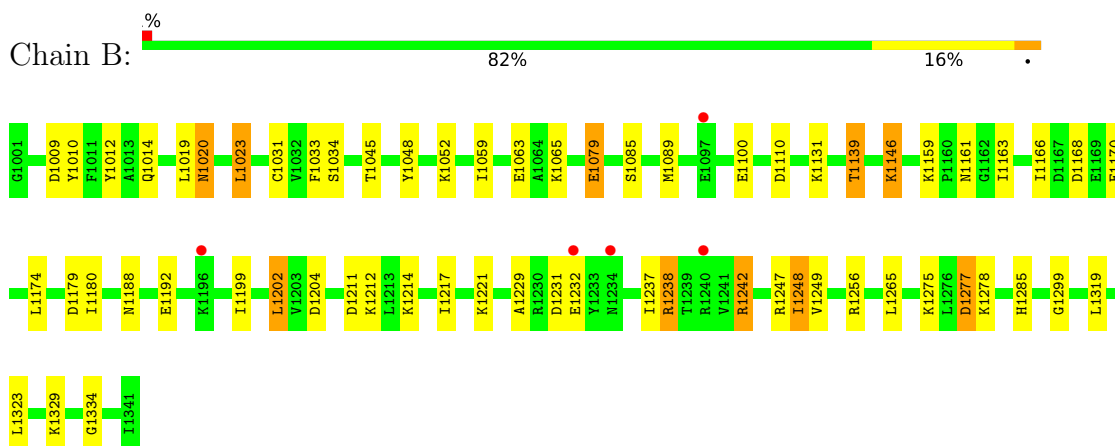
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: DNA polymerase IV



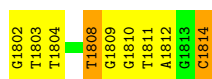
- Molecule 1: DNA polymerase IV



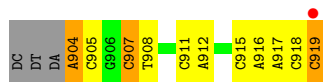
- Molecule 2: 5'-D(\*GP\*TP\*TP\*GP\*GP\*AP\*TP\*GP\*GP\*TP\*AP\*GP\*(DOC))-3'



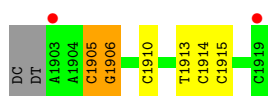
- Molecule 2: 5'-D(\*GP\*TP\*TP\*GP\*GP\*AP\*TP\*GP\*GP\*TP\*AP\*GP\*(DOC))-3'



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



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





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.75Å 98.80Å 98.64Å 90.00° 99.08° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 19.94 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.8 (20.00-2.70) 97.8 (19.94-2.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.56 (at 2.71Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.192 , 0.252 0.193 , 0.257	Depositor DCC
$R_{free}$ test set	1415 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.1	Xtrriage
Anisotropy	0.154	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 52.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6862	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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: CA, AF, DGT, DOC

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.45	0/2779	0.59	1/3731 (0.0%)
1	B	0.45	0/2779	0.56	0/3731
2	D	1.10	2/282 (0.7%)	1.47	5/436 (1.1%)
2	H	0.81	0/285	1.63	7/440 (1.6%)
3	E	1.09	1/353 (0.3%)	1.60	8/540 (1.5%)
3	J	0.91	0/377	1.56	7/577 (1.2%)
All	All	0.59	3/6855 (0.0%)	0.88	28/9455 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	802	DG	C6-N1	7.60	1.44	1.39
3	E	917	DA	C5-C4	5.71	1.42	1.38
2	D	802	DG	N7-C5	5.08	1.42	1.39

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1804	DT	O4'-C1'-N1	8.01	113.61	108.00
2	H	1802	DG	O4'-C4'-C3'	-7.69	101.39	106.00
2	D	805	DG	O4'-C1'-N9	7.62	113.33	108.00
3	J	1915	DC	O4'-C1'-N1	7.23	113.06	108.00
3	E	904	DA	P-O3'-C3'	7.09	128.20	119.70
2	D	810	DG	O4'-C4'-C3'	-6.86	101.76	104.50
3	E	915	DC	P-O3'-C3'	6.86	127.93	119.70
3	J	1906	DG	O4'-C1'-N9	6.77	112.74	108.00
3	J	1910	DC	O4'-C1'-N1	6.72	112.70	108.00
3	E	915	DC	O4'-C4'-C3'	-6.70	101.82	104.50
2	D	806	DG	O4'-C1'-N9	6.69	112.68	108.00
2	H	1810	DG	O4'-C1'-N9	6.44	112.51	108.00

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	1905	DC	O4'-C4'-C3'	-6.43	101.93	104.50
3	J	1914	DC	O4'-C1'-N1	6.35	112.45	108.00
2	H	1808	DT	C1'-O4'-C4'	-6.18	103.92	110.10
2	D	810	DG	C1'-O4'-C4'	-6.11	104.00	110.10
2	D	809	DG	O4'-C1'-N9	6.07	112.25	108.00
3	E	907	DC	C1'-O4'-C4'	-5.98	104.12	110.10
3	J	1913	DT	O4'-C1'-N1	5.85	112.09	108.00
2	H	1811	DT	O4'-C1'-N1	5.79	112.05	108.00
3	J	1906	DG	C1'-O4'-C4'	-5.57	104.53	110.10
1	A	287	VAL	CB-CA-C	-5.48	100.99	111.40
3	E	918	DC	O4'-C1'-N1	5.40	111.78	108.00
3	E	915	DC	C1'-O4'-C4'	-5.21	104.89	110.10
2	H	1803	DT	O4'-C1'-N1	5.19	111.63	108.00
2	H	1810	DG	P-O3'-C3'	5.17	125.91	119.70
3	E	916	DA	N1-C2-N3	-5.03	126.78	129.30
3	E	919	DC	C2-N1-C1'	5.03	124.33	118.80

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	2740	0	2883	34	0
1	B	2740	0	2880	39	0
2	D	269	0	149	0	0
2	H	272	0	148	4	0
3	E	316	0	180	10	0
3	J	337	0	191	4	0
4	A	31	0	12	5	0
4	B	31	0	12	3	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
6	E	14	0	9	2	0
6	J	14	0	10	1	0
7	A	33	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	38	0	0	3	0
7	D	2	0	0	0	0
7	E	7	0	0	1	0
7	H	9	0	0	3	0
7	J	3	0	0	0	0
All	All	6862	0	6474	89	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 (89) 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:309:GLU:HG3	7:A:520:HOH:O	1.35	1.21
1:B:1014:GLN:HE22	1:B:1139:THR:H	1.02	0.97
1:B:1012:TYR:HB2	1:B:1045:THR:HG21	1.49	0.95
1:B:1014:GLN:NE2	1:B:1139:THR:H	1.74	0.85
1:A:14:GLN:HE22	1:A:139:THR:H	1.27	0.83
1:A:291:GLU:HG3	7:A:540:HOH:O	1.81	0.81
1:B:1014:GLN:HE22	1:B:1139:THR:N	1.79	0.81
1:B:1012:TYR:HB2	1:B:1045:THR:CG2	2.11	0.80
1:B:1034:SER:HB3	3:J:1905:DC:H5'	1.65	0.77
1:A:285:HIS:HD2	7:A:529:HOH:O	1.70	0.74
4:A:414:DGT:HN2	3:E:904:DA:H5''	1.54	0.72
1:B:1079:GLU:CD	1:B:1079:GLU:H	1.93	0.72
4:B:1414:DGT:H5'	2:H:1814:DOC:H2''	1.72	0.71
4:B:1414:DGT:O2G	4:B:1414:DGT:O2B	2.09	0.69
1:B:1285:HIS:HD2	7:B:11:HOH:O	1.75	0.68
1:A:65:LYS:HE3	1:B:1161:ASN:O	1.94	0.67
4:A:414:DGT:H8	4:A:414:DGT:H5'A	1.78	0.65
1:B:1009:ASP:OD2	1:B:1139:THR:HG22	1.98	0.64
1:A:88:ILE:HG23	1:A:132:ILE:HD12	1.79	0.64
1:A:99:ILE:HG22	1:A:109:LEU:HG	1.80	0.63
1:A:152:LYS:HG2	1:A:184:PRO:HG3	1.81	0.62
1:A:10:TYR:HA	4:A:414:DGT:O1G	2.02	0.59
3:E:907:DC:H2'	3:E:908:DT:H72	1.84	0.58
1:A:277:ASP:O	1:A:278:LYS:HB2	2.03	0.58
1:A:188:ASN:O	1:A:192:GLU:HG2	2.04	0.58
1:B:1166:ILE:HD13	1:B:1174:LEU:HD11	1.86	0.57
1:B:1009:ASP:OD2	1:B:1139:THR:CG2	2.52	0.57
1:A:10:TYR:CZ	1:A:14:GLN:HG3	2.40	0.56

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1034:SER:HB3	3:J:1905:DC:C5'	2.35	0.56
1:A:10:TYR:CE1	1:A:14:GLN:HG3	2.41	0.56
1:A:289:VAL:HB	1:A:332:ARG:HB2	1.87	0.56
1:A:287:VAL:HG22	6:E:926:AF:C6	2.35	0.56
1:B:1247:ARG:CZ	1:B:1249:VAL:HG12	2.36	0.55
1:B:1166:ILE:HG23	1:B:1170:GLU:HB3	1.89	0.55
3:J:1906:DG:N7	6:J:1926:AF:H1	2.22	0.55
1:A:336:ARG:NH2	7:A:537:HOH:O	2.39	0.54
1:A:304:HIS:HD2	1:A:305:GLY:O	1.90	0.54
3:E:904:DA:H2''	3:E:905:DC:OP2	2.08	0.54
2:H:1812:DA:H5''	7:H:54:HOH:O	2.07	0.54
1:A:217:ILE:HD12	1:A:221:LYS:HB3	1.89	0.54
4:A:414:DGT:N2	3:E:904:DA:H5''	2.22	0.54
1:B:1014:GLN:HE22	1:B:1139:THR:HB	1.73	0.54
1:A:91:LEU:O	1:A:94:GLU:HG2	2.09	0.53
1:B:1242:ARG:NH2	7:B:7:HOH:O	2.42	0.53
1:B:1202:LEU:HD22	1:B:1229:ALA:HB2	1.91	0.52
1:B:1110:ASP:HB2	1:B:1237:ILE:HD13	1.93	0.51
1:A:251:MET:HG2	1:A:264:TYR:CG	2.46	0.51
1:B:1048:TYR:O	1:B:1052:LYS:HG2	2.11	0.50
1:B:1159:LYS:NZ	4:B:1414:DGT:O1B	2.36	0.50
1:B:1020:ASN:C	1:B:1020:ASN:HD22	2.15	0.50
1:B:1277:ASP:O	1:B:1278:LYS:HB2	2.12	0.49
1:B:1014:GLN:NE2	1:B:1139:THR:HB	2.28	0.49
3:E:911:DC:H4'	7:E:49:HOH:O	2.14	0.47
1:A:89:MET:SD	1:A:101:ILE:HD12	2.54	0.47
1:A:35:GLY:HA2	3:E:905:DC:H5	1.80	0.47
1:A:245:ILE:HG13	1:A:276:LEU:HD23	1.96	0.47
1:B:1188:ASN:O	1:B:1192:GLU:HG2	2.15	0.47
1:B:1285:HIS:HE1	7:H:93:HOH:O	1.98	0.46
1:B:1085:SER:O	1:B:1089:MET:HG2	2.15	0.46
3:E:919:DC:O2	3:E:919:DC:H2'	2.16	0.45
3:J:1905:DC:H2''	3:J:1906:DG:H5'	1.99	0.45
1:A:8:PHE:CD1	1:A:8:PHE:N	2.84	0.44
2:H:1812:DA:C5'	7:H:54:HOH:O	2.63	0.44
4:A:414:DGT:H8	4:A:414:DGT:C5'	2.45	0.44
1:B:1275:LYS:HA	7:B:55:HOH:O	2.17	0.44
1:B:1020:ASN:HB3	1:B:1023:LEU:HD22	2.00	0.44
1:B:1166:ILE:HD13	1:B:1174:LEU:CD1	2.47	0.44
1:B:1285:HIS:ND1	1:B:1299:GLY:HA3	2.32	0.43
3:E:911:DC:H2''	3:E:912:DA:C8	2.54	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1199:ILE:HG23	1:B:1204:ASP:HB2	2.00	0.43
1:B:1265:LEU:HD22	1:B:1319:LEU:HD22	2.01	0.43
1:A:35:GLY:HA2	3:E:905:DC:C5	2.54	0.43
1:A:10:TYR:HD1	1:A:13:ALA:HB3	1.84	0.42
1:A:291:GLU:OE1	1:A:328:ARG:HB3	2.18	0.42
1:B:1100:GLU:HG3	1:B:1238:ARG:O	2.19	0.42
1:B:1031:CYS:HB2	1:B:1033:PHE:CE2	2.54	0.42
3:E:907:DC:H2'	3:E:908:DT:C7	2.49	0.42
1:B:1248:ILE:HA	1:B:1334:GLY:HA3	2.01	0.42
1:A:48:TYR:HA	1:A:51:ARG:HB2	2.02	0.41
1:A:126:LEU:O	1:A:130:ASN:HB2	2.20	0.41
1:A:199:ILE:HG23	1:A:204:ASP:HB2	2.02	0.41
1:A:29:VAL:HG13	1:A:43:VAL:HG13	2.02	0.41
1:A:336:ARG:HB3	6:E:926:AF:C5	2.51	0.41
1:A:222:ALA:O	1:A:226:ILE:HG12	2.20	0.41
1:B:1146:LYS:HE2	1:B:1231:ASP:OD1	2.21	0.40
1:B:1217:ILE:HD12	1:B:1221:LYS:HB3	2.02	0.40
1:B:1256:ARG:HG3	1:B:1329:LYS:HA	2.03	0.40
1:A:152:LYS:HG2	1:A:184:PRO:CG	2.49	0.40
2:H:1808:DT:H2''	2:H:1809:DG:O5'	2.20	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	339/341 (99%)	322 (95%)	15 (4%)	2 (1%)	25	50
1	B	339/341 (99%)	330 (97%)	7 (2%)	2 (1%)	25	50
All	All	678/682 (99%)	652 (96%)	22 (3%)	4 (1%)	25	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	216	MET
1	B	1277	ASP
1	A	161	ASN
1	B	1010	TYR

### 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	299/299 (100%)	278 (93%)	21 (7%)	15	35
1	B	299/299 (100%)	276 (92%)	23 (8%)	13	30
All	All	598/598 (100%)	554 (93%)	44 (7%)	13	32

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	LEU
1	A	24	LYS
1	A	32	VAL
1	A	36	ARG
1	A	65	LYS
1	A	83	GLN
1	A	98	LYS
1	A	116	ARG
1	A	131	LYS
1	A	167	ASP
1	A	177	GLU
1	A	193	LYS
1	A	197	LEU
1	A	209	GLU
1	A	212	LYS
1	A	253	ARG
1	A	267	ARG
1	A	287	VAL
1	A	313	SER
1	A	323	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	336	ARG
1	B	1019	LEU
1	B	1020	ASN
1	B	1023	LEU
1	B	1059	ILE
1	B	1063	GLU
1	B	1065	LYS
1	B	1079	GLU
1	B	1131	LYS
1	B	1139	THR
1	B	1146	LYS
1	B	1163	ILE
1	B	1168	ASP
1	B	1179	ASP
1	B	1180	ILE
1	B	1202	LEU
1	B	1211	ASP
1	B	1212	LYS
1	B	1214	LYS
1	B	1232	GLU
1	B	1238	ARG
1	B	1242	ARG
1	B	1248	ILE
1	B	1323	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	GLN
1	A	123	ASN
1	A	200	ASN
1	A	285	HIS
1	A	304	HIS
1	B	1014	GLN
1	B	1020	ASN
1	B	1082	GLN
1	B	1083	GLN
1	B	1161	ASN
1	B	1188	ASN
1	B	1285	HIS
1	B	1304	HIS



### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	DOC	D	814	2,3	16,19,20	0.47	0	20,26,29	1.03	2 (10%)
2	DOC	H	1814	2,5,3	16,19,20	0.39	0	20,26,29	1.14	2 (10%)

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	D	814	2,3	-	0/7/18/19	0/2/2/2
2	DOC	H	1814	2,5,3	-	2/7/18/19	0/2/2/2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1814	DOC	C4'-O4'-C1'	-3.10	106.88	109.81
2	D	814	DOC	C3'-C2'-C1'	-2.66	99.71	102.78
2	H	1814	DOC	C3'-C2'-C1'	-2.61	99.76	102.78
2	D	814	DOC	C4'-O4'-C1'	-2.12	107.81	109.81

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	H	1814	DOC	O4'-C1'-N1-C6
2	H	1814	DOC	C2'-C1'-N1-C6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	1814	DOC	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 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
6	AF	J	1926	3	14,16,16	1.04	2 (14%)	19,23,23	2.02	5 (26%)
4	DGT	B	1414	5	26,33,33	0.98	3 (11%)	32,52,52	1.84	3 (9%)
6	AF	E	926	3	14,16,16	1.26	2 (14%)	19,23,23	1.76	4 (21%)
4	DGT	A	414	5	26,33,33	1.04	2 (7%)	32,52,52	2.67	8 (25%)

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
6	AF	J	1926	3	-	-	0/3/3/3
4	DGT	B	1414	5	-	6/18/34/34	0/3/3/3
6	AF	E	926	3	-	-	0/3/3/3

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	DGT	A	414	5	-	10/18/34/34	0/3/3/3

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	926	AF	C9-C9A	-2.88	1.50	1.54
4	A	414	DGT	C5-C6	-2.86	1.41	1.47
4	B	1414	DGT	C5-C6	-2.81	1.41	1.47
6	E	926	AF	C9-C8A	-2.74	1.51	1.54
4	A	414	DGT	C8-N7	-2.62	1.30	1.35
4	B	1414	DGT	C8-N7	-2.45	1.30	1.35
6	J	1926	AF	C9-C9A	-2.33	1.51	1.54
4	B	1414	DGT	C5-C4	-2.13	1.37	1.43
6	J	1926	AF	C9-C8A	-2.11	1.51	1.54

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1414	DGT	PB-O3B-PG	-7.47	107.18	132.83
4	A	414	DGT	O1G-PG-O3G	-7.07	83.01	110.68
4	A	414	DGT	O2G-PG-O3G	-5.97	87.30	110.68
4	A	414	DGT	PB-O3B-PG	-5.73	113.15	132.83
4	A	414	DGT	O1G-PG-O3B	5.52	123.14	104.64
4	B	1414	DGT	PA-O3A-PB	-5.39	114.34	132.83
4	A	414	DGT	O3B-PG-O3G	-5.05	83.19	111.19
6	J	1926	AF	C9A-C4A-C4B	4.86	109.86	107.01
6	E	926	AF	C8A-C4B-C4A	4.17	109.45	107.01
4	A	414	DGT	PA-O3A-PB	-3.96	119.24	132.83
6	J	1926	AF	C8A-C4B-C4A	3.54	109.08	107.01
4	A	414	DGT	O2G-PG-O1G	3.52	121.07	107.64
6	E	926	AF	C9A-C4A-C4B	3.44	109.03	107.01
6	E	926	AF	C9A-C9-C8A	3.38	105.66	102.85
6	J	1926	AF	C9A-C9-C8A	3.28	105.58	102.85
6	J	1926	AF	C3-C2-N	-2.64	115.99	120.91
4	A	414	DGT	O2G-PG-O3B	2.49	112.98	104.64
6	J	1926	AF	C2-C1-C9A	-2.47	118.75	121.18
4	B	1414	DGT	O6-C6-C5	2.21	128.68	124.37
6	E	926	AF	C9-C8A-C8	-2.02	131.87	134.47

There are no chirality outliers.

All (16) torsion outliers are listed below:

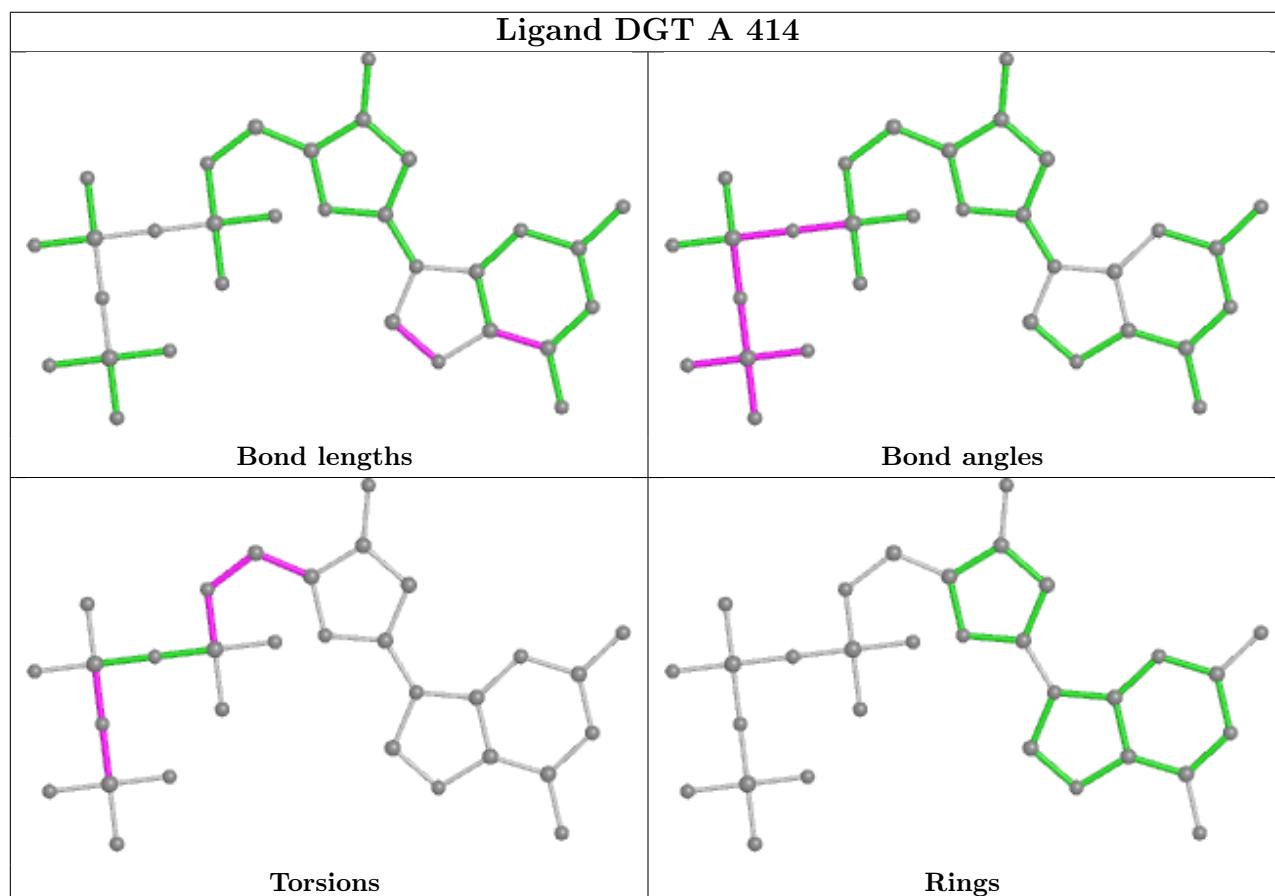
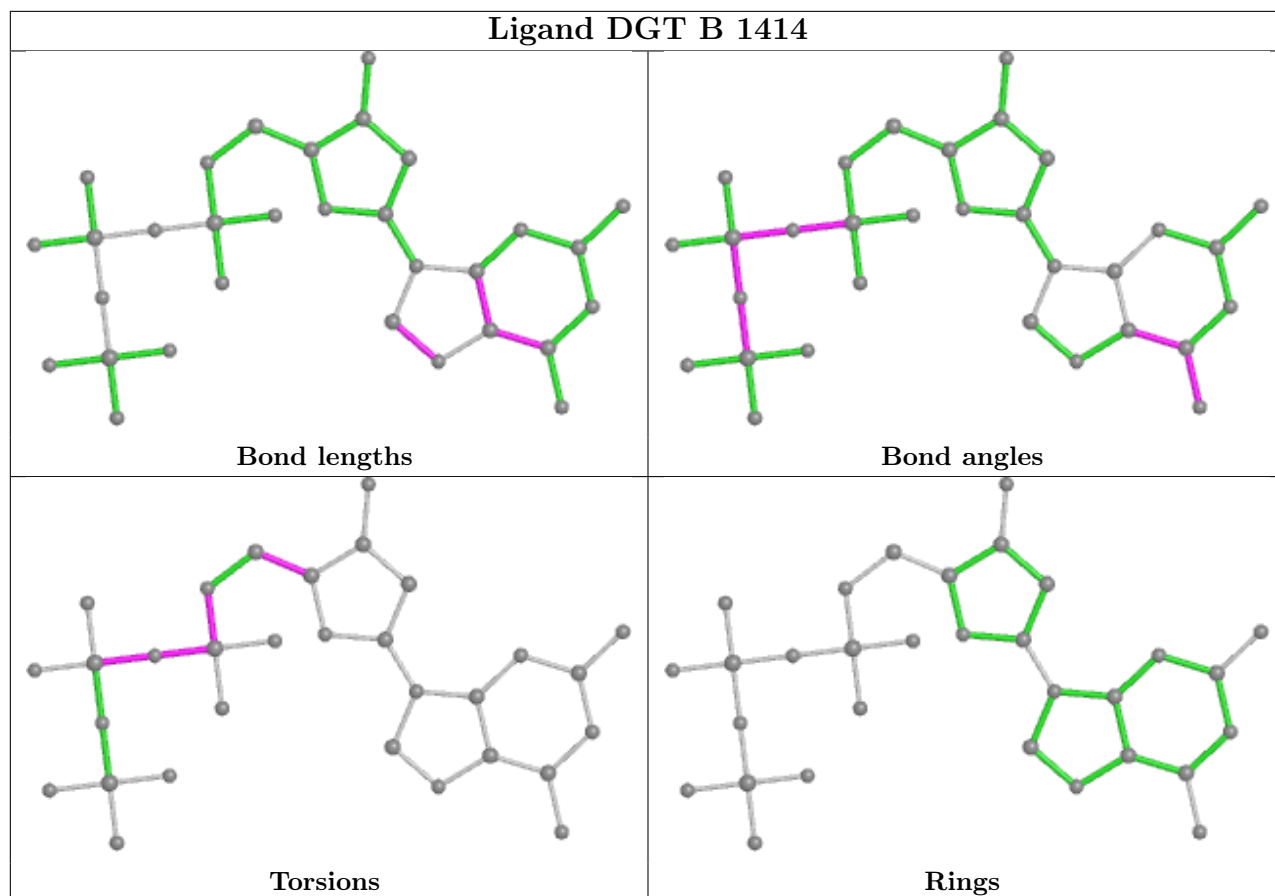
Mol	Chain	Res	Type	Atoms
4	A	414	DGT	C5'-O5'-PA-O3A
4	A	414	DGT	C4'-C5'-O5'-PA
4	A	414	DGT	C3'-C4'-C5'-O5'
4	B	1414	DGT	PB-O3A-PA-O5'
4	B	1414	DGT	C5'-O5'-PA-O3A
4	B	1414	DGT	C5'-O5'-PA-O1A
4	A	414	DGT	O4'-C4'-C5'-O5'
4	B	1414	DGT	O4'-C4'-C5'-O5'
4	B	1414	DGT	C3'-C4'-C5'-O5'
4	A	414	DGT	PG-O3B-PB-O3A
4	A	414	DGT	C5'-O5'-PA-O1A
4	A	414	DGT	C5'-O5'-PA-O2A
4	A	414	DGT	PB-O3B-PG-O3G
4	A	414	DGT	PG-O3B-PB-O1B
4	A	414	DGT	PG-O3B-PB-O2B
4	B	1414	DGT	PA-O3A-PB-O2B

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	J	1926	AF	1	0
4	B	1414	DGT	3	0
6	E	926	AF	2	0
4	A	414	DGT	5	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	341/341 (100%)	0.09	9 (2%) 56 57	14, 28, 40, 46	0
1	B	341/341 (100%)	-0.04	5 (1%) 73 76	14, 27, 39, 45	0
2	D	12/13 (92%)	-0.18	0 100 100	17, 23, 37, 38	0
2	H	12/13 (92%)	0.00	0 100 100	13, 24, 33, 42	0
3	E	16/19 (84%)	0.40	1 (6%) 20 19	23, 32, 50, 56	0
3	J	17/19 (89%)	0.64	2 (11%) 4 3	17, 33, 46, 57	0
All	All	739/746 (99%)	0.04	17 (2%) 60 62	13, 28, 40, 57	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	919	DC	7.0
3	J	1919	DC	4.4
3	J	1903	DA	3.4
1	A	36	ARG	3.3
1	A	238	ARG	2.8
1	A	234	ASN	2.7
1	A	211	ASP	2.5
1	B	1240	ARG	2.5
1	A	210	PHE	2.4
1	B	1232	GLU	2.4
1	B	1196	LYS	2.4
1	A	277	ASP	2.3
1	A	278	LYS	2.3
1	A	274	TYR	2.1
1	B	1234	ASN	2.1
1	A	176	ARG	2.1
1	B	1097	GLU	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. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	DOC	H	1814	18/19	0.96	0.20	21,31,34,34	0
2	DOC	D	814	18/19	0.98	0.19	19,24,25,26	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

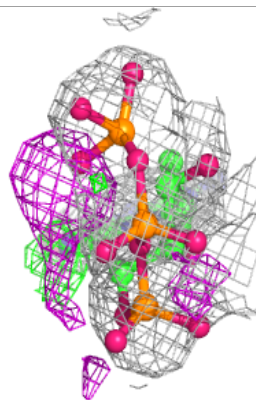
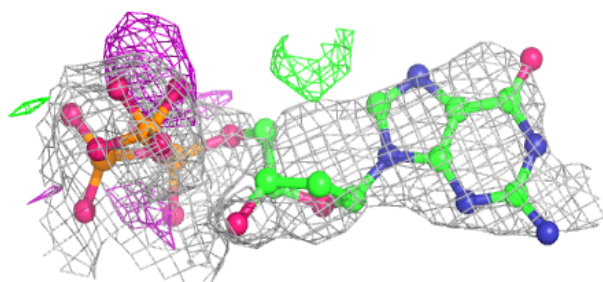
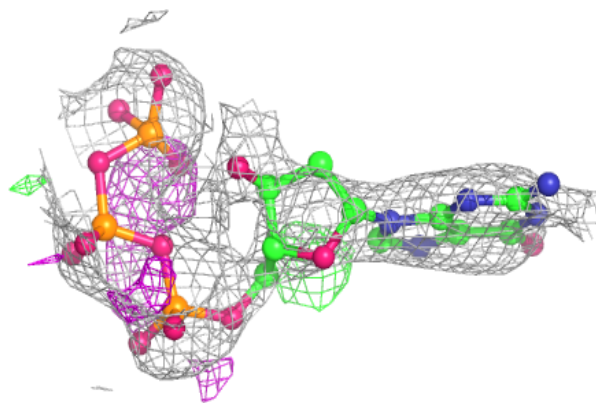
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	CA	A	415	1/1	0.83	0.11	61,61,61,61	0
4	DGT	A	414	31/31	0.88	0.26	42,63,67,67	0
5	CA	B	1417	1/1	0.90	0.15	45,45,45,45	0
4	DGT	B	1414	31/31	0.95	0.14	32,36,39,40	0
5	CA	A	417	1/1	0.96	0.20	43,43,43,43	0
6	AF	E	926	14/14	0.96	0.14	34,37,39,39	0
5	CA	B	1415	1/1	0.98	0.11	42,42,42,42	0
6	AF	J	1926	14/14	0.98	0.14	31,33,35,35	0
5	CA	B	1416	1/1	0.99	0.08	29,29,29,29	0
5	CA	A	416	1/1	0.99	0.04	37,37,37,37	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

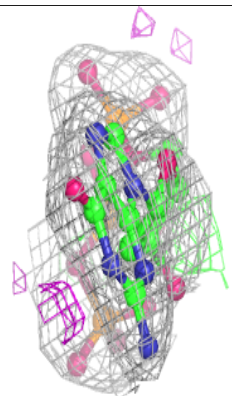
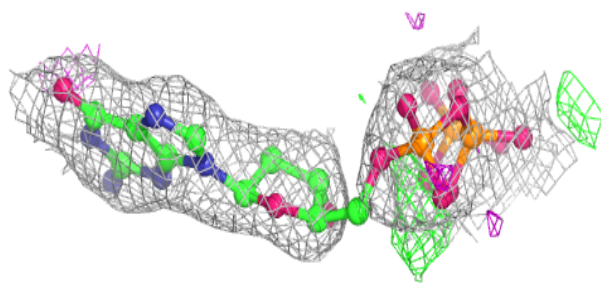
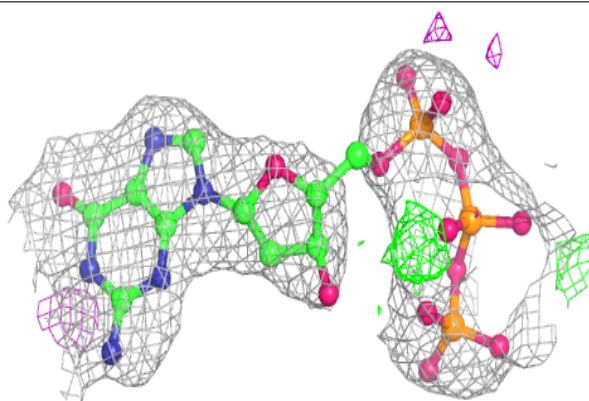


**Electron density around DGT A 414:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around DGT B 1414:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.