



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2023 – 11:56 PM EDT

PDB ID : 3KHR  
Title : Dpo4 post-extension ternary complex with the correct C opposite the 2-aminofluorene-guanine [AF]G lesion  
Authors : Rechkoblit, O.; Malinina, L.; Patel, D.J.  
Deposited on : 2009-10-30  
Resolution : 2.01 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

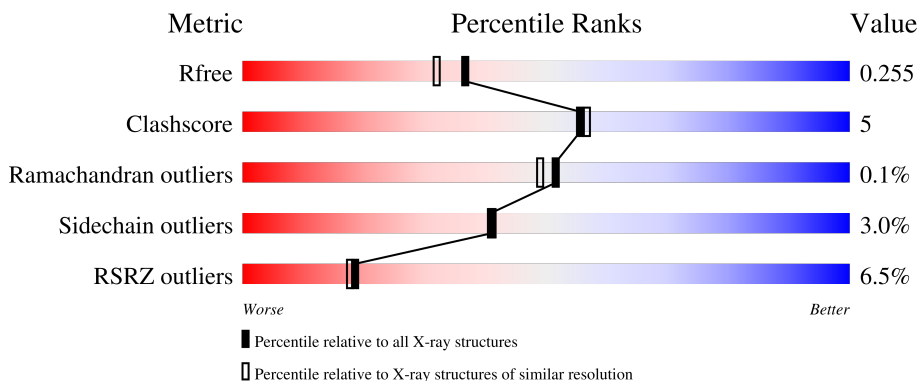
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.01 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	
1	B	341	
2	D	13	
2	H	13	
3	E	19	

*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 (32%), a green segment (53%), a yellow segment (37%), an orange segment (5%), and a grey segment (5%).</p>

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 7237 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	Total 2749	C 1763	N 474	O 506	S 6	0	1	0
1	B	341	Total 2761	C 1769	N 479	O 507	S 6	0	2	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(\*TP\*TP\*GP\*GP\*AP\*TP\*GP\*GP\*TP\*AP\*GP\*C P\*(DDG))-3'.

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

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	815	DDG	G	engineered mutation	PDB 3KHR
H	1815	DDG	G	engineered mutation	PDB 3KHR

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

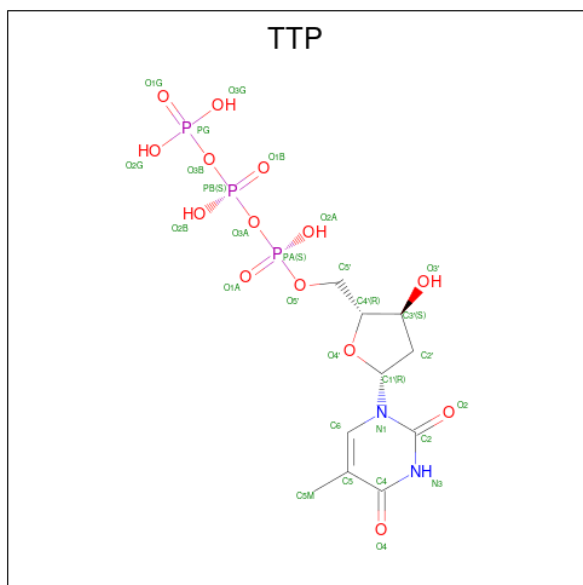
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	17	Total 338	C 163	N 62	O 97	P 16	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	J	18	357	172	65	103	17	0	0	0

- Molecule 4 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula:  $C_{10}H_{17}N_2O_{14}P_3$ ).

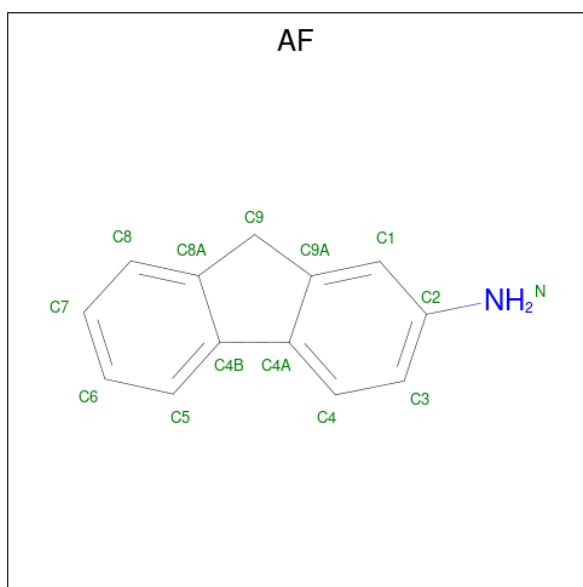


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	29	10	2	14	3	0	0
4	B	1	29	10	2	14	3	0	0

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

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

- Molecule 6 is 2-AMINOFUORENE (three-letter code: AF) (formula:  $C_{13}H_{11}N$ ).



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		

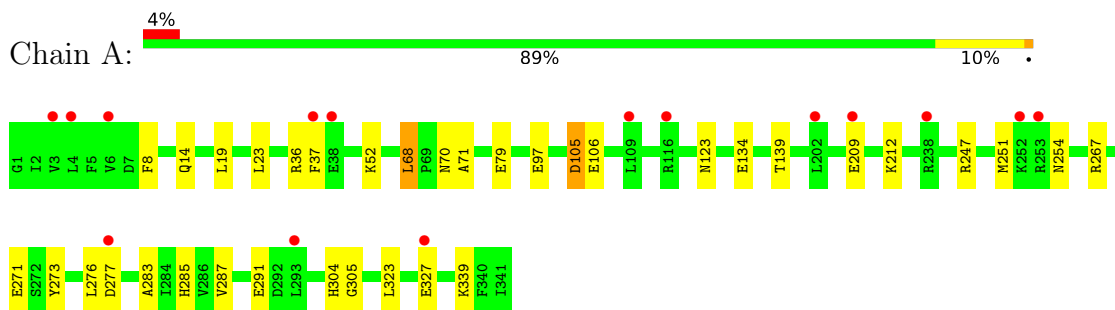
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	171	Total	O	0	0
			171	171		
7	D	13	Total	O	0	0
			13	13		
7	E	24	Total	O	0	0
			24	24		
7	B	147	Total	O	0	0
			147	147		
7	H	23	Total	O	0	0
			23	23		
7	J	24	Total	O	0	0
			24	24		

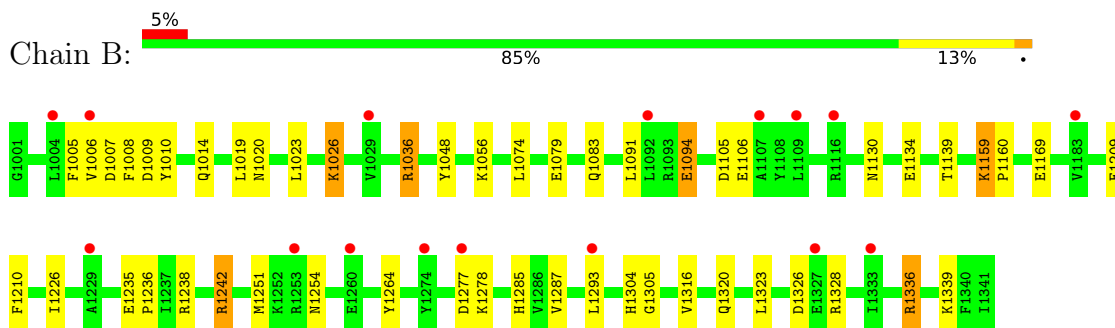
### 3 Residue-property plots [i](#)

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.

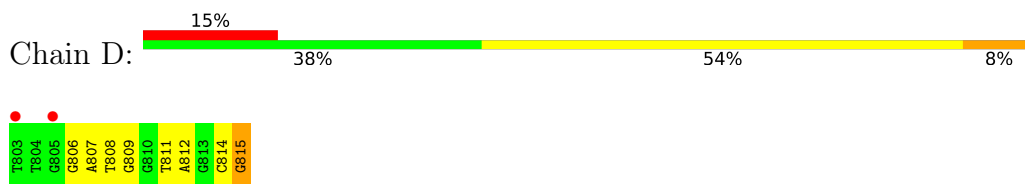
- Molecule 1: DNA polymerase IV



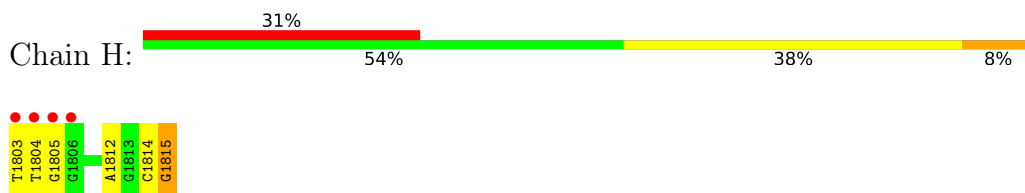
- Molecule 1: DNA polymerase IV



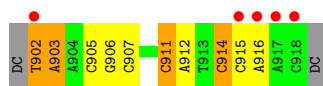
- Molecule 2: 5'-D(\*TP\*TP\*GP\*GP\*AP\*TP\*GP\*GP\*TP\*AP\*GP\*CP\*(DDG))-3'



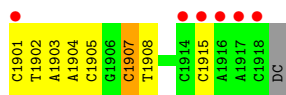
- Molecule 2: 5'-D(\*TP\*TP\*GP\*GP\*AP\*TP\*GP\*GP\*TP\*AP\*GP\*CP\*(DDG))-3'



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



- Molecule 3: 5'-D(\*CP\*C\*TP\*AP\*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	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.48Å 53.02Å 99.86Å 81.76° 76.76° 70.14°	Depositor
Resolution (Å)	20.00 – 2.01 19.72 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.9 (20.00-2.01) 95.9 (19.72-2.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.13 (at 2.01Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.206 , 0.246 0.214 , 0.255	Depositor DCC
$R_{free}$ test set	3185 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.0	Xtrriage
Anisotropy	0.188	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7237	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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, TTP, DDG

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	0.54	0/2788	0.62	0/3742
1	B	0.52	0/2801	0.62	0/3760
2	D	0.88	0/278	1.60	6/429 (1.4%)
2	H	0.86	0/278	1.59	5/429 (1.2%)
3	E	0.84	0/378	1.78	10/579 (1.7%)
3	J	0.93	0/399	1.60	5/611 (0.8%)
All	All	0.61	0/6922	0.94	26/9550 (0.3%)

There are no bond length outliers.

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	902	DT	O3'-P-O5'	-11.62	81.92	104.00
3	E	905	DC	O4'-C1'-N1	10.22	115.16	108.00
3	J	1901	DC	O4'-C1'-N1	9.19	114.44	108.00
2	H	1814	DC	O4'-C1'-N1	7.67	113.37	108.00
2	D	808	DT	O4'-C1'-N1	-7.66	102.64	108.00
3	E	903	DA	O5'-P-OP1	7.62	119.85	110.70
3	E	914	DC	O4'-C1'-N1	7.34	113.14	108.00
3	E	902	DT	OP1-P-O3'	-6.88	90.06	105.20
3	E	902	DT	OP2-P-O3'	-6.83	90.19	105.20
3	J	1905	DC	C1'-O4'-C4'	-6.72	103.38	110.10
2	D	811	DT	C4-C5-C7	6.56	122.94	119.00
2	D	814	DC	O4'-C1'-N1	6.51	112.56	108.00
2	H	1812	DA	O4'-C1'-N9	-6.50	103.45	108.00
2	H	1805	DG	P-O3'-C3'	6.33	127.30	119.70
3	E	903	DA	O5'-P-OP2	6.15	118.08	110.70
3	J	1907	DC	C1'-O4'-C4'	-6.14	103.96	110.10
3	E	907	DC	O4'-C4'-C3'	-6.09	102.06	104.50
3	E	906	DG	N3-C2-N2	-5.97	115.72	119.90

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	911	DC	O4'-C1'-N1	5.68	111.98	108.00
2	H	1804	DT	O4'-C4'-C3'	-5.65	102.24	104.50
2	D	811	DT	C6-C5-C7	-5.63	119.52	122.90
2	D	812	DA	O4'-C1'-N9	-5.59	104.09	108.00
3	J	1915	DC	O4'-C1'-N1	5.45	111.81	108.00
2	D	809	DG	O4'-C1'-N9	5.40	111.78	108.00
2	H	1803	DT	P-O3'-C3'	5.11	125.83	119.70
3	J	1904	DA	O4'-C1'-N9	-5.10	104.43	108.00

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

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	2749	0	2895	25	0
1	B	2761	0	2898	36	0
2	D	269	0	149	2	0
2	H	269	0	149	1	0
3	E	338	0	192	6	0
3	J	357	0	203	3	0
4	A	29	0	13	0	0
4	B	29	0	13	0	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
6	E	14	0	10	1	0
6	J	14	0	10	1	0
7	A	171	0	0	6	0
7	B	147	0	0	5	0
7	D	13	0	0	0	0
7	E	24	0	0	0	0
7	H	23	0	0	0	0
7	J	24	0	0	2	0
All	All	7237	0	6532	68	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 5.

All (68) 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:14:GLN:HE22	1:A:139:THR:H	1.08	1.01
1:B:1091:LEU:O	1:B:1094:GLU:HG2	1.64	0.97
1:B:1014:GLN:HE22	1:B:1139:THR:H	1.07	0.93
1:B:1336:ARG:HD3	7:B:229:HOH:O	1.75	0.87
1:A:123:ASN:HB3	7:A:533:HOH:O	1.79	0.82
3:E:902:DT:H2''	3:E:903:DA:OP2	1.82	0.80
1:A:283:ALA:HB2	1:A:339:LYS:HD2	1.65	0.76
1:B:1048:TYR:CE1	1:B:1160:PRO:HB3	2.22	0.74
3:E:915:DC:H2''	3:E:916:DA:C8	2.27	0.68
1:B:1130:ASN:O	1:B:1134:GLU:HG3	1.93	0.68
3:E:902:DT:C2''	3:E:903:DA:OP2	2.42	0.66
1:B:1014:GLN:NE2	1:B:1139:THR:H	1.89	0.65
1:A:14:GLN:NE2	1:A:139:THR:H	1.87	0.63
1:B:1079:GLU:CD	1:B:1079:GLU:H	2.02	0.63
1:B:1251:MET:HA	1:B:1264:TYR:CE1	2.37	0.59
1:A:70:ASN:O	1:A:70:ASN:OD1	2.20	0.59
1:A:134:GLU:HG2	7:A:622:HOH:O	2.03	0.58
1:B:1083:GLN:HG2	7:B:54:HOH:O	2.04	0.57
1:A:52[B]:LYS:HD2	7:A:551:HOH:O	2.04	0.57
1:A:267:ARG:HD2	7:A:388:HOH:O	2.05	0.57
1:B:1285:HIS:HD2	7:B:14:HOH:O	1.88	0.56
1:A:304:HIS:HD2	1:A:305:GLY:O	1.88	0.56
1:B:1106:GLU:OE1	2:H:1815:DDG:H5''	2.07	0.54
1:B:1210:PHE:CD1	1:B:1226:ILE:HD12	2.43	0.54
1:B:1210:PHE:HD1	1:B:1226:ILE:HD12	1.75	0.52
1:A:247:ARG:HG3	1:A:247:ARG:HH11	1.77	0.50
1:B:1036:ARG:NH2	1:B:1254:ASN:OD1	2.43	0.50
1:A:209:GLU:HG2	1:A:212:LYS:H	1.76	0.49
1:A:254:ASN:HD22	1:A:291:GLU:HG2	1.77	0.49
1:B:1023:LEU:O	1:B:1026:LYS:HG2	2.13	0.48
1:A:283:ALA:CB	1:A:339:LYS:HD2	2.38	0.48
1:B:1277:ASP:O	1:B:1278:LYS:HB2	2.13	0.48
1:A:209:GLU:HG3	1:B:1169:GLU:OE1	2.13	0.48
1:A:106:GLU:OE1	2:D:815:DDG:H5''	2.14	0.47
1:A:285:HIS:HD2	7:A:568:HOH:O	1.96	0.47
1:A:68:LEU:HD13	1:A:71:ALA:HB2	1.95	0.47
1:B:1304[B]:HIS:HD2	1:B:1305:GLY:O	1.97	0.47
1:B:1339:LYS:HD3	7:B:453:HOH:O	2.15	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:911:DC:H2''	3:E:912:DA:C8	2.49	0.47
3:J:1907:DC:H2''	3:J:1908:DT:H72	1.97	0.47
1:B:1159:LYS:HE2	7:B:25:HOH:O	2.15	0.46
1:A:287:VAL:HB	6:E:926:AF:C7	2.46	0.45
1:A:273:TYR:HA	1:A:276:LEU:HD12	1.98	0.45
1:B:1210:PHE:HD1	1:B:1226:ILE:CD1	2.30	0.45
3:J:1902:DT:H71	7:J:251:HOH:O	2.17	0.45
1:A:247:ARG:CZ	1:A:271:GLU:OE1	2.65	0.44
1:B:1316:VAL:O	1:B:1320:GLN:HG3	2.19	0.43
1:B:1020:ASN:HB3	1:B:1023:LEU:HD22	2.00	0.43
1:A:273:TYR:HE1	7:A:408:HOH:O	2.02	0.43
1:B:1326:ASP:OD1	1:B:1328:ARG:HB2	2.18	0.43
3:E:914:DC:H2''	3:E:915:DC:C5	2.54	0.42
1:B:1008:PHE:CD2	1:B:1105:ASP:HB2	2.54	0.42
1:A:36:ARG:HD3	1:A:251:MET:O	2.19	0.42
1:A:97:GLU:H	1:A:97:GLU:CD	2.23	0.42
1:B:1009:ASP:O	1:B:1010:TYR:C	2.58	0.42
1:B:1293:LEU:HD21	3:J:1903:DA:H2'	2.01	0.42
1:B:1242:ARG:HG3	7:J:50:HOH:O	2.20	0.42
1:B:1048:TYR:CE1	1:B:1160:PRO:CB	3.01	0.41
1:B:1023:LEU:HD21	1:B:1074:LEU:HD11	2.02	0.41
1:B:1048:TYR:HE1	1:B:1160:PRO:HB3	1.82	0.41
1:B:1006:VAL:O	1:B:1106:GLU:HA	2.20	0.41
1:B:1005:PHE:CE2	1:B:1007:ASP:HB2	2.56	0.41
1:B:1005:PHE:CE1	1:B:1106:GLU:HB3	2.56	0.41
1:B:1235:GLU:HA	1:B:1236:PRO:HD3	1.79	0.41
1:B:1287:VAL:HB	6:J:1926:AF:C7	2.51	0.40
2:D:806:DG:H2''	2:D:807:DA:C8	2.56	0.40
1:A:37:PHE:HZ	3:E:902:DT:H3'	1.86	0.40
1:A:8:PHE:CD2	1:A:105:ASP:HB2	2.56	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	340/341 (100%)	331 (97%)	8 (2%)	1 (0%)	41	37
1	B	341/341 (100%)	331 (97%)	10 (3%)	0	100	100
All	All	681/682 (100%)	662 (97%)	18 (3%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	277	ASP

### 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	300/299 (100%)	293 (98%)	7 (2%)	50	53
1	B	301/299 (101%)	290 (96%)	11 (4%)	34	32
All	All	601/598 (100%)	583 (97%)	18 (3%)	41	41

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	LEU
1	A	23	LEU
1	A	68	LEU
1	A	79	GLU
1	A	105	ASP
1	A	323	LEU
1	A	327	GLU
1	B	1019	LEU
1	B	1026	LYS
1	B	1036	ARG
1	B	1056	LYS
1	B	1094	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1159	LYS
1	B	1209	GLU
1	B	1238	ARG
1	B	1242	ARG
1	B	1323	LEU
1	B	1336	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	GLN
1	A	123	ASN
1	A	285	HIS
1	A	304	HIS
1	B	1014	GLN
1	B	1070	ASN
1	B	1285	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	DDG	D	815	3,2	17,23,24	1.30	3 (17%)	15,33,36	1.67	4 (26%)
2	DDG	H	1815	3,2	17,23,24	1.26	3 (17%)	15,33,36	1.52	4 (26%)

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	DDG	D	815	3,2	-	2/3/18/19	0/3/3/3
2	DDG	H	1815	3,2	-	2/3/18/19	0/3/3/3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	815	DDG	C5-C6	-3.15	1.41	1.47
2	H	1815	DDG	C5-C6	-2.45	1.42	1.47
2	H	1815	DDG	C5-C4	-2.37	1.37	1.43
2	H	1815	DDG	C8-N7	-2.26	1.31	1.35
2	D	815	DDG	C5-C4	-2.20	1.37	1.43
2	D	815	DDG	O5'-C5'	-2.16	1.39	1.44

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1815	DDG	O4'-C1'-C2'	3.11	110.04	106.67
2	D	815	DDG	C4'-O4'-C1'	-3.08	106.90	109.81
2	D	815	DDG	C2'-C1'-N9	2.85	117.86	112.48
2	D	815	DDG	C3'-C2'-C1'	-2.55	99.83	102.78
2	H	1815	DDG	O4'-C4'-C5'	-2.44	105.51	109.52
2	H	1815	DDG	O6-C6-N1	-2.25	117.99	120.65
2	D	815	DDG	O4'-C1'-C2'	2.20	109.05	106.67
2	H	1815	DDG	O6-C6-C5	2.02	128.32	124.37

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	H	1815	DDG	O4'-C4'-C5'-O5'
2	D	815	DDG	O4'-C4'-C5'-O5'
2	D	815	DDG	C3'-C4'-C5'-O5'
2	H	1815	DDG	C3'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	815	DDG	1	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	1815	DDG	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
4	TTP	B	1414	5	26,30,30	1.25	3 (11%)	39,47,47	1.94	10 (25%)
4	TTP	A	414	5	26,30,30	1.24	3 (11%)	39,47,47	1.80	9 (23%)
6	AF	E	926	3	14,16,16	1.02	2 (14%)	19,23,23	1.90	3 (15%)
6	AF	J	1926	3	14,16,16	0.98	1 (7%)	19,23,23	1.74	4 (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
4	TTP	B	1414	5	-	3/22/34/34	0/2/2/2
4	TTP	A	414	5	-	4/22/34/34	0/2/2/2
6	AF	E	926	3	-	-	0/3/3/3
6	AF	J	1926	3	-	-	0/3/3/3

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1414	TTP	C4-C5	3.05	1.49	1.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	414	TTP	C6-C5	2.92	1.39	1.34
4	B	1414	TTP	C4-N3	-2.87	1.33	1.38
4	B	1414	TTP	C6-C5	2.59	1.38	1.34
6	E	926	AF	C9-C8A	-2.40	1.51	1.54
4	A	414	TTP	C4-C5	2.32	1.48	1.44
4	A	414	TTP	C2-N1	2.24	1.42	1.38
6	E	926	AF	C9-C9A	-2.24	1.51	1.54
6	J	1926	AF	C9-C8A	-2.15	1.51	1.54

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1414	TTP	C4-N3-C2	-5.32	120.47	127.35
4	B	1414	TTP	C5-C4-N3	5.31	119.84	115.31
4	B	1414	TTP	N3-C2-N1	5.29	121.91	114.89
6	E	926	AF	C8A-C4B-C4A	4.57	109.69	107.01
6	J	1926	AF	C8A-C4B-C4A	4.54	109.67	107.01
4	A	414	TTP	C5-C4-N3	4.49	119.14	115.31
4	A	414	TTP	C4-N3-C2	-4.26	121.84	127.35
4	A	414	TTP	N3-C2-N1	4.02	120.23	114.89
4	A	414	TTP	O4-C4-C5	-3.84	120.45	124.90
6	E	926	AF	C9A-C4A-C4B	3.79	109.23	107.01
4	A	414	TTP	C5M-C5-C4	3.32	122.42	118.77
6	J	1926	AF	C9A-C9-C8A	3.22	105.53	102.85
4	B	1414	TTP	C5M-C5-C4	3.20	122.29	118.77
6	J	1926	AF	C9A-C4A-C4B	3.19	108.88	107.01
6	E	926	AF	C9A-C9-C8A	2.97	105.33	102.85
4	B	1414	TTP	O4-C4-C5	-2.95	121.49	124.90
4	B	1414	TTP	C5-C6-N1	-2.87	120.39	123.34
4	A	414	TTP	O3G-PG-O3B	2.57	113.27	104.64
4	A	414	TTP	C2'-C1'-N1	-2.54	107.91	113.77
4	B	1414	TTP	O2-C2-N1	-2.49	119.47	122.79
6	J	1926	AF	C9-C8A-C8	-2.37	131.44	134.47
4	A	414	TTP	O2-C2-N1	-2.22	119.83	122.79
4	B	1414	TTP	C2'-C1'-N1	-2.18	108.76	113.77
4	A	414	TTP	C5-C6-N1	-2.17	121.10	123.34
4	B	1414	TTP	C5M-C5-C6	-2.10	120.05	122.85
4	B	1414	TTP	O3G-PG-O3B	2.10	111.66	104.64

There are no chirality outliers.

All (7) torsion outliers are listed below:

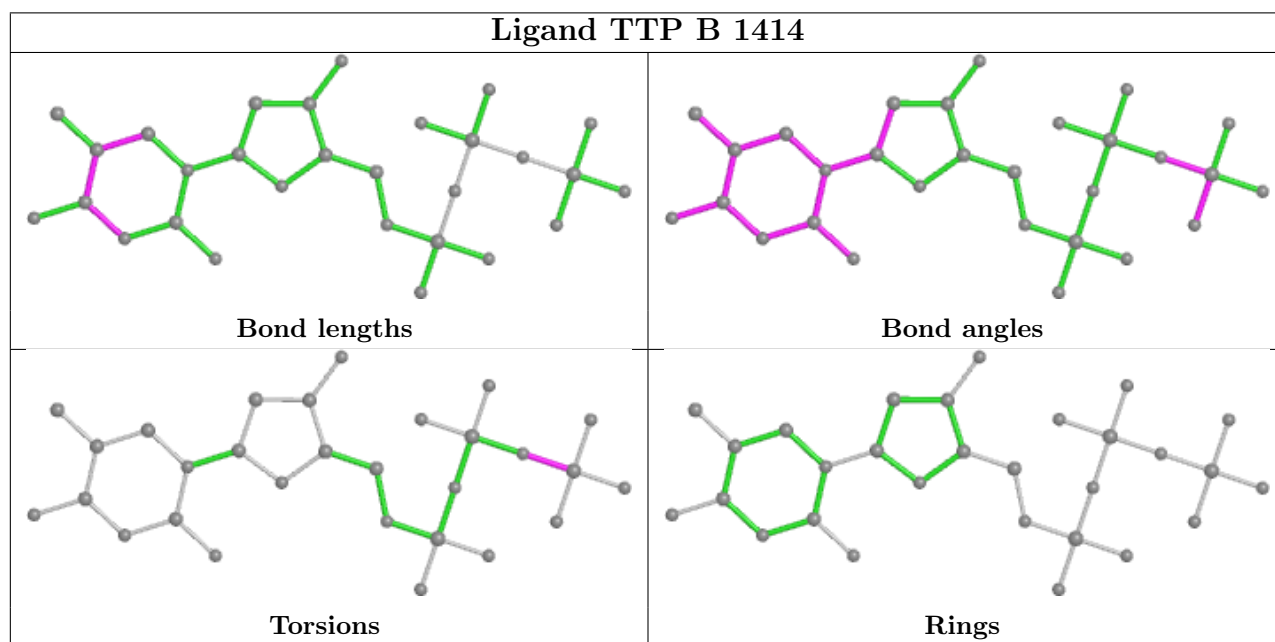
Mol	Chain	Res	Type	Atoms
4	A	414	TTP	PB-O3B-PG-O2G
4	A	414	TTP	PB-O3B-PG-O3G
4	B	1414	TTP	PB-O3B-PG-O2G
4	B	1414	TTP	PB-O3B-PG-O3G
4	A	414	TTP	PA-O3A-PB-O1B
4	A	414	TTP	PB-O3B-PG-O1G
4	B	1414	TTP	PB-O3B-PG-O1G

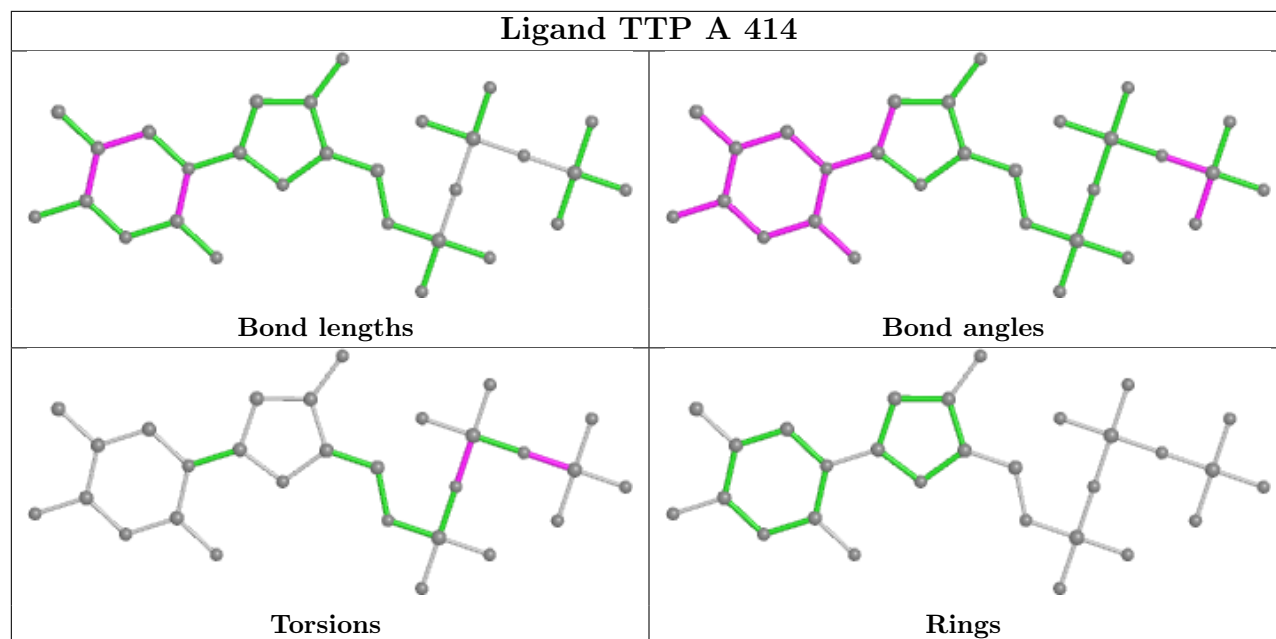
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	926	AF	1	0
6	J	1926	AF	1	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%)	0.37	15 (4%) 34 33	17, 30, 45, 52	0
1	B	341/341 (100%)	0.47	16 (4%) 31 30	18, 27, 42, 50	0
2	D	12/13 (92%)	0.52	2 (16%) 1 1	24, 30, 79, 85	0
2	H	12/13 (92%)	1.56	4 (33%) 0 0	22, 30, 98, 100	0
3	E	17/19 (89%)	0.97	5 (29%) 0 0	25, 31, 98, 99	0
3	J	18/19 (94%)	1.50	6 (33%) 0 0	21, 40, 97, 98	0
All	All	741/746 (99%)	0.48	48 (6%) 18 18	17, 29, 47, 100	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1803	DT	11.4
3	J	1918	DC	6.9
3	J	1901	DC	6.8
2	D	803	DT	6.5
3	E	917	DA	6.4
1	B	1327	GLU	5.4
3	E	918	DC	5.0
2	H	1804	DT	5.0
2	H	1805	DG	4.6
1	A	37	PHE	4.5
3	J	1916	DA	4.2
3	J	1917	DA	4.1
1	B	1116	ARG	3.9
1	A	327	GLU	3.6
1	B	1029	VAL	3.4
3	J	1914	DC	3.4
3	E	916	DA	3.4
3	J	1915	DC	3.2
1	A	116	ARG	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	3	VAL	3.1
1	B	1004	LEU	3.0
1	B	1274	TYR	2.8
1	A	4	LEU	2.8
2	H	1806	DG	2.7
1	B	1092	LEU	2.7
1	B	1253	ARG	2.7
1	A	202	LEU	2.6
1	A	293	LEU	2.6
1	A	238	ARG	2.6
1	B	1293	LEU	2.6
1	A	209	GLU	2.5
1	B	1229	ALA	2.5
1	A	252	LYS	2.4
1	A	6	VAL	2.4
1	B	1183	VAL	2.4
1	A	38	GLU	2.4
1	A	109	LEU	2.4
1	B	1109	LEU	2.4
1	A	253	ARG	2.4
1	B	1107	ALA	2.3
1	B	1006	VAL	2.2
1	B	1277	ASP	2.2
3	E	915	DC	2.2
1	B	1333	ILE	2.2
1	A	277	ASP	2.2
3	E	902	DT	2.1
1	B	1260	GLU	2.1
2	D	805	DG	2.1

## 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( $\text{\AA}^2$ )	Q<0.9
2	DDG	D	815	21/22	0.96	0.10	34,36,46,46	0
2	DDG	H	1815	21/22	0.96	0.09	30,33,39,41	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

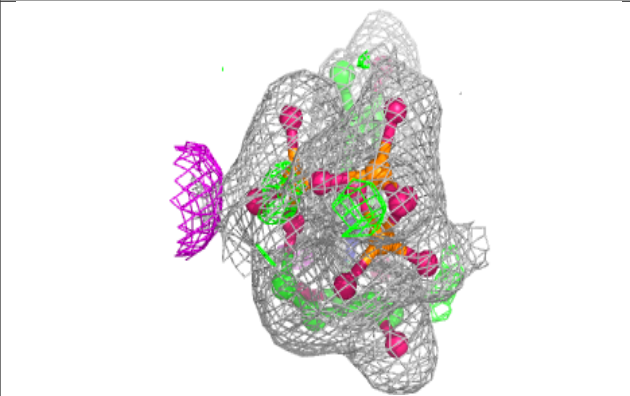
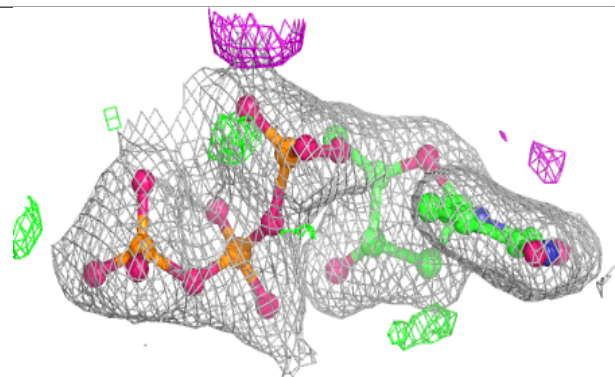
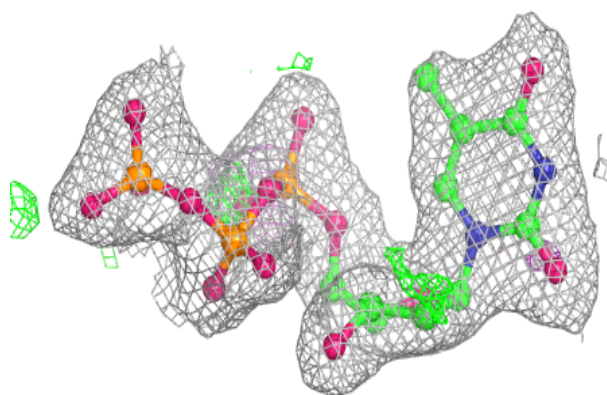
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	CA	B	1415	1/1	0.93	0.22	56,56,56,56	0
5	CA	A	415	1/1	0.94	0.43	48,48,48,48	0
6	AF	J	1926	14/14	0.95	0.13	30,34,37,37	0
6	AF	E	926	14/14	0.97	0.10	30,34,37,37	0
4	TTP	A	414	29/29	0.98	0.12	18,20,23,32	0
5	CA	B	1417	1/1	0.98	0.05	43,43,43,43	0
5	CA	B	1416	1/1	0.99	0.09	27,27,27,27	0
5	CA	A	416	1/1	0.99	0.09	28,28,28,28	0
5	CA	A	417	1/1	0.99	0.08	40,40,40,40	0
4	TTP	B	1414	29/29	0.99	0.11	14,18,21,22	0

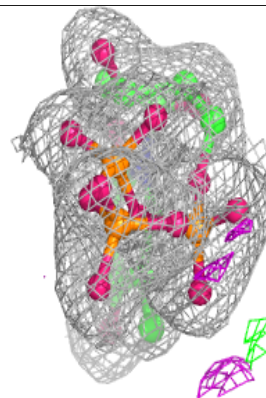
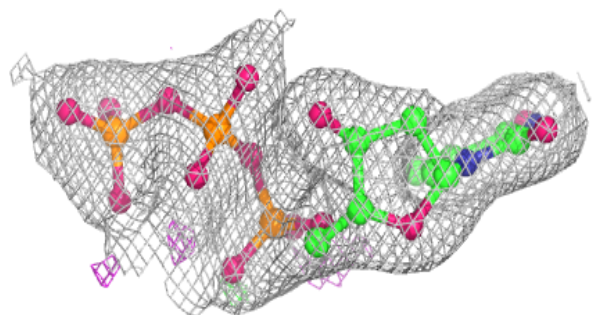
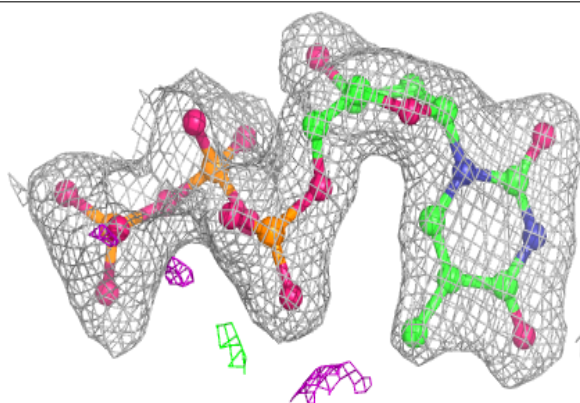
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around TTP 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 TTP 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.