WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Aug 16, 2023 – 08:14 PM EDT

PDB ID	:	2ASD
Title	:	oxoG-modified Insertion Ternary Complex
Authors	:	Rechkoblit, O.; Malinina, L.; Cheng, Y.; Kuryavyi, V.; Broyde, S.; Geacintov, N.E.; Patel, D.J.
Deposited on	:	2005-08-23
Resolution	:	1.95 Å(reported)

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

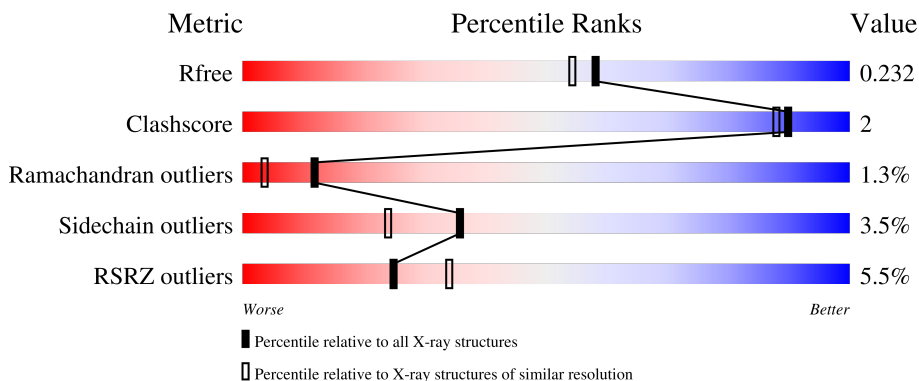|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 1.95 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.
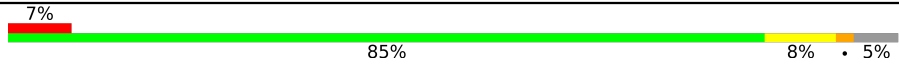
Metric	Percentile Ranks	Value
Rfree		0.232
Clashscore		2
Ramachandran outliers		1.3%
Sidechain outliers		3.5%
RSRZ outliers		5.5%

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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	D	13	85% 15%
1	H	13	8% 69% 23% 8%
2	E	19	11% 74% 16% 11%
2	J	19	16% 63% 32% 5%
3	A	360	3% 89% 5% · 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	360	

7%

85%   8%   ·   5%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 7301 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a DNA chain called 5'-D(*GP*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*A P*(DDG))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	13	Total	C	N	O	P	0	0	0
			272	130	53	77	12			
1	H	13	Total	C	N	O	P	0	0	0
			272	130	53	77	12			

- Molecule 2 is a DNA chain called 5'-D(*CP*T*AP*AP*CP*(8OG)P*CP*TP*AP*CP*CP* AP*TP*CP*CP*AP*AP*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	17	Total	C	N	O	P	0	0	0
			338	162	63	97	16			
2	J	19	Total	C	N	O	P	0	0	0
			377	181	68	110	18			

- Molecule 3 is a protein called DNA polymerase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	341	Total	C	N	O	S	0	0	0
			2739	1757	472	504	6			
3	B	341	Total	C	N	O	S	0	0	0
			2739	1757	472	504	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	GLY	-	cloning artifact	UNP Q97W02
A	-6	SER	-	cloning artifact	UNP Q97W02
A	-5	HIS	-	cloning artifact	UNP Q97W02
A	-4	MET	-	cloning artifact	UNP Q97W02
A	-3	GLY	-	cloning artifact	UNP Q97W02
A	-2	GLY	-	cloning artifact	UNP Q97W02

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	cloning artifact	UNP Q97W02
A	0	GLY	-	cloning artifact	UNP Q97W02
A	1	GLY	-	cloning artifact	UNP Q97W02
B	-7	GLY	-	cloning artifact	UNP Q97W02
B	-6	SER	-	cloning artifact	UNP Q97W02
B	-5	HIS	-	cloning artifact	UNP Q97W02
B	-4	MET	-	cloning artifact	UNP Q97W02
B	-3	GLY	-	cloning artifact	UNP Q97W02
B	-2	GLY	-	cloning artifact	UNP Q97W02
B	-1	GLY	-	cloning artifact	UNP Q97W02
B	0	GLY	-	cloning artifact	UNP Q97W02
B	1001	GLY	-	cloning artifact	UNP Q97W02

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	1	Total 1	Ca 1	0	0
4	A	3	Total 3	Ca 3	0	0
4	B	3	Total 3	Ca 3	0	0

- Molecule 5 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: $C_9H_{16}N_3O_{13}P_3$).
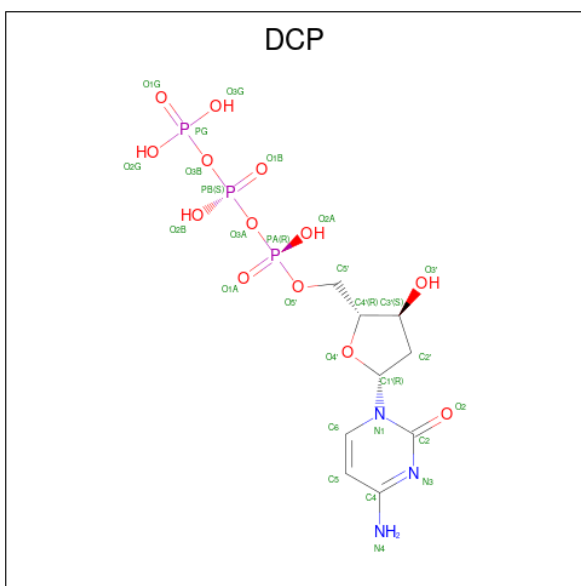
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total 28	C 9	N 3	O 13	P 3	0	0
5	B	1	Total 28	C 9	N 3	O 13	P 3	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	42	Total 42	O 42	1	0
6	E	37	Total 37	O 37	0	0
6	H	21	Total 21	O 21	0	0
6	J	26	Total 26	O 26	0	0
6	A	211	Total 211	O 211	2	0
6	B	164	Total 164	O 164	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: 5'-D(*GP*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*(DDG))-3'

Chain D:
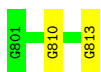
- Molecule 1: 5'-D(*GP*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*(DDG))-3'

Chain H:
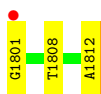
- Molecule 2: 5'-D(*CP*T*AP*AP*CP*(8OG)P*CP*TP*AP*CP*CP*AP*TP*CP*CP*AP*AP*CP*C)-3'

Chain E:
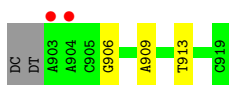
- Molecule 2: 5'-D(*CP*T*AP*AP*CP*(8OG)P*CP*TP*AP*CP*CP*AP*TP*CP*CP*AP*AP*CP*C)-3'

Chain J:

- Molecule 3: DNA polymerase IV

Chain A:

S297 L323 E327 I341 GLU ALA ILE GLY LEU ASP LYS PHE PHE ASP THR
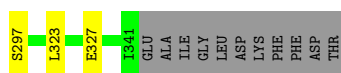
• Molecule 3: DNA polymerase IV

Chain B:

7%

85%   8%   • 5%

GLY SER HIS MET GLY GLY GLY G1001 Y1010 Q1014 Y1048 N1070 M1076 E1097 R1116 D1117 E1134 T1139 K1146 N1147 D1156 K1159 P1160 K1164 V1165 I1166 E1169 E1170 V1171 K1172 R1176 E1177 D1182 N1188 E1192 K1195 K1196 L1197 G1198 I1199

D1204 S1207 I1208 E1209 F1210 D1211 K1212 G1215 M1216 D1231 N1234 R1238 R1242 K1262 P1263 D1277 K1278 A1288 D1292 L1293 D1294 L1318 L1323 D1326 E1327 R1336 I1341 GLU ALA ILE GLY LEU ASP LYS PHE PHE ASP THR

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	75.95Å 101.11Å 84.12Å 90.00° 97.14° 90.00°	Depositor
Resolution (Å)	20.00 – 1.95 32.19 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (20.00-1.95) 99.0 (32.19-1.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.53 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.225 , 0.253 0.232 , 0.232	Depositor DCC
R_{free} test set	4954 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	26.3	Xtriage
Anisotropy	0.494	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7301	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.00 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.5173e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: DCP, 8OG, CA, 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	D	0.68	0/282	1.24	0/436
1	H	0.82	0/282	1.37	1/436 (0.2%)
2	E	0.65	0/351	1.26	3/534 (0.6%)
2	J	1.12	4/394 (1.0%)	1.45	2/600 (0.3%)
3	A	0.34	0/2778	0.69	6/3731 (0.2%)
3	B	0.33	0/2778	0.66	9/3731 (0.2%)
All	All	0.48	4/6865 (0.1%)	0.86	21/9468 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	A	0	1
3	B	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	1919	DC	C4-C5	7.31	1.48	1.43
2	J	1916	DA	N9-C4	5.67	1.41	1.37
2	J	1919	DC	N1-C6	5.61	1.40	1.37
2	J	1919	DC	N1-C2	5.43	1.45	1.40

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	1909	DA	O4'-C1'-N9	6.53	112.57	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	277	ASP	CB-CG-OD2	6.20	123.88	118.30
3	A	117	ASP	CB-CG-OD2	6.17	123.85	118.30
2	E	913	DT	O4'-C1'-N1	-5.95	103.84	108.00
3	A	292	ASP	CB-CG-OD2	5.85	123.56	118.30
3	A	182	ASP	CB-CG-OD2	5.82	123.54	118.30
2	E	909	DA	O4'-C1'-N9	5.76	112.03	108.00
2	J	1901	DC	O4'-C1'-N1	5.59	111.92	108.00
3	B	1277	ASP	CB-CG-OD2	5.56	123.30	118.30
3	B	1204	ASP	CB-CG-OD2	5.52	123.27	118.30
3	B	1156	ASP	CB-CG-OD2	5.39	123.15	118.30
3	A	7	ASP	CB-CG-OD2	5.31	123.08	118.30
3	A	204	ASP	CB-CG-OD2	5.25	123.02	118.30
3	B	1117	ASP	CB-CG-OD2	5.23	123.01	118.30
3	B	1211	ASP	CB-CG-OD2	5.18	122.96	118.30
3	B	1294	ASP	CB-CG-OD2	5.12	122.91	118.30
3	B	1182	ASP	CB-CG-OD2	5.11	122.90	118.30
2	E	909	DA	C1'-O4'-C4'	-5.09	105.01	110.10
1	H	1801	DG	C4-C5-N7	-5.08	108.77	110.80
3	B	1292	ASP	CB-CG-OD2	5.05	122.85	118.30
3	B	1326	ASP	CB-CG-OD2	5.04	122.84	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	159	LYS	Peptide
3	B	1159	LYS	Peptide

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	D	272	0	149	1	0
1	H	272	0	149	2	0
2	E	338	0	189	0	0
2	J	377	0	212	2	0
3	A	2739	0	2883	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	2739	0	2880	12	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
4	D	1	0	0	0	0
5	A	28	0	12	0	0
5	B	28	0	12	0	0
6	A	211	0	0	0	0
6	B	164	0	0	1	0
6	D	42	0	0	1	0
6	E	37	0	0	0	0
6	H	21	0	0	1	0
6	J	26	0	0	0	0
All	All	7301	0	6486	24	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 2.

All (24) 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:D:810:DG:OP2	6:D:713:HOH:O	1.80	0.97
3:A:14:GLN:HE22	3:A:139:THR:H	1.12	0.90
3:B:1014:GLN:HE22	3:B:1139:THR:H	1.23	0.82
3:B:1048:TYR:CZ	3:B:1160:PRO:HD3	2.26	0.70
3:B:1277:ASP:O	6:B:1778:HOH:O	2.15	0.64
3:B:1048:TYR:OH	3:B:1160:PRO:HD3	1.99	0.62
3:B:1166:ILE:HG23	3:B:1170:GLU:HB3	1.85	0.59
3:B:1234:ASN:HD22	3:B:1234:ASN:N	2.03	0.57
1:H:1808:DT:OP2	6:H:1789:HOH:O	2.17	0.54
3:B:1147:ASN:HD22	3:B:1147:ASN:C	2.11	0.54
3:B:1048:TYR:CZ	3:B:1160:PRO:CD	2.91	0.53
3:A:48:TYR:OH	3:A:160:PRO:HD3	2.10	0.52
3:B:1014:GLN:NE2	3:B:1139:THR:H	2.00	0.48
3:A:48:TYR:CZ	3:A:160:PRO:HD3	2.49	0.48
3:A:159:LYS:HB3	3:A:160:PRO:CD	2.46	0.46
3:A:79:GLU:H	3:A:79:GLU:CD	2.19	0.46
3:A:14:GLN:NE2	3:A:139:THR:H	1.95	0.45
2:J:1915:DC:C4	2:J:1916:DA:N6	2.84	0.45
3:B:1199:ILE:HD11	3:B:1208:ILE:HG21	2.00	0.43
2:J:1914:DC:C4	2:J:1915:DC:N4	2.87	0.42
3:B:1146:LYS:HE2	3:B:1231:ASP:OD1	2.19	0.42

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:1262:LYS:N	3:B:1263:PRO:HD2	2.34	0.42
3:A:49:GLU:HA	3:A:52:LYS:HE2	2.01	0.42
1:H:1812:DA:H2"	1:H:1813:DDG:H8	2.02	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	339/360 (94%)	328 (97%)	7 (2%)	4 (1%)	13	4
3	B	339/360 (94%)	328 (97%)	6 (2%)	5 (2%)	10	3
All	All	678/720 (94%)	656 (97%)	13 (2%)	9 (1%)	12	3

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	160	PRO
3	B	1160	PRO
3	A	10	TYR
3	B	1010	TYR
3	B	1277	ASP
3	B	1278	LYS
3	A	277	ASP
3	A	159	LYS
3	B	1159	LYS

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	A	299/311 (96%)	293 (98%)	6 (2%)	55	48
3	B	299/311 (96%)	284 (95%)	15 (5%)	24	11
All	All	598/622 (96%)	577 (96%)	21 (4%)	36	24

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	195	LYS
3	A	238	ARG
3	A	252	LYS
3	A	297	SER
3	A	323	LEU
3	A	327	GLU
3	B	1070	ASN
3	B	1076	MET
3	B	1116	ARG
3	B	1146	LYS
3	B	1147	ASN
3	B	1164	LYS
3	B	1176	ARG
3	B	1195	LYS
3	B	1234	ASN
3	B	1242	ARG
3	B	1262	LYS
3	B	1294	ASP
3	B	1318	LEU
3	B	1323	LEU
3	B	1336	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	14	GLN
3	A	285	HIS
3	B	1014	GLN
3	B	1083	GLN
3	B	1147	ASN
3	B	1234	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
1	DDG	H	1813	1,2,4	17,23,24	1.22	3 (17%)	15,33,36	1.10	1 (6%)
2	8OG	E	906	2	22,25,26	3.88	9 (40%)	30,37,40	3.95	13 (43%)
1	DDG	D	813	1,2,4	17,23,24	1.17	3 (17%)	15,33,36	0.95	0
2	8OG	J	1906	2	22,25,26	3.89	10 (45%)	30,37,40	3.88	13 (43%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	DDG	H	1813	1,2,4	-	0/3/18/19	0/3/3/3
2	8OG	E	906	2	-	2/7/21/22	0/3/3/3
1	DDG	D	813	1,2,4	-	0/3/18/19	0/3/3/3
2	8OG	J	1906	2	-	2/7/21/22	0/3/3/3

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	906	8OG	O3'-C3'	-9.14	1.24	1.43
2	J	1906	8OG	O3'-C3'	-8.94	1.24	1.43
2	J	1906	8OG	C1'-N9	-7.92	1.34	1.47
2	E	906	8OG	C5-C4	7.87	1.48	1.37
2	E	906	8OG	C1'-N9	-7.83	1.34	1.47
2	J	1906	8OG	C3'-C4'	-7.53	1.32	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	906	8OG	C3'-C4'	-7.49	1.32	1.53
2	J	1906	8OG	C5-C4	7.46	1.48	1.37
2	J	1906	8OG	C8-N9	-4.82	1.32	1.40
2	E	906	8OG	C8-N9	-4.80	1.32	1.40
2	J	1906	8OG	C4-N9	-3.67	1.32	1.39
2	E	906	8OG	C4-N9	-3.47	1.32	1.39
2	J	1906	8OG	O8-C8	3.36	1.29	1.23
2	E	906	8OG	O8-C8	2.84	1.28	1.23
1	H	1813	DDG	C5-C6	-2.79	1.41	1.47
2	J	1906	8OG	C6-N1	-2.77	1.33	1.38
1	D	813	DDG	C5-C6	-2.60	1.42	1.47
2	J	1906	8OG	C8-N7	-2.45	1.33	1.38
2	E	906	8OG	C6-N1	-2.42	1.34	1.38
1	D	813	DDG	C8-N7	-2.38	1.31	1.35
1	H	1813	DDG	C8-N7	-2.35	1.31	1.35
1	H	1813	DDG	C5-C4	-2.25	1.37	1.43
1	D	813	DDG	C5-C4	-2.18	1.37	1.43
2	J	1906	8OG	C5-C6	2.15	1.47	1.42
2	E	906	8OG	C8-N7	-2.03	1.34	1.38

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	1906	8OG	C2'-C1'-N9	12.37	130.66	116.01
2	E	906	8OG	C2'-C1'-N9	12.29	130.56	116.01
2	J	1906	8OG	N9-C4-N3	9.87	137.09	125.81
2	E	906	8OG	N9-C4-N3	9.84	137.06	125.81
2	J	1906	8OG	C5-C4-N3	-7.29	120.28	127.80
2	E	906	8OG	C5-C4-N3	-7.05	120.52	127.80
2	E	906	8OG	C4'-O4'-C1'	-5.04	97.27	109.45
2	E	906	8OG	N7-C8-N9	4.75	112.22	106.58
2	E	906	8OG	O4'-C1'-N9	4.66	112.97	108.29
2	E	906	8OG	C2-N3-C4	4.63	120.54	112.30
2	J	1906	8OG	C2-N3-C4	4.58	120.47	112.30
2	J	1906	8OG	N7-C8-N9	4.50	111.92	106.58
2	J	1906	8OG	C4'-O4'-C1'	-4.47	98.65	109.45
2	J	1906	8OG	O4'-C1'-N9	3.71	112.02	108.29
2	E	906	8OG	C5-N7-C8	-3.15	104.94	109.47
2	E	906	8OG	O6-C6-C5	-3.14	120.04	127.24
2	E	906	8OG	O4'-C1'-C2'	3.07	112.05	106.25
2	J	1906	8OG	O4'-C1'-C2'	2.89	111.72	106.25
2	J	1906	8OG	C5-N7-C8	-2.83	105.39	109.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	1906	8OG	O6-C6-C5	-2.82	120.77	127.24
2	E	906	8OG	O3'-C3'-C2'	2.59	120.17	110.90
2	J	1906	8OG	O3'-C3'-C2'	2.52	119.93	110.90
2	E	906	8OG	C3'-C2'-C1'	-2.37	96.60	102.54
2	J	1906	8OG	C3'-C2'-C1'	-2.32	96.72	102.54
1	H	1813	DDG	C2'-C1'-N9	2.16	116.55	112.48
2	E	906	8OG	O4'-C4'-C5'	2.08	116.22	109.37
2	J	1906	8OG	C5'-C4'-C3'	2.04	126.71	114.74

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	906	8OG	C2'-C1'-N9-C8
2	J	1906	8OG	C2'-C1'-N9-C8
2	E	906	8OG	O4'-C4'-C5'-O5'
2	J	1906	8OG	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	H	1813	DDG	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z|$ > 2	Counts	RMSZ	#$|Z|$ > 2
5	DCP	A	414	4	25,29,29	1.02	1 (4%)	37,45,45	1.04	3 (8%)
5	DCP	B	1414	4	25,29,29	1.10	3 (12%)	37,45,45	1.01	2 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DCP	A	414	4	-	5/22/34/34	0/2/2/2
5	DCP	B	1414	4	-	3/22/34/34	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	414	DCP	C4-N4	2.34	1.39	1.33
5	B	1414	DCP	PA-O5'	-2.11	1.50	1.59
5	B	1414	DCP	C2-N1	-2.10	1.35	1.40
5	B	1414	DCP	C4-N4	2.08	1.38	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	1414	DCP	O4'-C1'-N1	2.75	112.78	107.86
5	A	414	DCP	C2'-C1'-N1	-2.66	107.64	113.77
5	B	1414	DCP	C2'-C1'-N1	-2.57	107.86	113.77
5	A	414	DCP	O4'-C1'-N1	2.09	111.60	107.86
5	A	414	DCP	O2-C2-N3	-2.03	119.03	122.33

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	414	DCP	PB-O3A-PA-O1A
5	A	414	DCP	PB-O3B-PG-O1G
5	A	414	DCP	PB-O3B-PG-O2G
5	A	414	DCP	PB-O3B-PG-O3G
5	B	1414	DCP	PB-O3B-PG-O2G
5	B	1414	DCP	PB-O3B-PG-O3G
5	A	414	DCP	PB-O3A-PA-O2A

Continued on next page...

Continued from previous page...

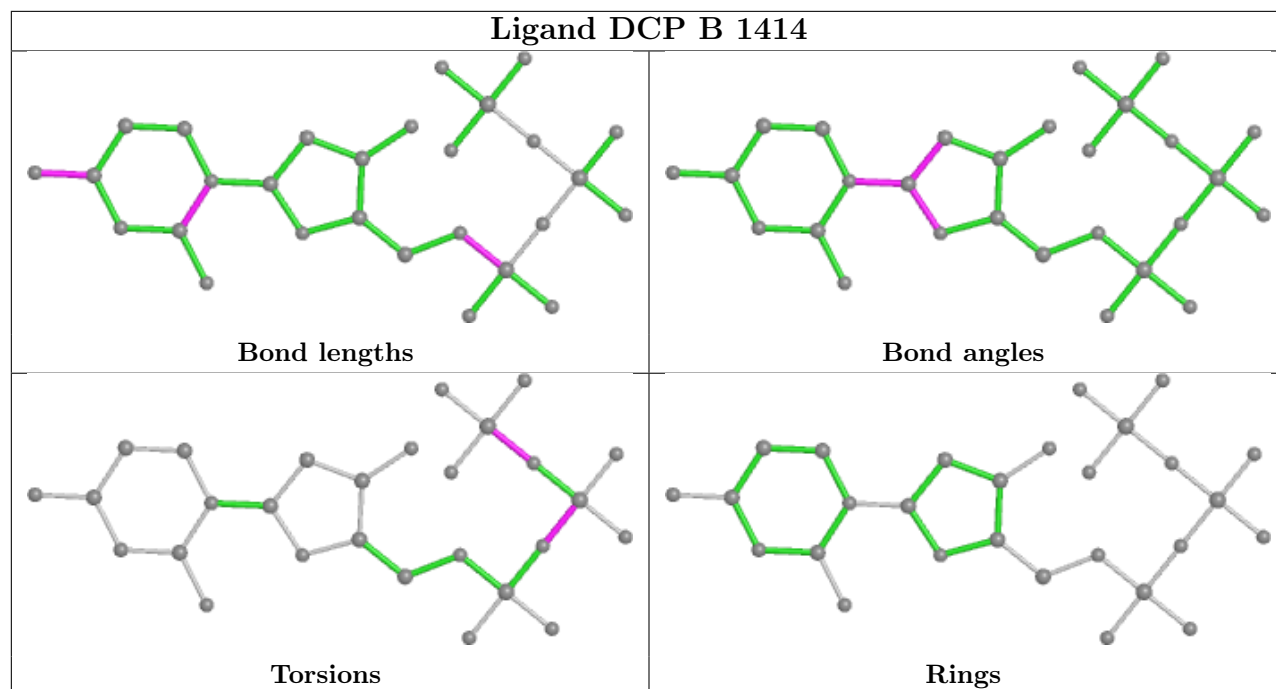| Mol	Chain	Res	Type	Atoms
5	B	1414	DCP	PA-O3A-PB-O2B

There are no ring outliers.

No monomer is involved in short contacts.

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 DCP A 414

Bond lengths

Bond angles

Torsions

Rings

Ligand DCP B 1414

Bond lengths

Bond angles

Torsions

Rings
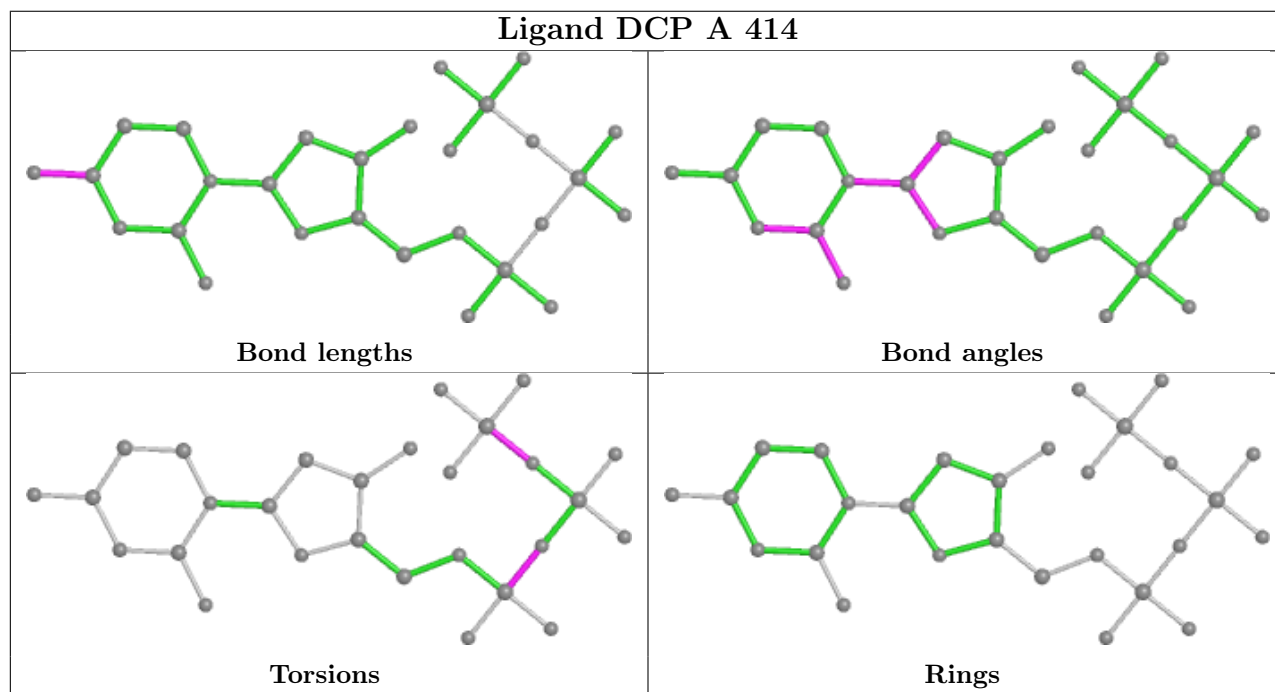
5.7 Other polymers ⓘ

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	D	12/13 (92%)	-0.29	0	100	100	12, 16, 24, 26	0
1	H	12/13 (92%)	0.32	1 (8%)	11	17	15, 17, 23, 26	0
2	E	16/19 (84%)	0.24	2 (12%)	3	6	10, 19, 43, 47	0
2	J	18/19 (94%)	1.17	3 (16%)	1	2	11, 18, 48, 52	0
3	A	341/360 (94%)	0.22	11 (3%)	47	57	8, 14, 24, 29	0
3	B	341/360 (94%)	0.46	24 (7%)	16	24	9, 16, 23, 27	0
All	All	740/784 (94%)	0.35	41 (5%)	25	34	8, 15, 24, 52	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	1901	DC	8.3
2	J	1902	DT	5.8
3	B	1196	LYS	4.6
3	B	1169	GLU	4.3
3	B	1234	ASN	4.1
3	A	25	GLY	4.1
2	E	903	DA	4.1
3	B	1210	PHE	3.7
2	E	904	DA	3.3
1	H	1801	DG	3.1
3	B	1188	ASN	3.1
3	B	1216	MET	3.0
3	A	21	PRO	3.0
3	B	1172	LYS	3.0
3	A	97	GLU	2.9
3	B	1197	LEU	2.9
3	A	238	ARG	2.7
3	A	38	GLU	2.7
3	A	287	VAL	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	B	1176	ARG	2.7
3	B	1215	GLY	2.6
3	B	1208	ILE	2.6
2	J	1903	DA	2.6
3	B	1288	ALA	2.4
3	A	277	ASP	2.4
3	A	274	TYR	2.4
3	B	1177	GLU	2.4
3	B	1327	GLU	2.4
3	A	48	TYR	2.4
3	B	1212	LYS	2.4
3	B	1134	GLU	2.4
3	B	1278	LYS	2.3
3	B	1097	GLU	2.3
3	A	169	GLU	2.3
3	B	1211	ASP	2.2
3	B	1207	SER	2.2
3	B	1209	GLU	2.2
3	B	1238	ARG	2.1
3	B	1341	ILE	2.1
3	A	241	VAL	2.1
3	B	1192	GLU	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.

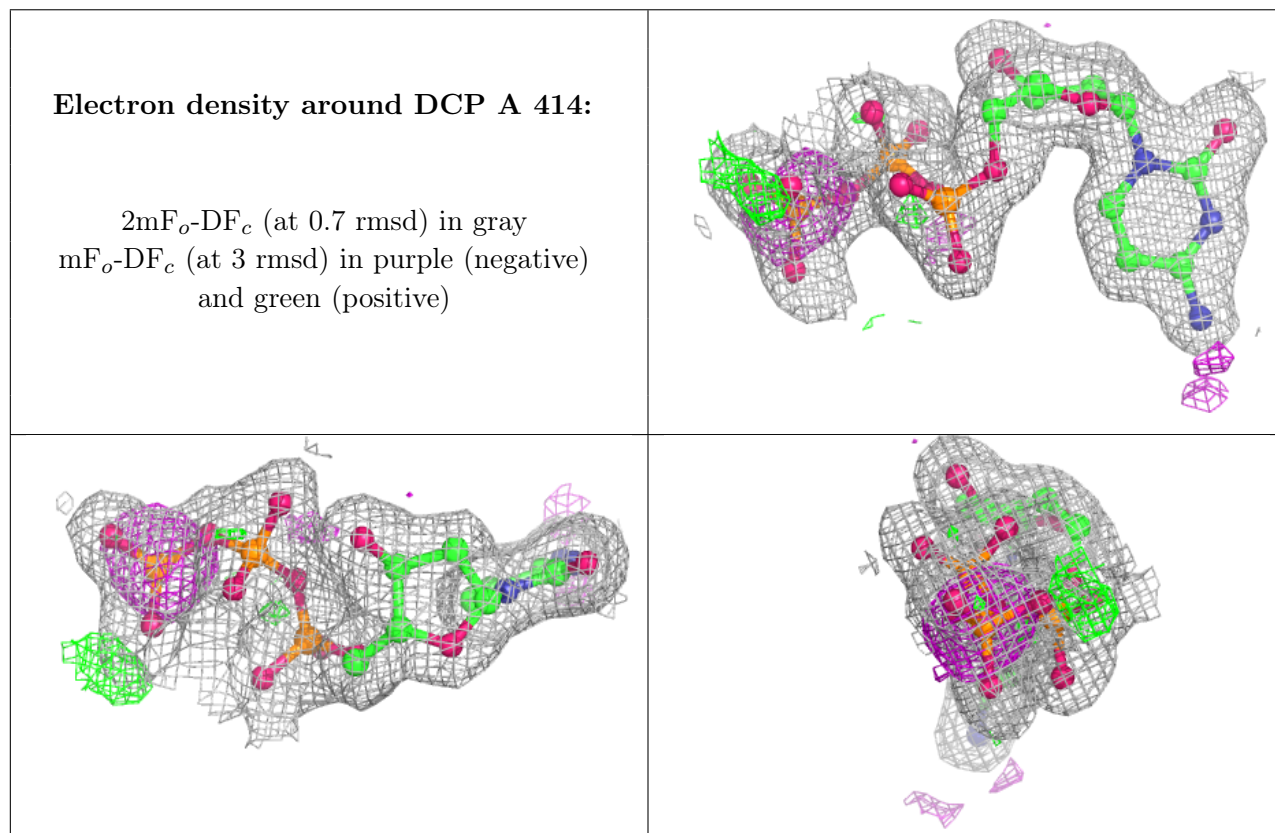| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	8OG	E	906	23/24	0.96	0.09	9,11,16,17	0
1	DDG	H	1813	21/22	0.96	0.10	10,12,20,20	0
1	DDG	D	813	21/22	0.97	0.11	12,14,16,17	0
2	8OG	J	1906	23/24	0.97	0.10	7,9,11,12	0

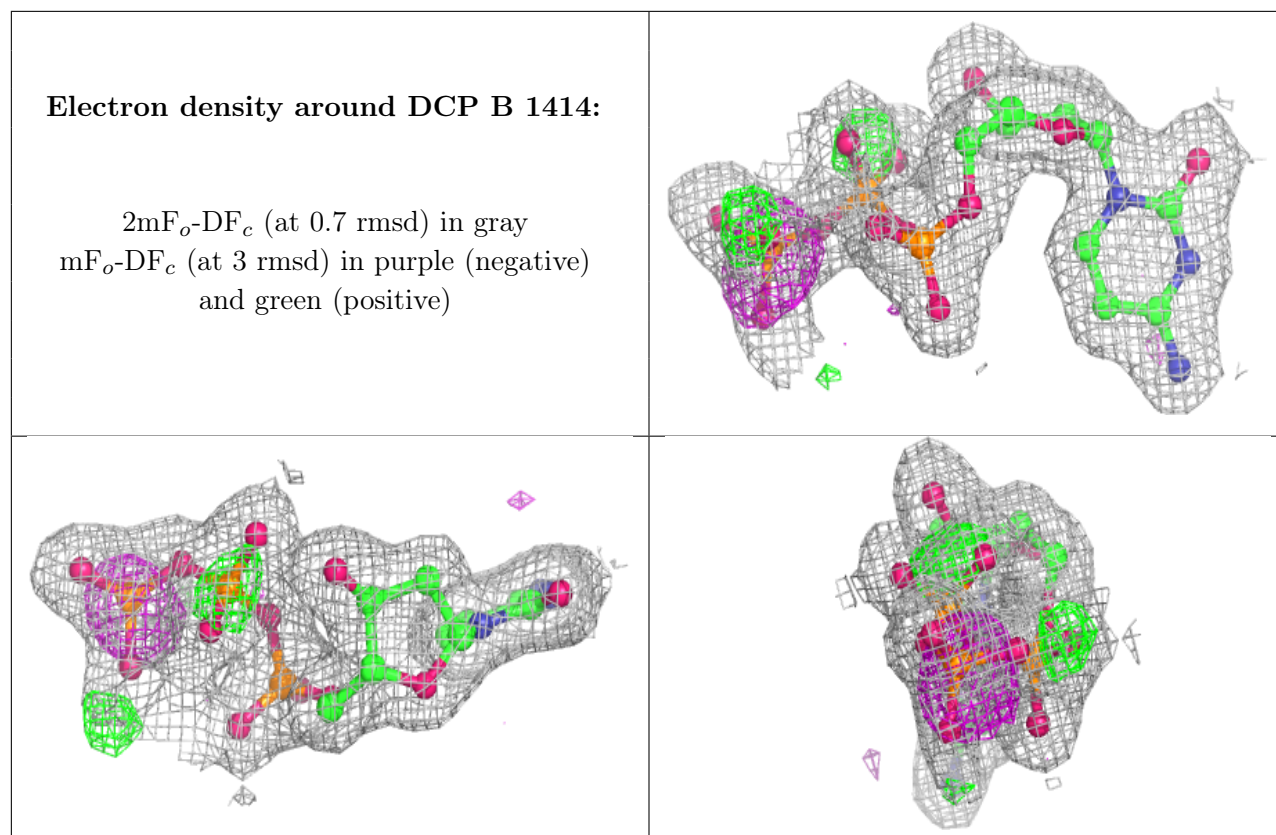## 6.3 Carbohydrates ⓘ

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	DCP	A	414	28/28	0.94	0.11	7,10,16,21	0
5	DCP	B	1414	28/28	0.95	0.11	7,10,14,20	0
4	CA	D	418	1/1	0.96	0.18	27,27,27,27	0
4	CA	A	417	1/1	0.98	0.05	14,14,14,14	0
4	CA	B	1416	1/1	0.99	0.11	11,11,11,11	0
4	CA	B	1417	1/1	0.99	0.06	25,25,25,25	0
4	CA	A	416	1/1	0.99	0.08	12,12,12,12	0
4	CA	B	1415	1/1	0.99	0.04	22,22,22,22	0
4	CA	A	415	1/1	1.00	0.08	17,17,17,17	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 DCP A 414:

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

Electron density around DCP 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 ⓘ

There are no such residues in this entry.