WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Aug 28, 2023 – 12:38 PM EDT

PDB ID	:	3KHL
Title	:	Dpo4 post-extension ternary complex with misinserted A opposite the 2-aminofluorene-guanine [AF]G lesion
Authors	:	Rechkoblit, O.; Malinina, L.; Patel, D.J.
Deposited on	:	2009-10-30
Resolution	:	2.10 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*
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:

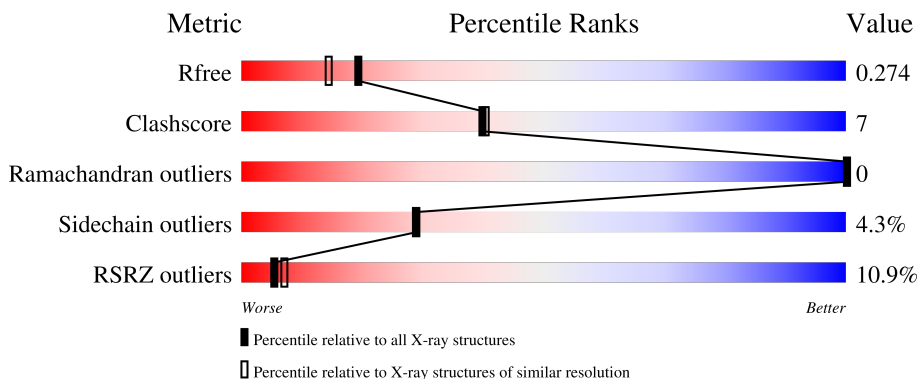|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.10 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.
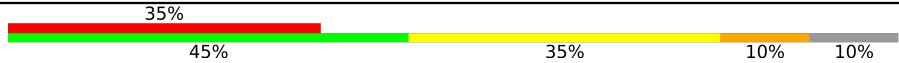
Metric	Percentile Ranks	Value
Rfree		0.274
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		4.3%
RSRZ outliers		10.9%

Worse ← → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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 <=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	9% / 85% / 13% ·
1	B	341	9% / 84% / 15% ·
2	D	13	38% / 38% / 38% / 23%
2	H	13	31% / 38% / 54% / 8%
3	E	20	20% / 25% / 40% / 20% / 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	20	

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 7145 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
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*AP*(DDG))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	13	Total 271	C 130	N 53	O 76	P 12	0	0	0
2	H	13	Total 261	C 125	N 48	O 76	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 3KHL
H	1815	DDG	G	engineered mutation	PDB 3KHL

- 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
3	E	17	Total 338	C 163	N 62	O 97	P 16	0	0	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

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

- Molecule 4 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: $C_{10}H_{17}N_2O_{14}P_3$).
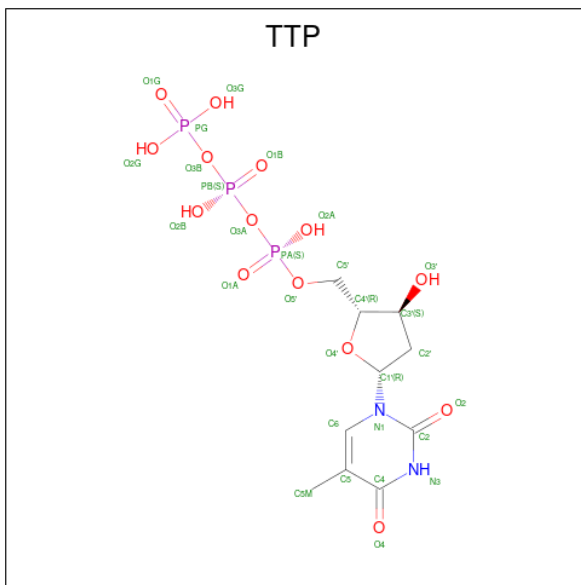
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 29	C 10	N 2	O 14	P 3	0	0
4	B	1	Total 29	C 10	N 2	O 14	P 3	0	0

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

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

- Molecule 6 is 2-AMINOFLUORENE (three-letter code: AF) (formula: $C_{13}H_{11}N$).
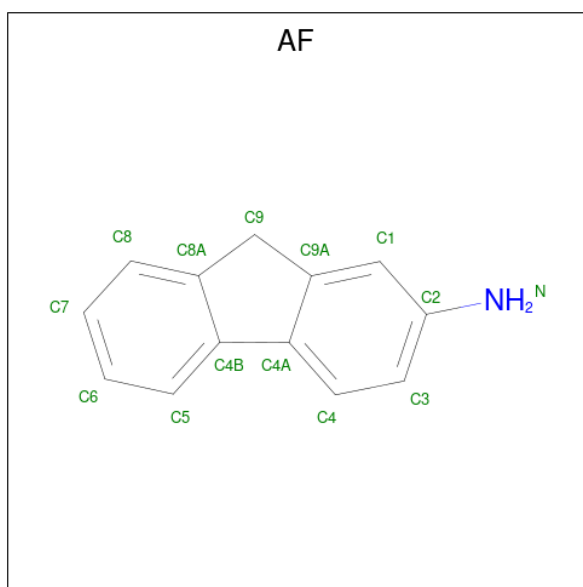
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total 14	C 13	N 1	0	0
6	J	1	Total 14	C 13	N 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	135	Total 135	O 135	0	0
7	D	15	Total 15	O 15	0	0
7	E	6	Total 6	O 6	0	0
7	B	126	Total 126	O 126	0	0
7	H	14	Total 14	O 14	0	0
7	J	20	Total 20	O 20	0	0

3 Residue-property plots ⓘ

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.

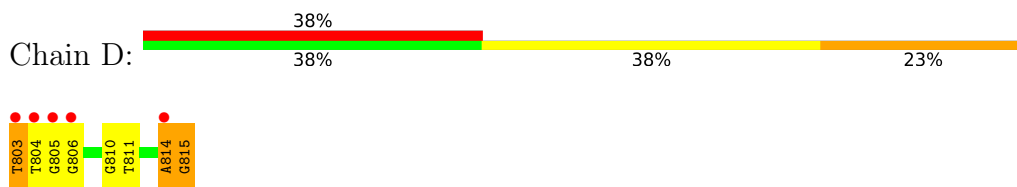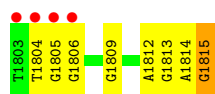- Molecule 1: DNA polymerase IV
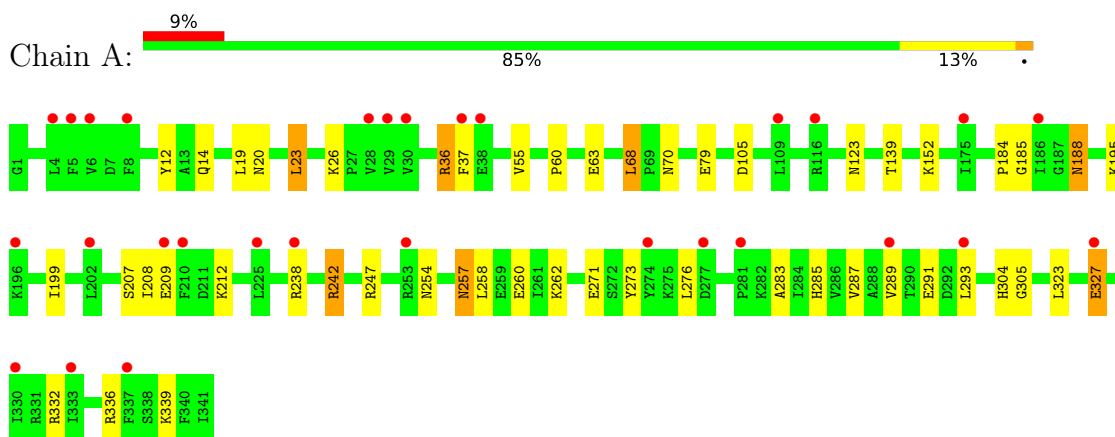
- Molecule 1: DNA polymerase IV
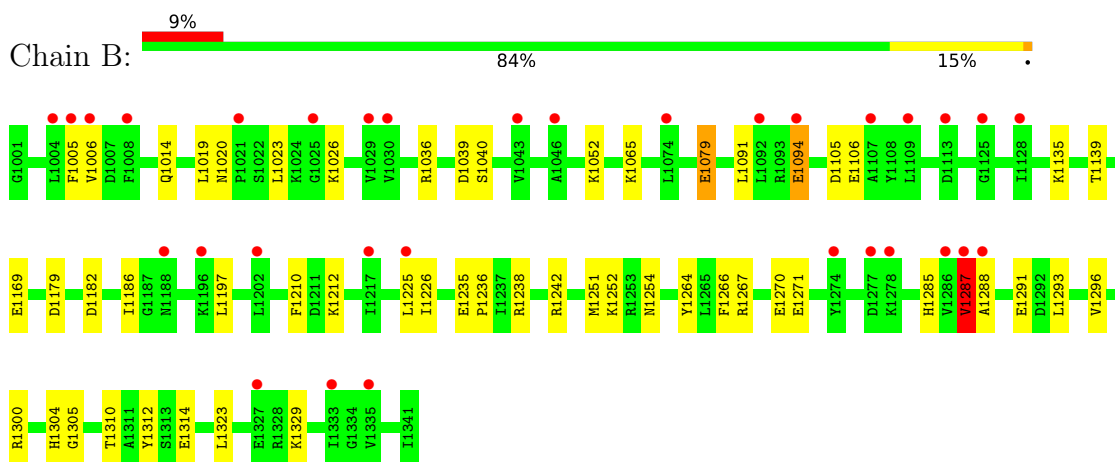
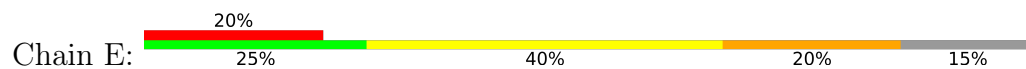
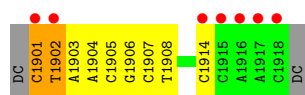- Molecule 2: 5'-D(*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*GP*AP*(DDG))-3'

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

Chain H:

• 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'

Chain E:

• 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'

Chain J:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	52.39Å　52.67Å　100.10Å 82.19°　76.40°　70.07°	Depositor
Resolution (Å)	20.00　–　2.10 19.76　–　2.10	Depositor EDS
% Data completeness (in resolution range)	97.2 (20.00-2.10) 97.2 (19.76-2.10)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.205　,　0.263 0.213　,　0.274	Depositor DCC
R_{free} test set	2797 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	41.3	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 52.0	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7145	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.38% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: AF, CA, 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	0.58	0/2788	0.66	0/3742
1	B	0.54	0/2801	0.65	1/3760 (0.0%)
2	D	0.93	0/281	1.53	4/434 (0.9%)
2	H	0.85	0/269	1.44	3/415 (0.7%)
3	E	0.87	0/378	1.78	14/579 (2.4%)
3	J	0.93	0/399	1.56	7/611 (1.1%)
All	All	0.64	0/6916	0.94	29/9541 (0.3%)

There are no bond length outliers.

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	902	DT	O3'-P-O5'	-11.03	83.04	104.00
3	E	916	DA	O4'-C1'-N9	7.94	113.56	108.00
3	E	902	DT	OP1-P-O3'	-7.85	87.92	105.20
3	J	1901	DC	O4'-C1'-N1	7.74	113.42	108.00
3	E	902	DT	P-O3'-C3'	7.67	128.91	119.70
3	E	903	DA	O5'-P-OP1	7.41	119.59	110.70
3	E	905	DC	C1'-O4'-C4'	-6.79	103.31	110.10
2	H	1813	DG	O4'-C1'-N9	6.59	112.61	108.00
3	E	906	DG	O4'-C1'-N9	-6.58	103.40	108.00
3	E	916	DA	C1'-O4'-C4'	-6.39	103.71	110.10
3	J	1905	DC	C1'-O4'-C4'	-6.38	103.72	110.10
1	B	1287	VAL	CB-CA-C	-6.37	99.30	111.40
3	E	914	DC	O4'-C1'-N1	6.15	112.31	108.00
2	D	803	DT	P-O3'-C3'	6.09	127.00	119.70
3	E	902	DT	OP2-P-O3'	-6.08	91.82	105.20
3	J	1904	DA	O4'-C1'-N9	-5.83	103.92	108.00
2	D	803	DT	O4'-C4'-C3'	-5.80	102.18	104.50
3	J	1906	DG	P-O3'-C3'	5.61	126.43	119.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	J	1901	DC	P-O3'-C3'	5.57	126.38	119.70
2	H	1809	DG	P-O3'-C3'	5.44	126.23	119.70
3	E	917	DA	O4'-C1'-N9	5.29	111.70	108.00
2	H	1812	DA	O4'-C1'-N9	-5.25	104.32	108.00
3	J	1901	DC	O4'-C1'-C2'	-5.21	101.73	105.90
2	D	811	DT	O4'-C1'-N1	-5.20	104.36	108.00
3	E	911	DC	O4'-C1'-N1	5.14	111.60	108.00
2	D	814	DA	O4'-C1'-N9	5.12	111.59	108.00
3	J	1902	DT	O4'-C1'-N1	5.10	111.57	108.00
3	E	916	DA	O4'-C1'-C2'	-5.09	101.83	105.90
3	E	903	DA	O5'-P-OP2	5.07	116.78	110.70

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	2749	0	2895	38	0
1	B	2761	0	2898	38	0
2	D	271	0	149	9	0
2	H	261	0	144	4	0
3	E	338	0	192	11	0
3	J	357	0	203	6	0
4	A	29	0	13	2	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	3	0
6	J	14	0	10	4	0
7	A	135	0	0	5	0
7	B	126	0	0	5	0
7	D	15	0	0	1	0
7	E	6	0	0	1	0
7	H	14	0	0	0	0
7	J	20	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7145	0	6527	97	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 (97) 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	1.04	0.95
1:B:1091:LEU:O	1:B:1094:GLU:HG2	1.69	0.93
3:E:902:DT:H2"	3:E:903:DA:OP2	1.66	0.92
1:B:1014:GLN:HE22	1:B:1139:THR:H	1.14	0.88
1:B:1296:VAL:HB	7:B:202:HOH:O	1.75	0.86
1:B:1293:LEU:HD21	3:J:1903:DA:H2'	1.68	0.76
1:B:1288:ALA:HB3	7:B:202:HOH:O	1.88	0.71
1:B:1251:MET:HA	1:B:1264:TYR:CE1	2.27	0.69
1:B:1287:VAL:HG13	6:J:1926:AF:C8	2.24	0.68
1:A:188:ASN:HD22	1:A:188:ASN:H	1.41	0.68
1:A:14:GLN:NE2	1:A:139:THR:H	1.86	0.68
1:B:1304[B]:HIS:HD2	1:B:1305:GLY:O	1.79	0.65
1:A:70:ASN:O	1:A:70:ASN:OD1	2.16	0.64
3:J:1914:DC:H6	7:J:95:HOH:O	1.81	0.63
1:B:1287:VAL:HG22	6:J:1926:AF:C5	2.28	0.62
1:B:1023:LEU:O	1:B:1026:LYS:HG2	2.00	0.62
1:A:60:PRO:HG2	1:A:63:GLU:HB2	1.81	0.61
1:A:304:HIS:HD2	1:A:305:GLY:O	1.83	0.60
1:A:285:HIS:HD2	7:A:521:HOH:O	1.85	0.59
1:A:283:ALA:HB2	1:A:339:LYS:HD2	1.83	0.58
1:A:123:ASN:HB3	7:A:549:HOH:O	2.04	0.58
1:A:209:GLU:HB3	7:A:555:HOH:O	2.05	0.57
3:E:907:DC:H2"	3:E:908:DT:H72	1.86	0.57
1:B:1179:ASP:HB3	1:B:1182:ASP:OD2	2.03	0.57
3:E:902:DT:C2'	3:E:903:DA:OP2	2.40	0.57
1:A:152:LYS:HE3	2:D:814:DA:H1'	1.85	0.56
1:A:273:TYR:HA	1:A:276:LEU:HD12	1.88	0.56
2:D:805:DG:H2"	2:D:806:DG:N7	2.20	0.56
2:D:804:DT:H1'	2:D:805:DG:C8	2.40	0.56
2:H:1805:DG:H2"	2:H:1806:DG:O4'	2.06	0.56
2:H:1814:DA:H2"	2:H:1815:DDG:H8	1.87	0.55
1:A:257:ASN:HB3	1:A:260:GLU:HB3	1.88	0.55
1:B:1052:LYS:HE2	7:B:225:HOH:O	2.04	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1291:GLU:HG3	1:B:1329:LYS:HB2	1.89	0.55
1:A:188:ASN:HD22	1:A:188:ASN:N	2.05	0.54
1:B:1020:ASN:HB3	1:B:1023:LEU:HD22	1.89	0.54
1:B:1285:HIS:HD2	7:B:9:HOH:O	1.91	0.53
1:B:1036:ARG:NH2	1:B:1254:ASN:OD1	2.33	0.53
1:A:247:ARG:NE	1:A:271:GLU:OE1	2.31	0.53
1:A:257:ASN:C	1:A:257:ASN:HD22	2.12	0.52
1:A:36:ARG:NH1	1:A:254:ASN:OD1	2.32	0.52
1:B:1186:ILE:HD11	1:B:1225:LEU:HD21	1.91	0.52
1:A:258:LEU:HG	1:A:262:LYS:HD2	1.92	0.51
1:B:1197:LEU:HD22	1:B:1212:LYS:HE3	1.93	0.51
1:B:1254:ASN:ND2	1:B:1291:GLU:HG2	2.26	0.50
1:B:1023:LEU:HA	1:B:1026:LYS:HE2	1.93	0.50
3:J:1901:DC:H2'	3:J:1902:DT:H6	1.76	0.50
7:A:595:HOH:O	2:D:810:DG:H2'	2.12	0.49
3:J:1907:DC:H2"	3:J:1908:DT:H72	1.93	0.49
3:E:915:DC:H2"	7:E:68:HOH:O	2.14	0.48
1:B:1287:VAL:HG13	6:J:1926:AF:C7	2.44	0.47
1:B:1079:GLU:CD	1:B:1079:GLU:H	2.17	0.47
1:A:12:TYR:CD2	4:A:414:TTP:H2'1	2.50	0.47
1:B:1293:LEU:CD2	3:J:1903:DA:H2'	2.42	0.46
1:A:20:ASN:HB3	1:A:23:LEU:HD22	1.96	0.46
1:B:1270:GLU:OE2	1:B:1312:TYR:OH	2.32	0.46
2:D:815:DDG:OP1	7:D:510:HOH:O	2.20	0.46
1:A:327:GLU:H	1:A:327:GLU:CD	2.18	0.46
1:A:293:LEU:HD11	3:E:903:DA:C8	2.51	0.45
1:B:1287:VAL:CG1	6:J:1926:AF:C8	2.94	0.45
1:B:1267:ARG:O	1:B:1271:GLU:HG3	2.16	0.45
1:A:185:GLY:HA2	2:D:814:DA:C2	2.52	0.45
1:B:1005:PHE:CE1	1:B:1106:GLU:HB3	2.52	0.44
2:D:803:DT:H2"	2:D:804:DT:OP2	2.17	0.44
3:E:911:DC:H2'	3:E:912:DA:C8	2.53	0.44
1:B:1006:VAL:O	1:B:1106:GLU:HA	2.18	0.44
1:A:209:GLU:HG2	1:A:212:LYS:H	1.83	0.43
1:A:55:VAL:HG21	1:A:68:LEU:HD12	2.01	0.43
1:B:1039:ASP:CG	1:B:1065:LYS:HE3	2.39	0.43
1:A:289:VAL:HB	1:A:332:ARG:HB2	2.01	0.43
1:A:199:ILE:HD11	1:A:208:ILE:HG21	2.01	0.43
2:D:805:DG:OP2	2:D:805:DG:H2'	2.18	0.43
1:B:1091:LEU:O	1:B:1094:GLU:CG	2.53	0.43
1:A:242:ARG:HD2	7:A:345:HOH:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:336:ARG:HB3	6:E:926:AF:C6	2.49	0.43
1:A:257:ASN:HD22	1:A:258:LEU:N	2.17	0.43
1:B:1266:PHE:CD1	1:B:1312:TYR:HD1	2.37	0.43
1:A:336:ARG:HB2	6:E:926:AF:C5	2.50	0.42
1:A:37:PHE:HZ	3:E:902:DT:H5"	1.84	0.42
1:B:1300:ARG:HD3	1:B:1314:GLU:OE1	2.20	0.42
1:B:1235:GLU:HA	1:B:1236:PRO:HD3	1.93	0.42
3:J:1901:DC:H2'	3:J:1902:DT:C6	2.55	0.42
3:E:914:DC:H2"	3:E:915:DC:C4	2.55	0.42
1:B:1310:THR:O	1:B:1314:GLU:HB2	2.20	0.42
2:H:1804:DT:H1'	2:H:1805:DG:N7	2.35	0.42
1:A:37:PHE:HZ	3:E:902:DT:H3'	1.84	0.41
1:A:184:PRO:HA	2:D:814:DA:H5'	2.03	0.41
1:B:1036:ARG:HB2	1:B:1040:SER:OG	2.20	0.41
1:A:12:TYR:CE2	4:A:414:TTP:H2'1	2.55	0.41
1:B:1169:GLU:HG2	7:B:240:HOH:O	2.20	0.41
1:B:1251:MET:HA	1:B:1264:TYR:CZ	2.55	0.41
1:B:1210:PHE:CD1	1:B:1226:ILE:HD12	2.56	0.41
3:E:914:DC:H2"	3:E:915:DC:C5	2.55	0.41
1:A:37:PHE:CZ	3:E:902:DT:H5"	2.57	0.40
2:H:1805:DG:H2'	2:H:1806:DG:C8	2.57	0.40
1:A:188:ASN:N	1:A:188:ASN:ND2	2.70	0.40
1:A:287:VAL:HB	6:E:926:AF:C7	2.50	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	340/341 (100%)	331 (97%)	9 (3%)	0	100	100
1	B	341/341 (100%)	331 (97%)	10 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	681/682 (100%)	662 (97%)	19 (3%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	300/299 (100%)	284 (95%)	16 (5%)	22 20
1	B	301/299 (101%)	291 (97%)	10 (3%)	38 40
All	All	601/598 (100%)	575 (96%)	26 (4%)	29 29

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	LEU
1	A	23	LEU
1	A	26	LYS
1	A	36	ARG
1	A	68	LEU
1	A	79	GLU
1	A	105	ASP
1	A	188	ASN
1	A	195	LYS
1	A	207	SER
1	A	238	ARG
1	A	242	ARG
1	A	257	ASN
1	A	291	GLU
1	A	323	LEU
1	A	327	GLU
1	B	1019	LEU
1	B	1079	GLU
1	B	1094	GLU
1	B	1105	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1135	LYS
1	B	1238	ARG
1	B	1242	ARG
1	B	1252	LYS
1	B	1287	VAL
1	B	1323	LEU

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

Mol	Chain	Res	Type
1	A	14	GLN
1	A	123	ASN
1	A	188	ASN
1	A	257	ASN
1	A	285	HIS
1	A	304	HIS
1	B	1014	GLN
1	B	1070	ASN
1	B	1188	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	H	1815	2,3,5	17,23,24	0.91	1 (5%)	15,33,36	1.42	2 (13%)
2	DDG	D	815	2,3	17,23,24	1.10	2 (11%)	15,33,36	1.31	3 (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
2	DDG	H	1815	2,3,5	-	0/3/18/19	0/3/3/3
2	DDG	D	815	2,3	-	0/3/18/19	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	815	DDG	C5-C6	-2.49	1.42	1.47
2	D	815	DDG	C5-C4	-2.05	1.37	1.43
2	H	1815	DDG	C5-C6	-2.04	1.43	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	1815	DDG	O6-C6-C5	2.69	129.62	124.37
2	D	815	DDG	O4'-C1'-C2'	2.63	109.52	106.67
2	D	815	DDG	O6-C6-C5	2.51	129.28	124.37
2	H	1815	DDG	C3'-C2'-C1'	2.18	105.30	102.78
2	D	815	DDG	C2'-C1'-N9	2.04	116.32	112.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	1815	DDG	1	0
2	D	815	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	A	414	5	26,30,30	1.31	3 (11%)	39,47,47	1.92	9 (23%)
6	AF	J	1926	3	14,16,16	1.22	2 (14%)	19,23,23	1.69	4 (21%)
6	AF	E	926	3	14,16,16	1.21	2 (14%)	19,23,23	1.67	4 (21%)
4	TTP	B	1414	5	26,30,30	1.32	4 (15%)	39,47,47	2.02	11 (28%)

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	A	414	5	-	4/22/34/34	0/2/2/2
6	AF	J	1926	3	-	-	0/3/3/3
6	AF	E	926	3	-	-	0/3/3/3
4	TTP	B	1414	5	-	4/22/34/34	0/2/2/2

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	414	TTP	C6-C5	3.42	1.40	1.34
4	B	1414	TTP	C6-C5	3.06	1.39	1.34
4	A	414	TTP	C4-C5	2.91	1.49	1.44
4	B	1414	TTP	C4-C5	2.75	1.49	1.44
6	E	926	AF	C9-C8A	-2.70	1.51	1.54
6	J	1926	AF	C9-C8A	-2.61	1.51	1.54
4	B	1414	TTP	O2-C2	2.54	1.27	1.23
6	E	926	AF	C9-C9A	-2.49	1.51	1.54
4	B	1414	TTP	C2-N1	2.31	1.42	1.38
6	J	1926	AF	C9-C9A	-2.25	1.51	1.54
4	A	414	TTP	C4-N3	-2.02	1.35	1.38

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	1414	TTP	C5-C4-N3	5.55	120.05	115.31
4	B	1414	TTP	C4-N3-C2	-5.12	120.72	127.35
4	A	414	TTP	N3-C2-N1	4.95	121.46	114.89
4	A	414	TTP	C4-N3-C2	-4.64	121.34	127.35
6	J	1926	AF	C8A-C4B-C4A	4.55	109.68	107.01
4	B	1414	TTP	O4-C4-C5	-4.43	119.77	124.90
6	E	926	AF	C8A-C4B-C4A	4.24	109.50	107.01
4	B	1414	TTP	N3-C2-N1	4.20	120.47	114.89
4	A	414	TTP	C5-C4-N3	4.11	118.82	115.31
4	B	1414	TTP	C5M-C5-C4	3.80	122.94	118.77
4	A	414	TTP	O4'-C1'-N1	3.46	114.04	107.86
4	A	414	TTP	C5-C6-N1	-3.46	119.78	123.34
6	J	1926	AF	C9A-C9-C8A	3.44	105.71	102.85
4	A	414	TTP	O4-C4-C5	-3.25	121.13	124.90
6	E	926	AF	C9A-C4A-C4B	3.19	108.88	107.01
6	E	926	AF	C9A-C9-C8A	3.05	105.39	102.85
4	A	414	TTP	C2'-C1'-N1	-3.01	106.83	113.77
4	A	414	TTP	O3G-PG-O3B	3.01	114.71	104.64
6	J	1926	AF	C9A-C4A-C4B	2.93	108.73	107.01
4	B	1414	TTP	C5-C6-N1	-2.90	120.36	123.34
4	B	1414	TTP	O2-C2-N1	-2.65	119.26	122.79
4	B	1414	TTP	C2'-C1'-N1	-2.58	107.83	113.77
4	A	414	TTP	O2-C2-N1	-2.26	119.78	122.79
6	E	926	AF	C9-C8A-C8	-2.25	131.59	134.47
6	J	1926	AF	C9-C8A-C8	-2.13	131.74	134.47
4	B	1414	TTP	PB-O3B-PG	-2.09	125.65	132.83
4	B	1414	TTP	C5M-C5-C6	-2.06	120.10	122.85
4	B	1414	TTP	O3G-PG-O3B	2.01	111.37	104.64

There are no chirality outliers.
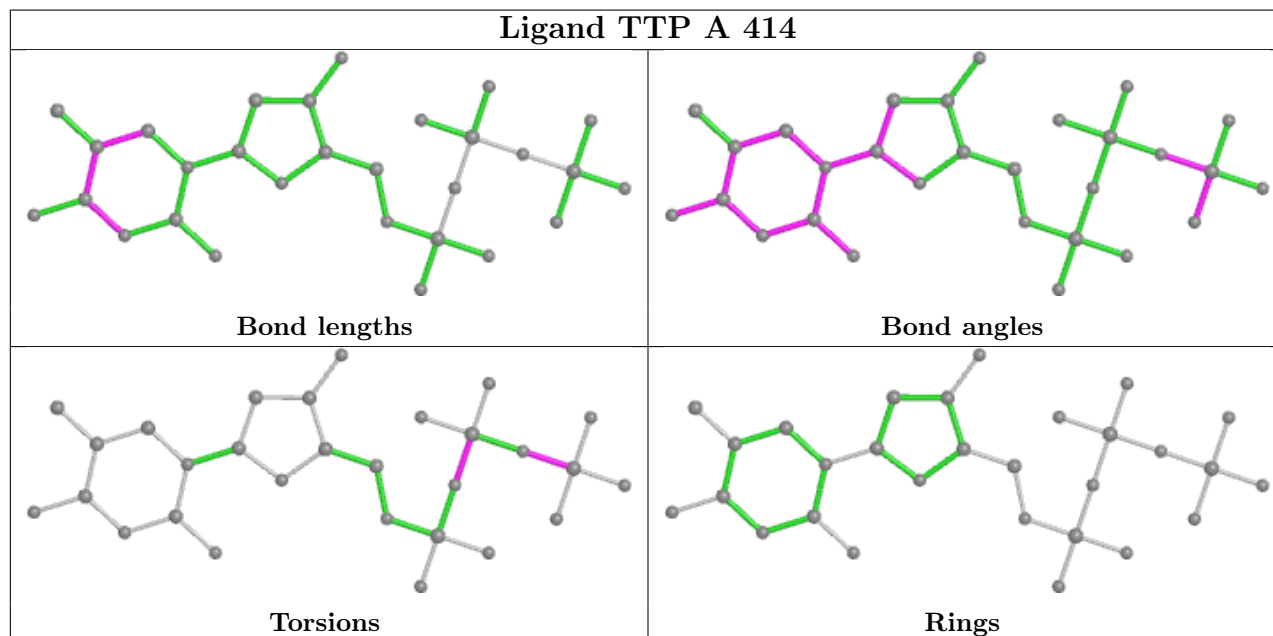
All (8) torsion outliers are listed below:

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

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	414	TTP	2	0
6	J	1926	AF	4	0
6	E	926	AF	3	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 then 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.

Ligand TTP A 414
Bond lengths   Bond angles   Torsions   Rings

Ligand TTP B 1414

Bond lengths

Bond angles

Torsions

Rings
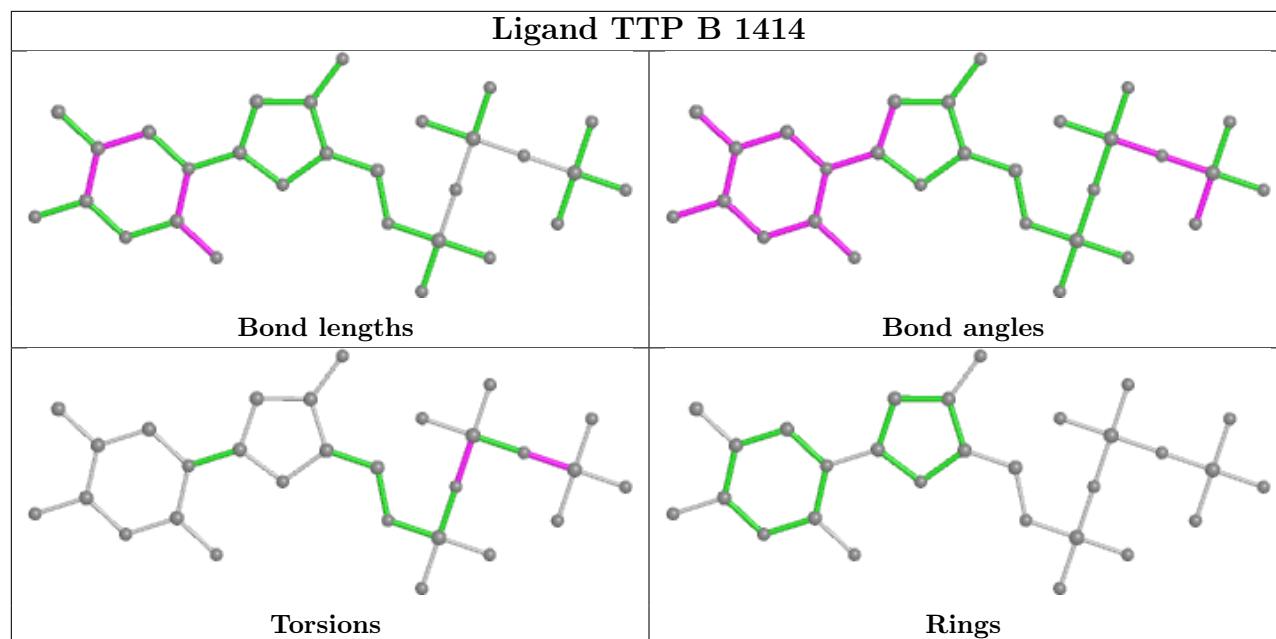
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	341/341 (100%)	0.57	29 (8%)	10	13	15, 29, 44, 52	0
1	B	341/341 (100%)	0.61	32 (9%)	8	11	18, 30, 43, 51	0
2	D	12/13 (92%)	1.52	5 (41%)	0	0	20, 29, 80, 82	0
2	H	12/13 (92%)	1.28	4 (33%)	0	0	16, 26, 86, 89	0
3	E	17/20 (85%)	1.56	4 (23%)	0	0	22, 33, 91, 95	0
3	J	18/20 (90%)	1.76	7 (38%)	0	0	19, 36, 94, 95	0
All	All	741/748 (99%)	0.67	81 (10%)	5	7	15, 30, 46, 95	0

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	803	DT	8.7
3	E	918	DC	7.8
2	H	1803	DT	7.4
3	J	1918	DC	7.3
3	J	1901	DC	7.1
3	E	916	DA	5.1
3	E	917	DA	4.8
3	E	915	DC	4.8
3	J	1916	DA	4.7
3	J	1917	DA	4.6
1	A	6	VAL	4.4
1	B	1274	TYR	4.1
3	J	1915	DC	3.8
1	B	1004	LEU	3.8
1	A	274	TYR	3.8
1	B	1029	VAL	3.8
2	D	805	DG	3.7
1	A	116	ARG	3.6
1	A	4	LEU	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	277	ASP	3.5
1	B	1021	PRO	3.3
1	A	293	LEU	3.3
1	B	1327	GLU	3.2
2	H	1805	DG	3.2
1	B	1109	LEU	3.1
2	H	1804	DT	3.1
1	A	202	LEU	3.1
1	A	5	PHE	2.9
1	A	333	ILE	2.8
1	A	37	PHE	2.8
1	A	38	GLU	2.8
1	B	1006	VAL	2.8
2	D	804	DT	2.8
2	D	806	DG	2.7
1	A	253	ARG	2.7
1	B	1030	VAL	2.6
1	B	1287	VAL	2.6
1	A	327	GLU	2.6
1	A	30	VAL	2.6
2	H	1806	DG	2.6
1	B	1196	LYS	2.6
1	B	1128	ILE	2.5
1	B	1333	ILE	2.5
3	J	1914	DC	2.5
1	A	225	LEU	2.5
1	A	28	VAL	2.5
1	B	1278	LYS	2.4
1	A	330	ILE	2.4
1	B	1074	LEU	2.4
1	A	8	PHE	2.4
1	B	1225	LEU	2.4
1	A	29	VAL	2.3
1	A	186	ILE	2.3
1	A	209	GLU	2.3
1	A	196	LYS	2.3
1	A	337	PHE	2.3
1	B	1202	LEU	2.3
1	B	1335	VAL	2.3
1	A	238	ARG	2.3
1	A	109	LEU	2.2
1	B	1107	ALA	2.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	1217	ILE	2.2
1	A	175	ILE	2.2
1	B	1092	LEU	2.2
1	A	281	PRO	2.2
1	B	1043	VAL	2.2
1	B	1094	GLU	2.2
1	B	1288	ALA	2.2
1	B	1008	PHE	2.1
1	B	1286	VAL	2.1
2	D	814	DA	2.1
3	J	1902	DT	2.1
1	B	1025	GLY	2.1
1	B	1188	ASN	2.1
1	A	210	PHE	2.0
1	B	1005	PHE	2.0
1	B	1125	GLY	2.0
1	B	1113	ASP	2.0
1	B	1046	ALA	2.0
1	A	289	VAL	2.0
1	B	1277	ASP	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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	DDG	D	815	21/22	0.86	0.19	49,50,64,65	0
2	DDG	H	1815	21/22	0.90	0.15	44,51,61,62	0

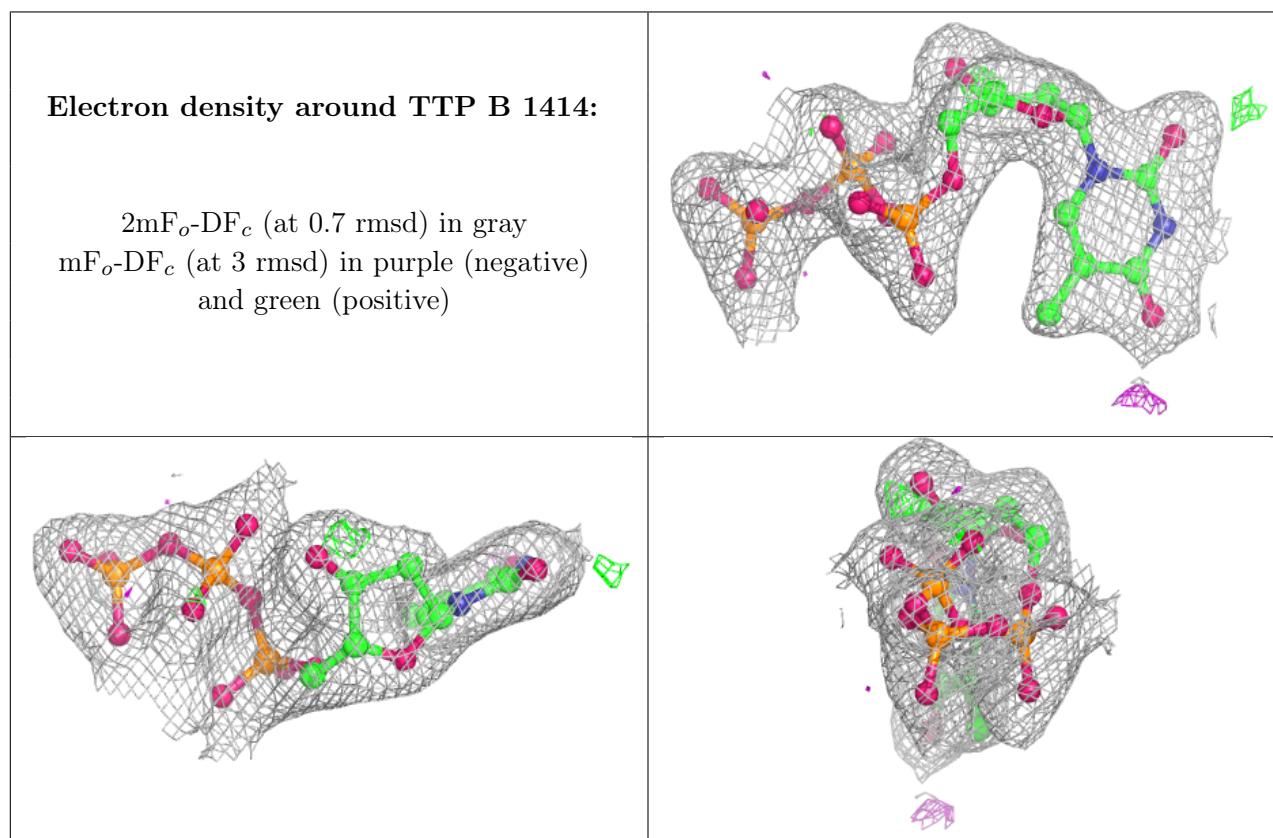## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.
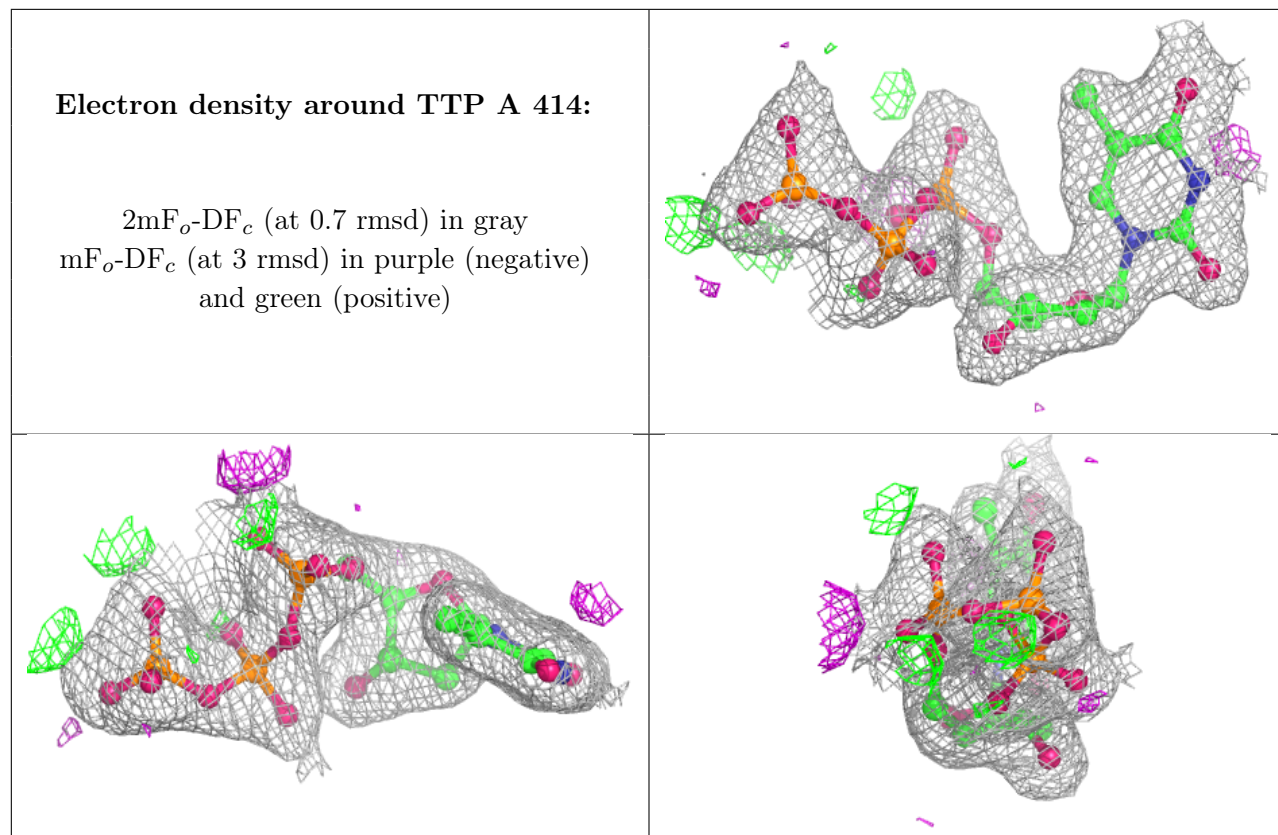
6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CA	B	1417	1/1	0.88	0.11	54,54,54,54	0
6	AF	J	1926	14/14	0.90	0.14	41,44,46,47	0
6	AF	E	926	14/14	0.91	0.13	40,44,45,45	0
5	CA	A	417	1/1	0.96	0.18	51,51,51,51	0
4	TTP	B	1414	29/29	0.96	0.12	7,13,15,18	0
4	TTP	A	414	29/29	0.97	0.12	8,13,18,30	0
5	CA	B	1415	1/1	0.98	0.07	54,54,54,54	0
5	CA	A	416	1/1	0.98	0.15	28,28,28,28	0
5	CA	B	1416	1/1	0.99	0.13	27,27,27,27	0
5	CA	A	415	1/1	0.99	0.15	44,44,44,44	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 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)

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)

6.5 Other polymers ⓘ

There are no such residues in this entry.