Full wwPDB X-ray Structure Validation Report (i)

Jun 15, 2024 – 07:39 PM EDT

PDB ID	:	4Q43
Title	:	Polymerase-damaged DNA complex
Authors	:	Kottur, J.; Sharma, A.; Nair, D.T.
Deposited on	:	2014-04-13
Resolution	:	2.45 Å(reported)

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

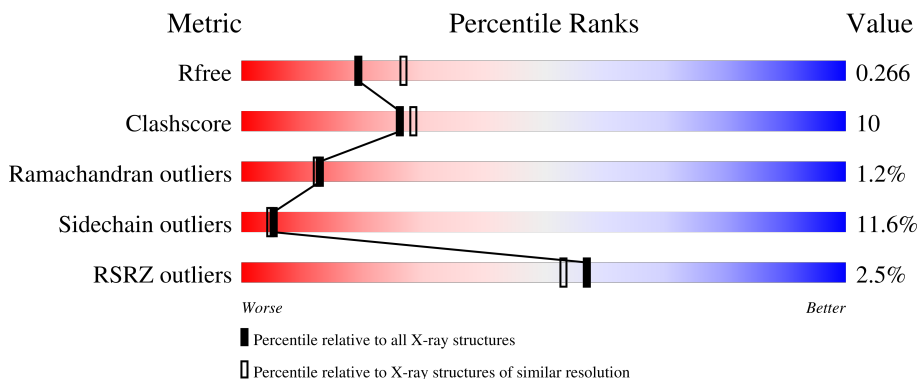|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.45 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		0.266
Clashscore		10
Ramachandran outliers		1.2%
Sidechain outliers		11.6%
RSRZ outliers		2.5%

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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	352	4% — 68% / 22% / 6% / ·
1	F	352	.% — 72% / 21% / · ·
2	B	18	50% / 39% / 11%
2	G	18	6% — 39% / 56% / 6%
3	C	18	6% — 72% / 11% / 11% / 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	18	44% 22% 11% 22%

2 Entity composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 6921 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	342	Total	C	N	O	S	0	0	0
			2683	1692	493	484	14			
1	F	342	Total	C	N	O	S	0	0	0
			2687	1695	494	484	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q47155
A	1	SER	-	expression tag	UNP Q47155
F	0	GLY	-	expression tag	UNP Q47155
F	1	SER	-	expression tag	UNP Q47155

- Molecule 2 is a DNA chain called DNA (5'-D(*TP*CP*TP*(RDG)P*GP*GP*GP*TP*CP *CP*TP*AP*GP*GP*AP*CP*CP*C)-3').
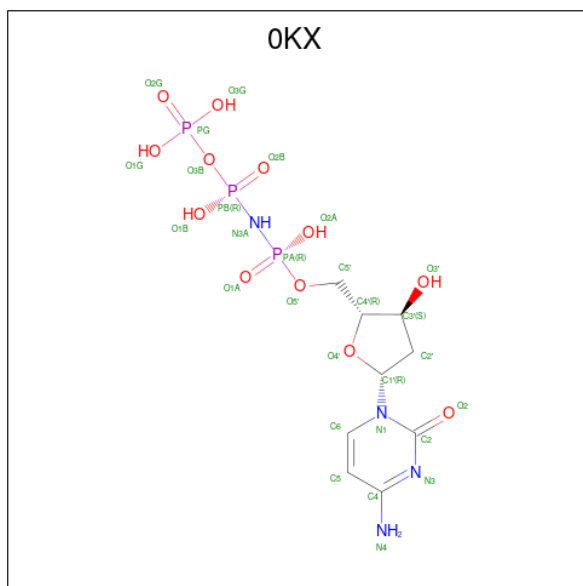
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	18	Total	C	N	O	P	0	0	0
			371	179	66	109	17			
2	G	18	Total	C	N	O	P	0	0	0
			371	179	66	109	17			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	17	Total	C	N	O	P	0	0	0
			344	164	64	100	16			
3	H	14	Total	C	N	O	P	0	0	0
			284	135	54	82	13			

- Molecule 4 is 2'-deoxy-5'-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]amino}p

hosphoryl]cytidine (three-letter code: 0KX) (formula: $C_9H_{17}N_4O_{12}P_3$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 28	C 9	N 4	O 12	P 3	0	0
4	F	1	Total 28	C 9	N 4	O 12	P 3	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total 2	Mg 2	0	0
5	F	2	Total 2	Mg 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	17	Total 17	O 17	0	0
6	F	59	Total 59	O 59	0	0
6	B	6	Total 6	O 6	0	0
6	C	13	Total 13	O 13	0	0
6	G	19	Total 19	O 19	0	0

Continued on next page...

Continued from previous page...

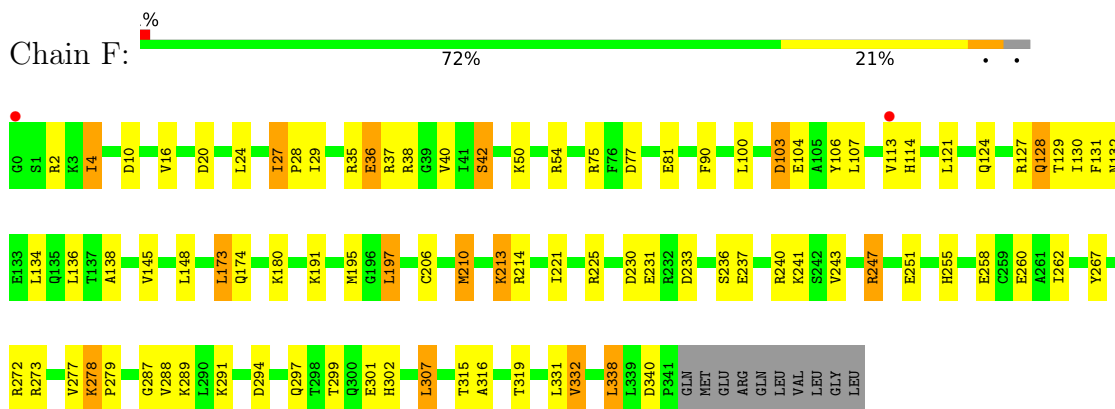| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	7	Total 7	O 7	0	0

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.
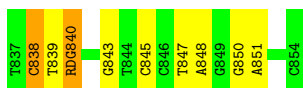
• Molecule 1: DNA polymerase IV
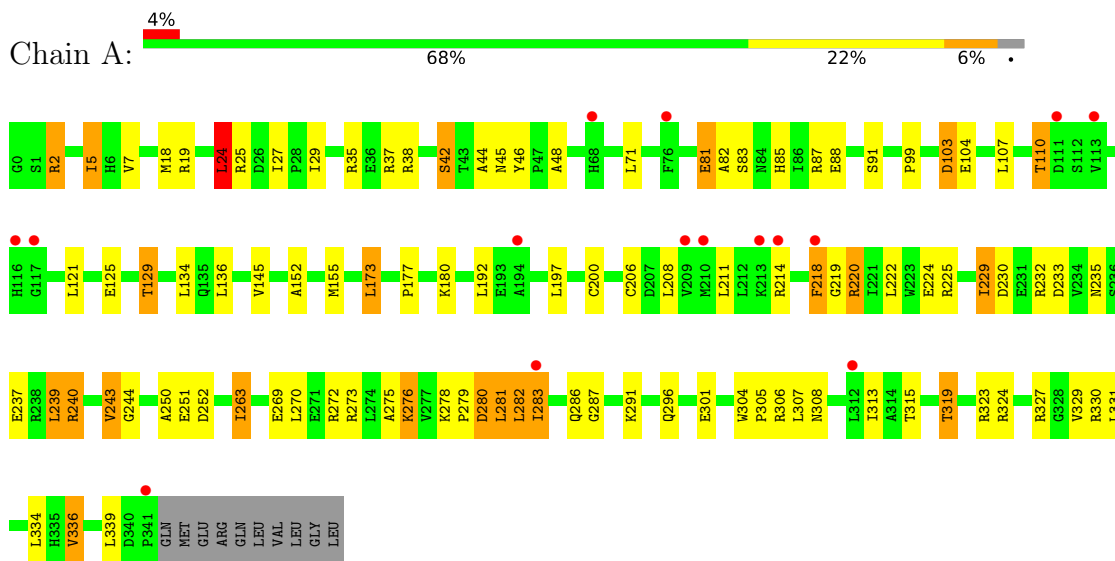
• Molecule 1: DNA polymerase IV

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

T837 C838 T839 RDG840 G843 T844 C845 C846 T847 A848 G849 G850 A851 C854

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

Chain G:

6%

39%   56%   6%

T837 C838 T839 RDG840 G843 T844 C845 A848 G849 G850 C853 C854

● Molecule 3: DNA (5'-D(*T*CP*TP*AP*GP*GP*GP*TP*CP*CP*TP*AP*GP*GP*AP*CP*CP*C)-3')

Chain C:

6%

72%   11%   11%   6%

DT C857 G861 C865 T866 G869 C873

● Molecule 3: DNA (5'-D(*T*CP*TP*AP*GP*GP*GP*TP*CP*CP*TP*AP*GP*GP*AP*CP*CP*C)-3')

Chain H:

44%   22%   11%   22%
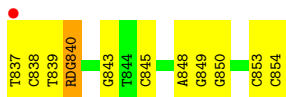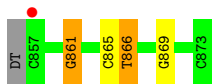
DT DC DT DA G860 G861 G862 T863 C864 C865 T866 C872 C873
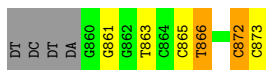
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	85.83Å 56.79Å 112.02Å 90.00° 91.96° 90.00°	Depositor
Resolution (Å)	40.52 – 2.45 40.52 – 2.45	Depositor EDS
% Data completeness (in resolution range)	93.8 (40.52-2.45) 98.9 (40.52-2.45)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.45 (at 2.45Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8_1069)	Depositor
R, R_{free}	0.214 , 0.268 0.215 , 0.266	Depositor DCC
R_{free} test set	1988 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	47.6	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 46.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6921	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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: MG, 0KX, RDG

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	0.33	0/2734	0.57	0/3695
1	F	0.38	0/2738	0.57	0/3699
2	B	0.63	0/382	1.42	2/585 (0.3%)
2	G	0.61	0/382	1.35	1/585 (0.2%)
3	C	0.67	0/385	1.38	4/592 (0.7%)
3	H	0.65	0/318	1.39	4/489 (0.8%)
All	All	0.43	0/6939	0.83	11/9645 (0.1%)

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	861	DG	O4'-C1'-N9	7.47	113.23	108.00
2	G	853	DC	O4'-C1'-N1	7.14	113.00	108.00
2	B	845	DC	O4'-C1'-N1	6.38	112.46	108.00
3	C	861	DG	O4'-C1'-C2'	-6.32	100.84	105.90
3	H	872	DC	O4'-C1'-N1	6.09	112.27	108.00
3	H	866	DT	N3-C4-O4	5.84	123.41	119.90
2	B	838	DC	O4'-C1'-N1	5.71	111.99	108.00
3	C	861	DG	C3'-C2'-C1'	-5.33	96.11	102.50
3	H	861	DG	O4'-C4'-C3'	-5.31	102.38	104.50
3	C	866	DT	N3-C4-O4	5.25	123.05	119.90
3	H	863	DT	N3-C4-O4	5.21	123.03	119.90

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	2683	0	2728	71	0
1	F	2687	0	2739	49	0
2	B	371	0	208	8	0
2	G	371	0	208	10	1
3	C	344	0	192	3	0
3	H	284	0	157	2	0
4	A	28	0	16	1	0
4	F	28	0	15	1	0
5	A	2	0	0	0	0
5	F	2	0	0	0	0
6	A	17	0	0	6	0
6	B	6	0	0	0	0
6	C	13	0	0	2	0
6	F	59	0	0	7	0
6	G	19	0	0	1	0
6	H	7	0	0	0	0
All	All	6921	0	6263	132	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 10.

All (132) 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:F:128:GLN:O	1:F:130:ILE:N	2.10	0.85
1:A:46:TYR:OH	6:A:510:HOH:O	1.98	0.81
2:G:850:DG:OP2	6:G:909:HOH:O	2.00	0.78
1:F:294:ASP:OD2	6:F:517:HOH:O	2.03	0.76
1:F:24:LEU:HB3	1:F:29:ILE:HG21	1.71	0.73
1:A:278:LYS:HG2	1:A:282:LEU:HB2	1.76	0.68
1:A:2:ARG:NH2	1:A:232:ARG:O	2.27	0.67
1:A:85:HIS:HB3	1:A:134:LEU:HD21	1.76	0.66
1:F:288:VAL:HG11	1:F:316:ALA:HB2	1.78	0.66
1:F:38:ARG:NH1	6:F:557:HOH:O	2.26	0.65
1:F:174:GLN:OE1	6:F:518:HOH:O	2.15	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:282:LEU:HD11	1:A:339:LEU:HG	1.79	0.64
1:F:302:HIS:HE1	1:F:315:THR:HG21	1.63	0.63
1:A:283:ILE:HG13	1:A:307:LEU:N	2.12	0.63
1:A:206:CYS:SG	1:A:211:LEU:HD11	2.39	0.62
1:A:5:ILE:HG23	1:A:107:LEU:HB2	1.81	0.62
1:F:241:LYS:NZ	2:G:845:DC:OP2	2.22	0.62
1:A:83:SER:OG	1:A:87:ARG:NH1	2.33	0.62
1:F:10:ASP:O	6:F:510:HOH:O	2.16	0.62
1:F:302:HIS:CE1	1:F:315:THR:HG21	2.35	0.62
1:F:240:ARG:NH1	2:G:843:DG:OP1	2.34	0.61
1:F:4:ILE:HD11	1:F:106:TYR:CD1	2.37	0.59
1:A:283:ILE:HD12	1:A:286:GLN:HB3	1.84	0.59
1:A:38:ARG:NH2	6:A:507:HOH:O	2.35	0.59
1:A:252:ASP:OD1	6:A:507:HOH:O	2.17	0.59
1:F:2:ARG:NH1	6:F:524:HOH:O	2.28	0.59
1:A:280:ASP:OD1	1:A:280:ASP:N	2.34	0.59
1:A:287:GLY:HA3	1:A:301:GLU:HG2	1.84	0.59
1:A:263:ILE:HG21	1:A:313:ILE:HG12	1.85	0.58
1:F:4:ILE:HD11	1:F:106:TYR:HD1	1.69	0.58
1:F:145:VAL:HB	1:F:230:ASP:HB3	1.84	0.58
1:A:240:ARG:NH2	2:B:843:DG:OP1	2.35	0.57
1:A:296:GLN:NE2	1:A:327:ARG:HH12	2.02	0.57
1:A:315:THR:O	1:A:319:THR:OG1	2.22	0.57
1:A:224:GLU:HB3	1:A:229:ILE:HG13	1.87	0.56
1:A:88:GLU:O	1:A:91:SER:OG	2.22	0.56
2:B:847:DT:H2"	2:B:848:DA:C8	2.42	0.55
1:A:243:VAL:HG13	1:A:336:VAL:HG22	1.88	0.55
1:A:278:LYS:HD3	1:A:282:LEU:HD22	1.88	0.55
1:A:19:ARG:HD3	1:A:136:LEU:HD13	1.88	0.55
1:F:90:PHE:HB3	1:F:107:LEU:HD21	1.89	0.55
1:A:304:TRP:CD2	1:A:305:PRO:HD2	2.42	0.55
1:F:291:LYS:HB3	1:F:331:LEU:HB3	1.89	0.54
1:A:269:GLU:OE2	1:A:272:ARG:NH2	2.30	0.54
1:A:177:PRO:HD2	1:A:180:LYS:HD2	1.90	0.53
1:F:42:SER:OG	4:F:401:0KX:O2	2.24	0.53
1:A:304:TRP:NE1	1:A:308:ASN:OD1	2.41	0.53
1:A:44:ALA:HB1	1:A:48:ALA:HB3	1.90	0.53
1:A:291:LYS:HB3	1:A:331:LEU:HB3	1.91	0.53
2:G:840:RDG:H12	2:G:840:RDG:OP2	2.09	0.53
1:F:287:GLY:HA3	1:F:301:GLU:HG2	1.91	0.52
1:A:211:LEU:HD12	1:A:211:LEU:H	1.74	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:283:ILE:CD1	1:A:286:GLN:HB3	2.39	0.52
1:F:210:MET:HA	1:F:213:LYS:HB2	1.92	0.51
1:A:239:LEU:HD22	1:A:239:LEU:H	1.76	0.51
1:A:42:SER:OG	4:A:401:0KX:O2	2.28	0.51
1:A:145:VAL:HB	1:A:230:ASP:HB3	1.92	0.51
2:B:840:RDG:H12	2:B:840:RDG:OP2	2.10	0.51
1:A:330:ARG:HH22	2:B:839:DT:H5"	1.75	0.50
1:A:24:LEU:HD12	1:A:27:ILE:HD11	1.92	0.50
1:F:35:ARG:HH21	2:G:838:DC:P	2.34	0.50
1:A:283:ILE:HD13	1:A:336:VAL:HB	1.92	0.50
1:F:127:ARG:HD2	1:F:138:ALA:O	2.12	0.50
1:F:273:ARG:HD3	6:F:541:HOH:O	2.11	0.50
3:C:869:DG:N7	6:C:909:HOH:O	2.35	0.50
1:F:255:HIS:N	1:F:258:GLU:OE1	2.37	0.49
1:A:278:LYS:HG2	1:A:282:LEU:H	1.78	0.49
1:A:304:TRP:CG	1:A:305:PRO:HD2	2.48	0.49
1:F:233:ASP:OD1	6:F:544:HOH:O	2.20	0.49
2:G:848:DA:H2'	2:G:849:DG:C8	2.47	0.49
1:A:29:ILE:HG12	1:A:45:ASN:ND2	2.28	0.48
1:A:125:GLU:O	1:A:129:THR:HG22	2.13	0.48
2:B:850:DG:H2"	2:B:851:DA:C8	2.49	0.47
1:A:244:GLY:O	1:A:273:ARG:NH2	2.47	0.47
1:A:87:ARG:HH21	1:A:99:PRO:HB3	1.79	0.47
1:A:152:ALA:HA	1:A:155:MET:HE3	1.96	0.47
1:A:218:PHE:N	6:A:515:HOH:O	2.46	0.47
1:A:235:ASN:O	1:A:237:GLU:N	2.41	0.47
1:F:132:ASN:O	1:F:134:LEU:N	2.47	0.47
1:A:220:ARG:N	6:A:515:HOH:O	2.48	0.46
1:A:103:ASP:OD1	1:A:104:GLU:HG3	2.16	0.46
1:A:18:MET:HG2	1:A:25:ARG:HA	1.97	0.46
1:F:40:VAL:HG21	2:G:840:RDG:C5	2.46	0.46
1:F:20:ASP:OD2	1:F:75:ARG:NH2	2.43	0.46
3:C:865:DC:H2'	3:C:866:DT:C6	2.51	0.45
1:A:275:ALA:HA	1:A:279:PRO:HA	1.99	0.45
1:A:81:GLU:HG2	1:A:82:ALA:N	2.31	0.45
1:A:278:LYS:HG3	1:A:280:ASP:OD1	2.16	0.45
1:F:77:ASP:O	1:F:81:GLU:HG2	2.17	0.45
1:A:35:ARG:NH2	2:B:838:DC:OP1	2.49	0.45
1:A:2:ARG:NE	6:A:517:HOH:O	2.10	0.45
1:A:270:LEU:HD22	1:A:334:LEU:HB3	1.98	0.44
1:A:291:LYS:CB	1:A:331:LEU:HB3	2.46	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:338:LEU:HD12	1:F:338:LEU:HA	1.81	0.44
1:A:2:ARG:HG3	1:A:110:THR:HG21	1.99	0.44
1:F:195:MET:HG3	1:F:197:LEU:HD22	1.99	0.44
3:C:861:DG:N3	6:C:912:HOH:O	2.36	0.44
1:A:278:LYS:CG	1:A:282:LEU:HB2	2.45	0.44
1:A:330:ARG:NH2	2:B:839:DT:H5"	2.31	0.44
1:F:124:GLN:OE1	1:F:127:ARG:NH2	2.51	0.44
1:A:208:LEU:HA	1:A:211:LEU:HD13	1.99	0.44
1:F:50:LYS:HE2	1:F:50:LYS:HB3	1.82	0.44
1:F:100:LEU:HD11	1:F:106:TYR:CE2	2.53	0.43
1:F:148:LEU:HD21	1:F:173:LEU:HD11	1.99	0.43
1:A:173:LEU:O	1:A:200:CYS:HB2	2.19	0.43
1:A:38:ARG:O	2:B:839:DT:H5'	2.18	0.43
1:F:262:ILE:HG21	1:F:332:VAL:HG13	2.00	0.43
3:H:872:DC:H2'	3:H:873:DC:C6	2.53	0.42
1:F:27:ILE:HD12	1:F:28:PRO:HD2	2.00	0.42
1:F:35:ARG:NH2	2:G:837:DT:O3'	2.40	0.42
1:A:24:LEU:HG	1:A:29:ILE:HG21	2.01	0.42
1:F:38:ARG:HD2	2:G:839:DT:OP2	2.20	0.42
2:G:838:DC:H1'	2:G:839:DT:C6	2.54	0.42
1:F:278:LYS:HA	1:F:279:PRO:HD3	1.87	0.42
1:A:87:ARG:HE	1:A:99:PRO:HB3	1.85	0.42
3:H:865:DC:H2'	3:H:866:DT:C6	2.55	0.42
1:F:221:ILE:O	1:F:225:ARG:HG2	2.19	0.41
1:F:103:ASP:OD1	1:F:104:GLU:HG3	2.21	0.41
1:A:5:ILE:HD11	1:A:7:VAL:HG22	2.03	0.41
1:A:282:LEU:O	1:A:306:ARG:HA	2.21	0.41
1:A:25:ARG:HA	1:A:25:ARG:HD3	1.88	0.41
1:A:225:ARG:NH1	1:A:229:ILE:O	2.54	0.41
1:A:269:GLU:O	1:A:273:ARG:HG3	2.20	0.41
1:F:36:GLU:H	1:F:36:GLU:HG3	1.44	0.41
1:F:247:ARG:HE	1:F:247:ARG:HA	1.85	0.41
1:F:289:LYS:HG3	1:F:299:THR:HG22	2.02	0.41
1:F:267:TYR:CZ	1:F:307:LEU:HD13	2.55	0.41
1:F:16:VAL:HG22	1:F:136:LEU:HD11	2.03	0.41
1:A:38:ARG:HD2	1:A:250:ALA:O	2.22	0.40
1:A:323:ARG:O	1:A:327:ARG:NH2	2.54	0.40
1:F:127:ARG:C	1:F:128:GLN:O	2.58	0.40
1:F:131:PHE:HA	1:F:136:LEU:O	2.20	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:838:DC:N4	2:G:854:DC:OP1[1_545]	2.10	0.10

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	340/352 (97%)	310 (91%)	24 (7%)	6 (2%)	8	6
1	F	340/352 (97%)	320 (94%)	18 (5%)	2 (1%)	25	29
All	All	680/704 (97%)	630 (93%)	42 (6%)	8 (1%)	13	12

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	LEU
1	F	129	THR
1	A	24	LEU
1	A	276	LYS
1	F	128	GLN
1	A	281	LEU
1	A	283	ILE
1	A	219	GLY

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	285/297 (96%)	253 (89%)	32 (11%)	6	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	F	286/297 (96%)	252 (88%)	34 (12%)	5 4
All	All	571/594 (96%)	505 (88%)	66 (12%)	5 5

All (66) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	A	5	ILE
1	A	24	LEU
1	A	37	ARG
1	A	42	SER
1	A	71	LEU
1	A	81	GLU
1	A	103	ASP
1	A	110	THR
1	A	121	LEU
1	A	129	THR
1	A	173	LEU
1	A	192	LEU
1	A	197	LEU
1	A	214	ARG
1	A	218	PHE
1	A	220	ARG
1	A	222	LEU
1	A	229	ILE
1	A	233	ASP
1	A	239	LEU
1	A	240	ARG
1	A	243	VAL
1	A	251	GLU
1	A	263	ILE
1	A	276	LYS
1	A	280	ASP
1	A	281	LEU
1	A	319	THR
1	A	324	ARG
1	A	329	VAL
1	A	336	VAL
1	F	4	ILE
1	F	27	ILE
1	F	36	GLU
1	F	37	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	42	SER
1	F	54	ARG
1	F	103	ASP
1	F	113	VAL
1	F	114	HIS
1	F	121	LEU
1	F	173	LEU
1	F	180	LYS
1	F	191	LYS
1	F	197	LEU
1	F	206	CYS
1	F	210	MET
1	F	213	LYS
1	F	214	ARG
1	F	231	GLU
1	F	236	SER
1	F	237	GLU
1	F	243	VAL
1	F	247	ARG
1	F	251	GLU
1	F	260	GLU
1	F	272	ARG
1	F	277	VAL
1	F	278	LYS
1	F	297	GLN
1	F	307	LEU
1	F	319	THR
1	F	332	VAL
1	F	338	LEU
1	F	340	ASP

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

Mol	Chain	Res	Type
1	A	255	HIS
1	A	296	GLN
1	F	21	ASN
1	F	174	GLN
1	F	297	GLN
1	F	302	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	RDG	G	840	2	21,31,32	2.35	8 (38%)	19,44,47	1.58	5 (26%)
2	RDG	B	840	2	21,31,32	2.34	9 (42%)	19,44,47	1.68	6 (31%)

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	RDG	G	840	2	-	2/6/26/27	0/4/4/4
2	RDG	B	840	2	-	2/6/26/27	0/4/4/4

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	840	RDG	O6-C6	6.07	1.35	1.23
2	B	840	RDG	O6-C6	5.90	1.35	1.23
2	B	840	RDG	C2-N2	4.19	1.43	1.34
2	G	840	RDG	O4'-C1'	3.84	1.51	1.42
2	G	840	RDG	C2-N2	3.71	1.42	1.34
2	B	840	RDG	O4'-C1'	3.28	1.49	1.42
2	B	840	RDG	O4'-C4'	-3.03	1.38	1.45
2	G	840	RDG	O4'-C4'	-2.89	1.38	1.45
2	G	840	RDG	O5'-C5'	-2.85	1.37	1.44
2	B	840	RDG	O5'-C5'	-2.78	1.38	1.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	840	RDG	O3'-C3'	-2.70	1.37	1.43
2	B	840	RDG	C2'-C3'	-2.54	1.46	1.52
2	B	840	RDG	C5-C6	-2.22	1.42	1.47
2	B	840	RDG	O3'-C3'	-2.22	1.38	1.43
2	B	840	RDG	C5-C4	-2.12	1.37	1.43
2	G	840	RDG	C2'-C3'	-2.11	1.47	1.52
2	G	840	RDG	C5-C4	-2.02	1.37	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	840	RDG	C5-C6-N1	3.01	119.27	113.95
2	B	840	RDG	O6-C6-C5	-2.98	118.55	124.37
2	B	840	RDG	C2'-C1'-N9	2.86	120.87	114.27
2	B	840	RDG	C5-C6-N1	2.80	118.89	113.95
2	G	840	RDG	C8-N7-C5	2.74	108.21	102.99
2	G	840	RDG	C2'-C1'-N9	2.70	120.49	114.27
2	B	840	RDG	C9-C7-C12	-2.43	104.26	112.92
2	B	840	RDG	C8-N7-C5	2.36	107.48	102.99
2	G	840	RDG	C9-C7-C12	-2.33	104.62	112.92
2	B	840	RDG	O5'-C5'-C4'	2.20	116.48	108.99
2	G	840	RDG	O6-C6-C5	-2.05	120.37	124.37

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	840	RDG	O4'-C4'-C5'-O5'
2	G	840	RDG	O4'-C4'-C5'-O5'
2	B	840	RDG	C3'-C4'-C5'-O5'
2	G	840	RDG	C3'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	840	RDG	2	0
2	B	840	RDG	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 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	$\#|Z| > 2$
4	0KX	F	401	5	28,29,29	1.47	3 (10%)	40,45,45	1.40	7 (17%)
4	0KX	A	401	5	28,29,29	1.45	5 (17%)	40,45,45	1.00	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
4	0KX	F	401	5	-	5/19/34/34	0/2/2/2
4	0KX	A	401	5	-	4/19/34/34	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	401	0KX	PB-O2B	4.09	1.52	1.46
4	A	401	0KX	PB-O2B	3.94	1.52	1.46
4	A	401	0KX	PA-O1A	3.70	1.52	1.46
4	F	401	0KX	PA-O1A	3.67	1.52	1.46
4	A	401	0KX	PB-O3B	2.25	1.61	1.59
4	F	401	0KX	C6-C5	2.06	1.39	1.35
4	A	401	0KX	PA-O2A	-2.02	1.51	1.56
4	A	401	0KX	C6-C5	2.00	1.39	1.35

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	401	0KX	O1B-PB-O3B	3.77	117.22	104.64
4	F	401	0KX	O1A-PA-N3A	-3.16	107.12	111.77
4	F	401	0KX	O2B-PB-N3A	-3.08	107.23	111.77
4	F	401	0KX	PB-O3B-PG	-2.62	123.39	132.62
4	F	401	0KX	O3B-PB-N3A	-2.26	100.31	106.59
4	F	401	0KX	C2'-C1'-N1	-2.22	108.66	113.77
4	A	401	0KX	PB-O3B-PG	-2.21	124.84	132.62
4	A	401	0KX	C2'-C1'-N1	-2.12	108.88	113.77
4	F	401	0KX	O2A-PA-O1A	2.08	114.28	109.92

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	0KX	PG-O3B-PB-O1B
4	A	401	0KX	PG-O3B-PB-O2B
4	F	401	0KX	PG-O3B-PB-O2B
4	F	401	0KX	PA-N3A-PB-O2B
4	F	401	0KX	PB-N3A-PA-O5'
4	A	401	0KX	O4'-C4'-C5'-O5'
4	A	401	0KX	C3'-C4'-C5'-O5'
4	F	401	0KX	O4'-C4'-C5'-O5'
4	F	401	0KX	PG-O3B-PB-O1B

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	401	0KX	1	0
4	A	401	0KX	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 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 0KX F 401

Bond lengths

Bond angles

Torsions

Rings

Ligand 0KX A 401

Bond lengths

Bond angles

Torsions

Rings
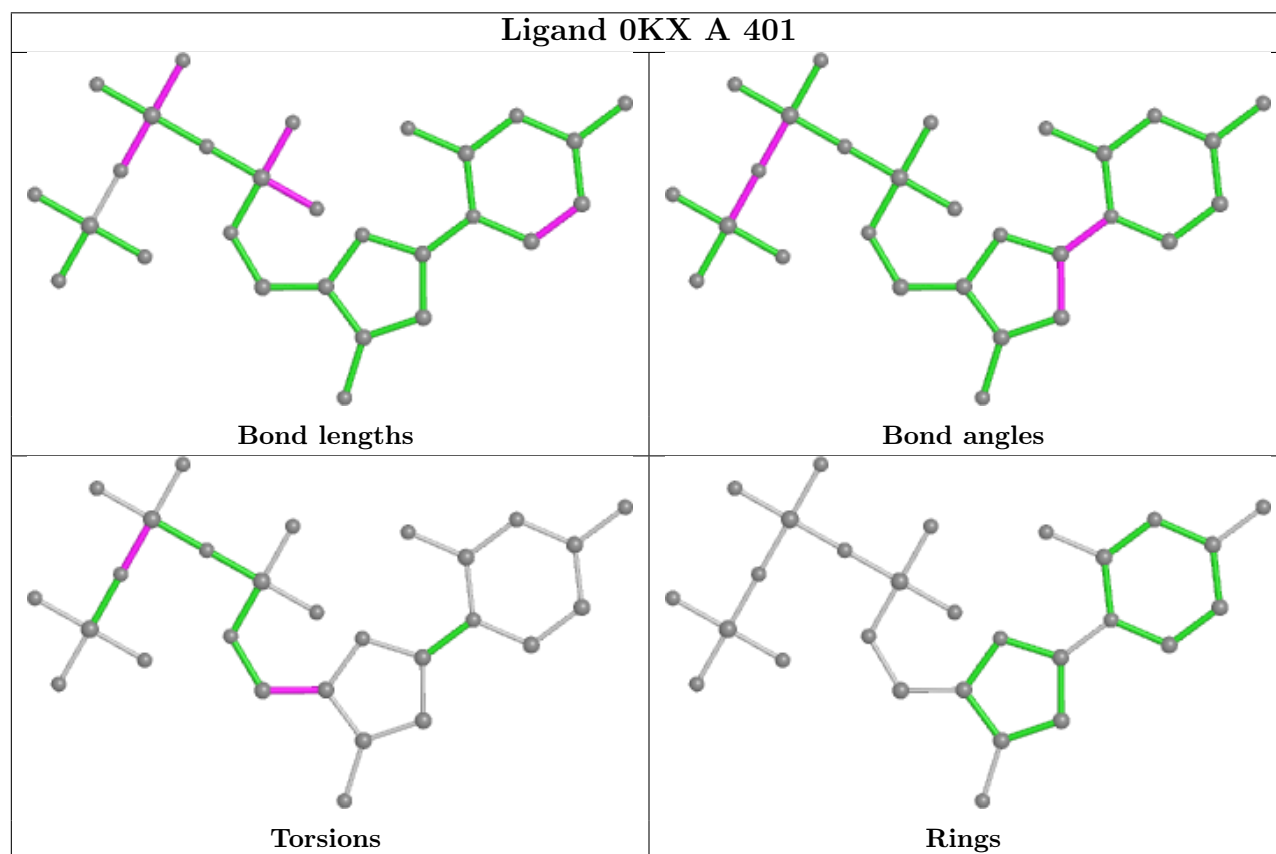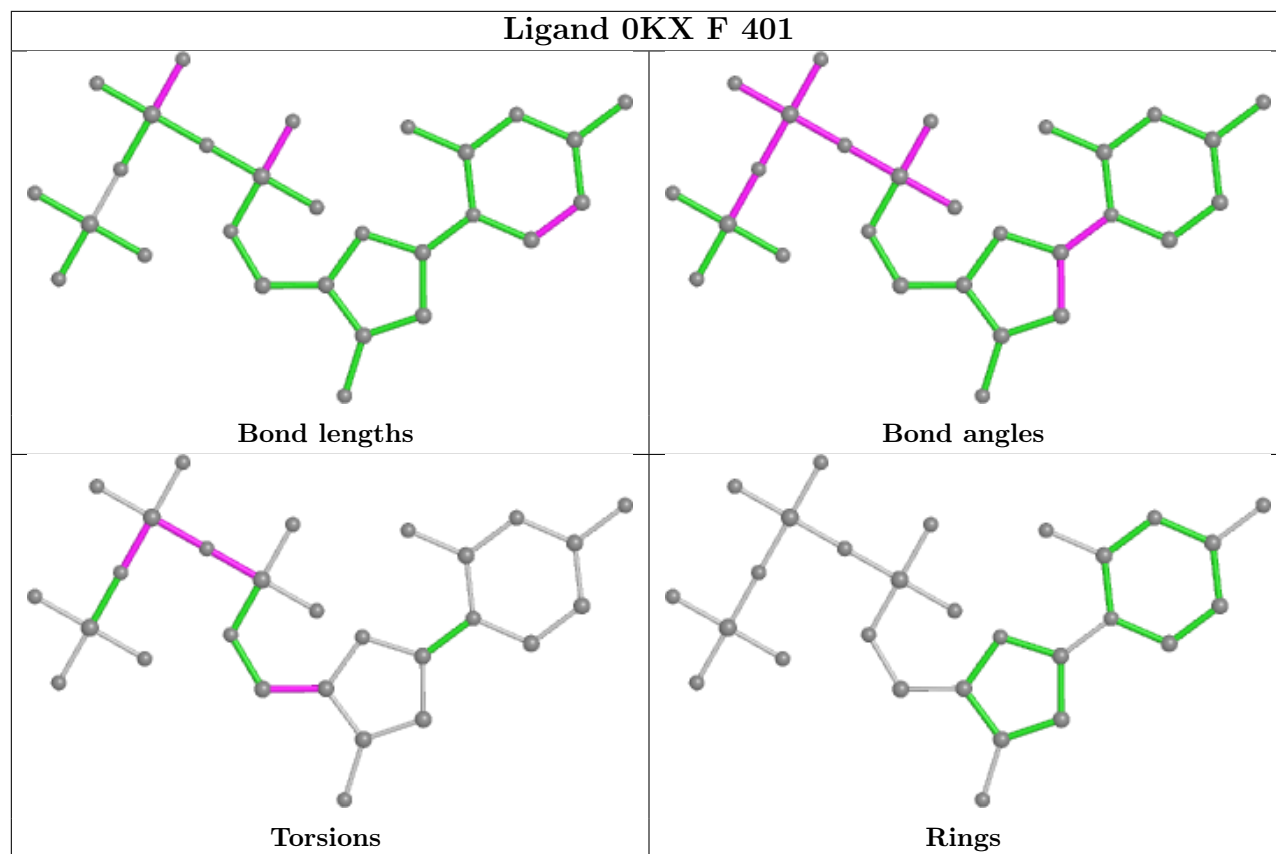
5.7 Other polymers ⓘ

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	342/352 (97%)	0.15	15 (4%)	34	32	43, 62, 82, 110	0
1	F	342/352 (97%)	-0.19	2 (0%)	89	89	25, 41, 60, 83	0
2	B	17/18 (94%)	-0.42	0	100	100	47, 54, 78, 81	0
2	G	17/18 (94%)	-0.08	1 (5%)	22	19	32, 48, 84, 85	0
3	C	17/18 (94%)	-0.06	1 (5%)	22	19	43, 56, 69, 112	0
3	H	14/18 (77%)	-0.11	0	100	100	31, 46, 69, 74	0
All	All	749/776 (96%)	-0.03	19 (2%)	57	53	25, 53, 77, 112	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	857	DC	5.7
1	F	0	GLY	5.4
1	A	113	VAL	4.9
1	A	341	PRO	3.9
1	A	214	ARG	3.7
1	A	283	ILE	3.3
1	A	117	GLY	3.2
1	A	213	LYS	3.2
1	A	218	PHE	3.1
1	A	116	HIS	3.1
2	G	837	DT	2.9
1	A	209	VAL	2.8
1	A	68	HIS	2.7
1	A	76	PHE	2.4
1	A	312	LEU	2.4
1	A	194	ALA	2.2
1	A	111	ASP	2.2
1	F	113	VAL	2.1
1	A	210	MET	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	RDG	B	840	28/29	0.96	0.16	33,43,59,65	0
2	RDG	G	840	28/29	0.96	0.14	23,34,49,56	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	MG	F	403	1/1	0.93	0.45	50,50,50,50	0
5	MG	A	403	1/1	0.95	0.11	65,65,65,65	0
4	0KX	A	401	28/28	0.96	0.17	40,48,59,64	0
5	MG	A	402	1/1	0.96	0.19	38,38,38,38	0
4	0KX	F	401	28/28	0.98	0.17	19,29,34,37	0
5	MG	F	402	1/1	0.99	0.19	15,15,15,15	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 0KX A 401:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
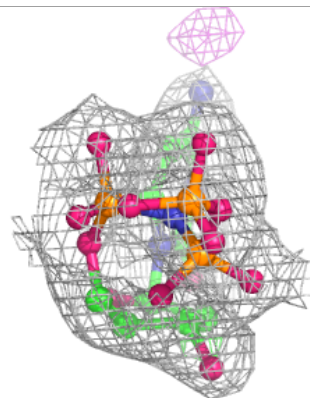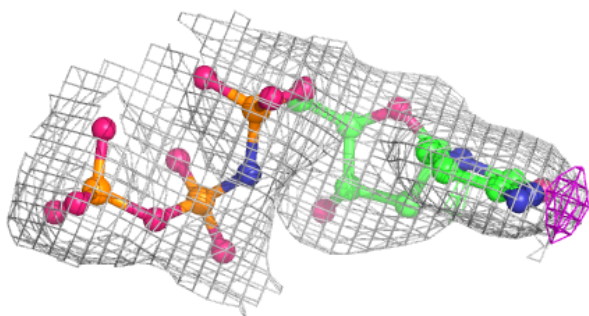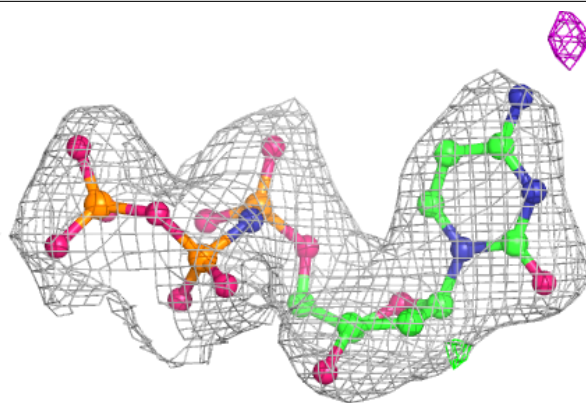
Electron density around 0KX F 401:

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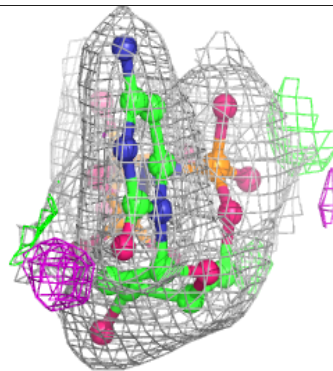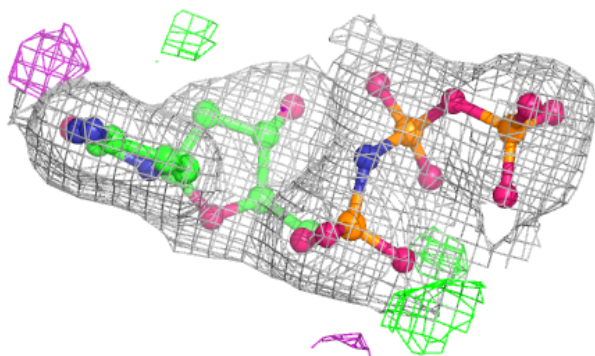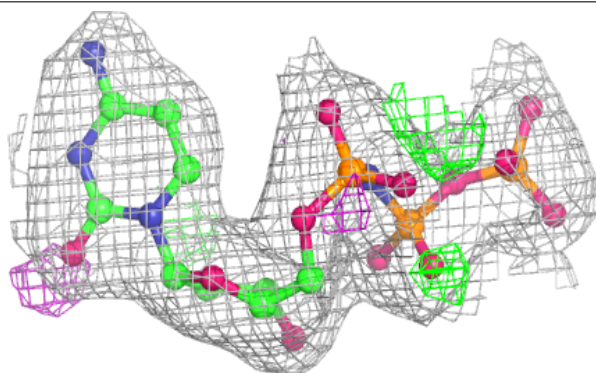
6.5 Other polymers (i)

There are no such residues in this entry.