Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2024 – 11:42 AM EDT

PDB ID	:	4KPY
Title	:	DNA binding protein and DNA complex structure
Authors	:	Sheng, G.; Zhao, H.; Wang, J.; Rao, Y.; Wang, Y.
Deposited on	:	2013-05-14
Resolution	:	2.41 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:

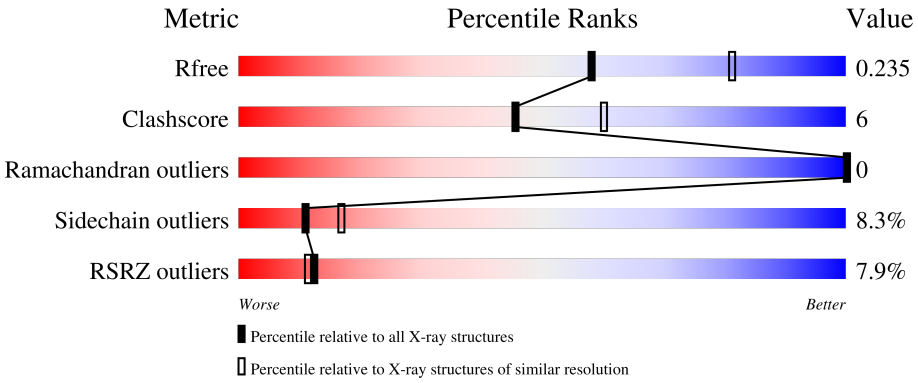
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.41 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



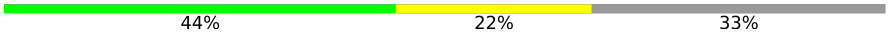
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	685	8% 80% 16%
1	B	685	9% 76% 20%
2	C	21	33% 38% 5% 24%
2	E	21	43% 29% 5% 24%
3	D	10	60% 40%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	10	
4	M	9	
4	N	9	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12268 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	676	Total	C	N	O	S	0	0	0
			5278	3372	992	907	7			
1	B	675	Total	C	N	O	S	0	1	0
			5272	3371	991	904	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	16	Total	C	N	O	P	0	0	0
			338	160	62	100	16			
2	E	16	Total	C	N	O	P	0	0	0
			338	160	62	100	16			

- Molecule 3 is a DNA chain called DNA (5'-D(P*TP*AP*CP*TP*AP*CP*CP*TP*CP*G)-3').

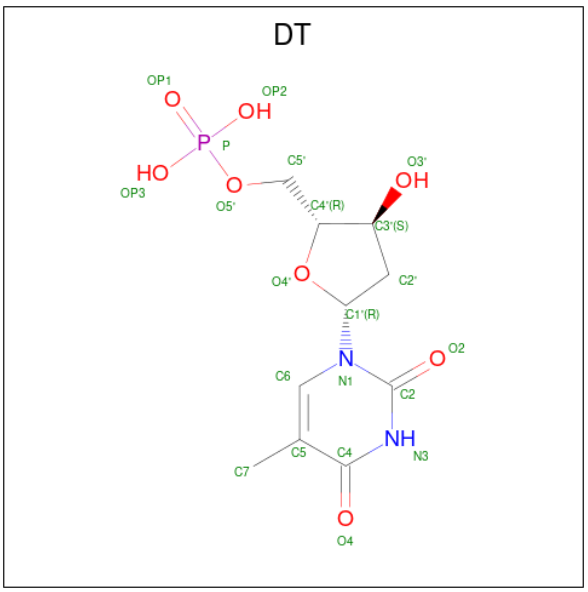
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	10	Total	C	N	O	P	0	0	0
			201	96	33	62	10			
3	F	10	Total	C	N	O	P	0	0	0
			201	96	33	62	10			

- Molecule 4 is a DNA chain called DNA (5'-D(*TP*AP*TP*AP*CP*AP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	6	Total	C	N	O	P	0	0	0
			120	57	24	33	6			
4	N	6	Total	C	N	O	P	0	0	0
			116	57	24	30	5			

- Molecule 5 is THYMIDINE-5'-MONOPHOSPHATE (three-letter code: DT) (formula:

$C_{10}H_{15}N_2O_8P$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O P 21 10 2 8 1	0	0
5	B	1	Total C N O P 21 10 2 8 1	0	0

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Mn 2 2	0	0
6	B	2	Total Mn 2 2	0	0
6	C	1	Total Mn 1 1	0	0
6	E	1	Total Mn 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	138	Total O 138 138	0	0
7	B	161	Total O 161 161	0	0
7	C	17	Total O 17 17	0	0

Continued on next page...

Continued from previous page...

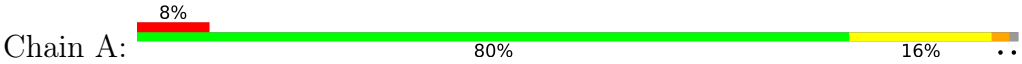
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	9	Total O 9 9	0	0
7	E	20	Total O 20 20	0	0
7	F	8	Total O 8 8	0	0
7	M	3	Total O 3 3	0	0

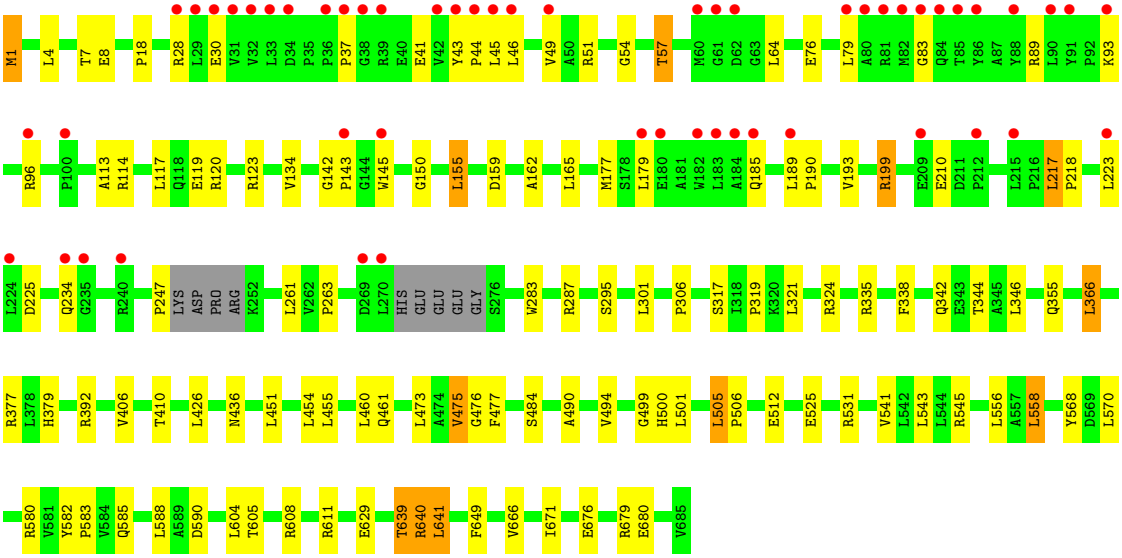
3 Residue-property plots (i)

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

• Molecule 1: Uncharacterized protein

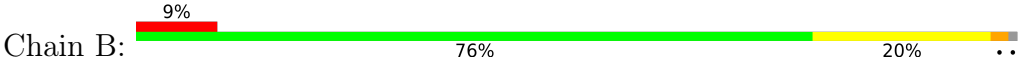
Chain A:

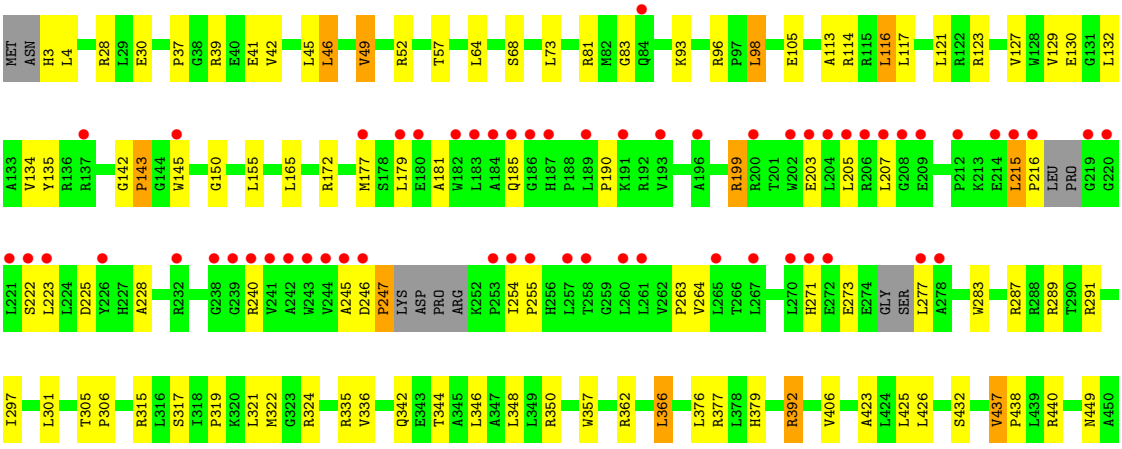


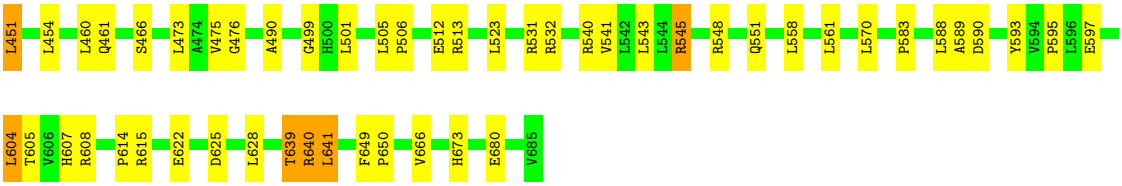


• Molecule 1: Uncharacterized protein

Chain B:

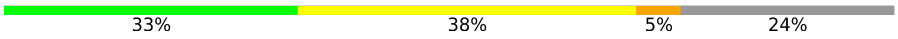






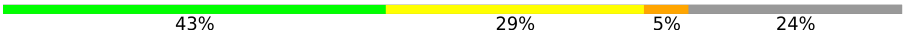
• Molecule 2: DNA (5'-D(P*TP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*TP*AP*GP*T)-3')

Chain C:



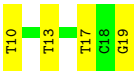
• Molecule 2: DNA (5'-D(P*TP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*TP*AP*GP*T)-3')

Chain E:



• Molecule 3: DNA (5'-D(P*TP*AP*CP*TP*AP*CP*CP*TP*CP*G)-3')

Chain D:

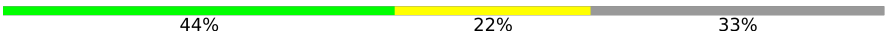


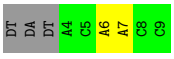
• Molecule 3: DNA (5'-D(P*TP*AP*CP*TP*AP*CP*CP*TP*CP*G)-3')

Chain F:

• Molecule 4: DNA (5'-D(*TP*AP*TP*AP*CP*AP*AP*CP*C)-3')

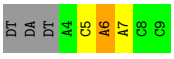
Chain M:





• Molecule 4: DNA (5'-D(*TP*AP*TP*AP*CP*AP*AP*CP*C)-3')

Chain N:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.33Å 118.36Å 160.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.31 – 2.41 48.31 – 2.41	Depositor EDS
% Data completeness (in resolution range)	97.1 (48.31-2.41) 97.0 (48.31-2.41)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.190 , 0.235 0.192 , 0.235	Depositor DCC
R_{free} test set	4033 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	44.7	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 48.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12268	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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: MN

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	0/5402	0.55	2/7333 (0.0%)
1	B	0.34	0/5399	0.56	3/7329 (0.0%)
2	C	0.87	1/379 (0.3%)	1.35	6/584 (1.0%)
2	E	0.81	1/379 (0.3%)	1.34	4/584 (0.7%)
3	D	0.92	1/223 (0.4%)	1.26	2/339 (0.6%)
3	F	0.89	1/223 (0.4%)	1.29	2/339 (0.6%)
4	M	0.68	0/134	1.31	1/203 (0.5%)
4	N	0.54	0/130	1.23	2/198 (1.0%)
All	All	0.43	4/12269 (0.0%)	0.70	22/16909 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1	DT	OP3-P	-11.03	1.48	1.61
2	C	1	DT	OP3-P	-10.51	1.48	1.61
3	D	10	DT	OP3-P	-10.45	1.48	1.61
3	F	10	DT	OP3-P	-9.81	1.49	1.61

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	11	DA	O4'-C1'-N9	8.26	113.78	108.00
2	E	1	DT	OP1-P-OP2	-8.13	107.41	119.60
4	N	7	DA	O4'-C1'-N9	7.36	113.15	108.00
2	C	1	DT	OP1-P-OP2	-7.33	108.60	119.60
2	C	15	DG	O4'-C1'-N9	6.23	112.36	108.00
1	B	247	PRO	N-CA-CB	6.10	110.62	103.30
2	C	13	DT	N3-C4-O4	6.08	123.55	119.90
2	E	9	DT	O4'-C1'-N1	6.00	112.20	108.00
2	E	1	DT	O4'-C1'-N1	-5.99	103.81	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	247	PRO	N-CA-CB	5.87	110.34	103.30
1	B	143	PRO	N-CA-CB	5.81	110.27	103.30
1	A	143	PRO	N-CA-CB	5.75	110.20	103.30
4	M	7	DA	O4'-C1'-N9	5.73	112.01	108.00
2	C	13	DT	C5-C4-O4	-5.68	120.92	124.90
1	B	116	LEU	CA-CB-CG	5.50	127.96	115.30
3	F	19	DG	O4'-C1'-N9	-5.49	104.15	108.00
2	E	3	DA	O4'-C1'-N9	-5.48	104.17	108.00
2	C	16	DT	O4'-C1'-N1	5.42	111.79	108.00
2	C	14	DT	N3-C4-O4	5.20	123.02	119.90
4	N	6	DA	O4'-C4'-C3'	-5.19	102.42	104.50
3	D	13	DT	O4'-C1'-N1	5.01	111.51	108.00
3	D	17	DT	C1'-O4'-C4'	-5.01	105.09	110.10

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	5278	0	5338	65	0
1	B	5272	0	5330	82	0
2	C	338	0	183	5	0
2	E	338	0	183	4	0
3	D	201	0	114	1	0
3	F	201	0	114	1	0
4	M	120	0	66	2	0
4	N	116	0	64	2	0
5	A	21	0	13	1	0
5	B	21	0	13	0	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
6	C	1	0	0	0	0
6	E	1	0	0	0	0
7	A	138	0	0	2	0
7	B	161	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	C	17	0	0	0	0
7	D	9	0	0	0	0
7	E	20	0	0	1	0
7	F	8	0	0	0	0
7	M	3	0	0	0	0
All	All	12268	0	11418	150	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 6.

All (150) 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:319:PRO:HG2	1:A:640:ARG:HD2	1.63	0.81
1:B:344:THR:HG21	1:B:460:LEU:HD11	1.66	0.76
1:B:513:ARG:NH2	1:B:551:GLN:O	2.17	0.75
1:B:597:GLU:O	7:B:900:HOH:O	2.05	0.73
1:A:512:GLU:OE2	1:A:545:ARG:NH2	2.23	0.71
2:E:14:DT:H3	4:M:6:DA:H2	1.37	0.71
1:A:679:ARG:NH1	7:A:852:HOH:O	2.23	0.71
1:A:28:ARG:NH1	1:A:93:LYS:O	2.25	0.69
1:A:49:VAL:HG22	1:A:79:LEU:HD11	1.76	0.67
1:A:344:THR:HG21	1:A:460:LEU:HD11	1.76	0.67
1:A:28:ARG:NH2	1:A:96:ARG:HB2	2.10	0.67
1:B:319:PRO:HG3	1:B:640:ARG:HD2	1.77	0.67
1:B:512:GLU:OE2	1:B:545:ARG:NH2	2.27	0.66
1:A:583:PRO:HD3	1:A:588:LEU:HD13	1.78	0.66
2:C:3:DA:H2'	2:C:4:DG:C8	2.32	0.65
1:A:30:GLU:OE2	1:A:93:LYS:NZ	2.30	0.64
1:B:190:PRO:HG3	1:B:263:PRO:HB3	1.79	0.64
5:A:701:DT:O2	1:B:39:ARG:NH1	2.30	0.63
1:B:121:LEU:HD22	1:B:134:VAL:HG21	1.82	0.62
2:E:3:DA:H2'	2:E:4:DG:C8	2.37	0.60
1:A:4:LEU:HD22	1:A:317:SER:HA	1.82	0.60
1:A:37:PRO:HB3	1:A:45:LEU:HD23	1.85	0.59
1:B:377:ARG:HD2	1:B:379:HIS:NE2	2.19	0.58
1:B:593:TYR:CZ	1:B:595:PRO:HG3	2.38	0.58
1:A:525:GLU:OE1	1:B:392:ARG:NH1	2.37	0.57
1:A:41:GLU:O	1:A:45:LEU:HB2	2.04	0.57
1:A:114:ARG:HH12	4:N:6:DA:H3'	1.70	0.56
1:B:461:GLN:HG3	1:B:499:GLY:O	2.04	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:190:PRO:HG3	1:A:263:PRO:HB3	1.88	0.56
1:B:590:ASP:OD2	3:F:19:DG:N1	2.28	0.56
1:B:37:PRO:HB3	1:B:45:LEU:HD23	1.88	0.55
1:B:423:ALA:HB1	1:B:673:HIS:CE1	2.41	0.55
1:B:607:HIS:O	1:B:608:ARG:HD2	2.07	0.55
1:B:130:GLU:OE1	1:B:172:ARG:NH1	2.40	0.55
1:B:41:GLU:O	1:B:45:LEU:HB2	2.07	0.54
1:B:362:ARG:NH2	7:B:904:HOH:O	2.40	0.54
1:A:120:ARG:NH1	1:A:301:LEU:O	2.41	0.54
1:B:135:TYR:HA	1:B:150:GLY:HA3	1.90	0.53
1:A:639:THR:HG21	1:A:640:ARG:HH21	1.72	0.53
1:B:283:TRP:HZ3	1:B:287:ARG:HH21	1.56	0.53
1:B:319:PRO:CG	1:B:640:ARG:HD2	2.38	0.53
1:A:45:LEU:O	1:A:49:VAL:HG23	2.07	0.53
1:B:98:LEU:HG	1:B:105:GLU:HB3	1.90	0.53
1:B:425:LEU:HD12	1:B:432:SER:HB3	1.91	0.53
1:B:350:ARG:NH1	7:B:825:HOH:O	2.42	0.52
1:B:639:THR:HG22	1:B:640:ARG:HE	1.75	0.52
1:B:615:ARG:HD3	7:E:207:HOH:O	2.10	0.51
1:B:583:PRO:HD3	1:B:588:LEU:HD13	1.92	0.51
1:B:639:THR:HG21	1:B:640:ARG:HH21	1.76	0.51
1:B:506:PRO:HG2	1:B:666:VAL:HG21	1.92	0.51
1:A:461:GLN:HG3	1:A:499:GLY:O	2.10	0.51
1:A:611:ARG:HD3	2:C:5:DG:H1'	1.93	0.50
1:B:114:ARG:NH1	4:M:6:DA:OP2	2.43	0.50
1:B:437:VAL:HG22	1:B:438:PRO:HA	1.94	0.50
1:A:119:GLU:O	1:A:123:ARG:HG3	2.12	0.50
1:A:134:VAL:O	1:A:150:GLY:HA3	2.12	0.49
1:B:246:ASP:OD1	1:B:247:PRO:N	2.45	0.49
1:A:142:GLY:HA3	1:A:145:TRP:CE2	2.48	0.49
1:A:177:MET:HE1	1:A:185:GLN:HG3	1.95	0.49
1:A:113:ALA:HB1	1:A:155:LEU:HD13	1.96	0.48
1:A:475:VAL:HG22	1:A:477:PHE:CE1	2.49	0.47
1:B:199:ARG:N	1:B:199:ARG:HD2	2.27	0.47
1:A:640:ARG:HG3	1:A:649:PHE:CE1	2.50	0.47
1:B:203:GLU:HB3	1:B:245:ALA:HB3	1.96	0.47
1:B:532:ARG:NH1	7:B:934:HOH:O	2.47	0.47
1:B:177:MET:HE1	1:B:185:GLN:HG3	1.96	0.47
1:B:322:MET:HE2	1:B:466:SER:HB2	1.97	0.47
1:B:532:ARG:NH2	7:B:905:HOH:O	2.46	0.47
1:A:295:SER:HA	1:A:306:PRO:HG3	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:605:THR:O	1:B:640:ARG:NH2	2.46	0.47
1:A:640:ARG:HG3	1:A:649:PHE:CD1	2.50	0.46
1:A:83:GLY:HA3	1:B:83:GLY:HA3	1.97	0.46
1:A:199:ARG:NH1	7:A:855:HOH:O	2.49	0.46
1:A:611:ARG:HD3	2:C:5:DG:C1'	2.45	0.46
1:A:193:VAL:HG21	1:A:261:LEU:HB3	1.97	0.46
1:A:338:PHE:CZ	1:A:455:LEU:HD13	2.51	0.46
1:A:287:ARG:HG3	1:A:582:TYR:CG	2.51	0.46
1:A:377:ARG:HD2	1:A:379:HIS:NE2	2.31	0.46
1:B:4:LEU:HD13	1:B:317:SER:HA	1.98	0.46
1:A:225:ASP:OD1	1:A:225:ASP:N	2.49	0.45
1:B:540:ARG:NH2	1:B:625:ASP:OD1	2.48	0.45
1:A:410:THR:O	1:A:436:ASN:HA	2.16	0.45
1:B:461:GLN:HE21	1:B:499:GLY:HA2	1.82	0.45
1:B:523:LEU:HD21	1:B:561:LEU:HD11	1.99	0.45
1:A:159:ASP:OD1	1:A:159:ASP:N	2.49	0.45
1:B:207:LEU:HD13	1:B:240:ARG:HE	1.81	0.45
1:A:639:THR:CG2	1:A:640:ARG:HH21	2.30	0.44
1:B:68:SER:HB2	1:B:73:LEU:HD21	1.99	0.44
1:A:18:PRO:HA	1:A:162:ALA:HA	1.99	0.44
1:B:39:ARG:O	1:B:42:VAL:HG12	2.18	0.44
1:B:315:ARG:HH12	1:B:589:ALA:HB3	1.83	0.44
1:B:476:GLY:O	1:B:490:ALA:HA	2.17	0.44
1:A:76:GLU:HG2	1:A:89:ARG:HG3	1.99	0.44
1:A:605:THR:O	1:A:640:ARG:NH2	2.50	0.44
1:B:473:LEU:HB3	1:B:541:VAL:HG12	2.00	0.44
1:B:113:ALA:O	1:B:116:LEU:HB3	2.17	0.44
1:B:215:LEU:HD13	1:B:216:PRO:HD2	1.98	0.44
1:B:215:LEU:O	1:B:222:SER:HA	2.18	0.44
1:B:225:ASP:HA	1:B:228:ALA:HB3	2.00	0.44
1:A:45:LEU:HA	1:A:45:LEU:HD12	1.69	0.43
1:A:366:LEU:HD12	1:A:366:LEU:HA	1.87	0.43
1:A:501:LEU:HD21	1:A:641:LEU:HD13	1.99	0.43
1:A:505:LEU:HD22	1:A:671:ILE:HD11	2.00	0.43
1:B:342:GLN:O	1:B:344:THR:HG23	2.18	0.43
1:B:366:LEU:HD13	1:B:376:LEU:HD23	2.00	0.43
1:B:449:ASN:ND2	2:E:2:DG:H21	2.17	0.43
1:B:640:ARG:HG3	1:B:649:PHE:CE1	2.54	0.43
1:B:639:THR:CG2	1:B:640:ARG:HH21	2.31	0.43
1:B:177:MET:HB3	1:B:181:ALA:HB3	2.00	0.43
1:B:287:ARG:HH11	1:B:291:ARG:HH22	1.66	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:30:GLU:OE1	1:B:93:LYS:NZ	2.47	0.43
1:B:501:LEU:HD21	1:B:641:LEU:HD13	2.01	0.43
2:C:4:DG:H2'	2:C:5:DG:C8	2.54	0.43
1:B:305:THR:HA	1:B:306:PRO:HD3	1.87	0.42
1:B:28:ARG:NH2	1:B:96:ARG:HB2	2.34	0.42
1:B:46:LEU:O	1:B:49:VAL:HG12	2.19	0.42
2:C:12:DG:H2'	2:C:13:DT:H72	2.00	0.42
1:A:28:ARG:HH21	1:A:96:ARG:HB2	1.81	0.42
1:B:143:PRO:O	1:B:145:TRP:CD1	2.72	0.42
1:A:1:MET:HG2	1:A:629:GLU:HG2	2.02	0.42
1:A:558:LEU:HG	1:A:568:TYR:CE1	2.55	0.42
1:B:127:VAL:HG23	1:B:129:VAL:HG23	2.02	0.42
1:A:506:PRO:HG2	1:A:666:VAL:HG21	2.02	0.41
1:A:476:GLY:O	1:A:490:ALA:HA	2.20	0.41
1:B:142:GLY:HA3	1:B:145:TRP:CE2	2.55	0.41
1:A:117:LEU:HD22	1:A:155:LEU:HB2	2.02	0.41
1:A:342:GLN:H	1:A:342:GLN:HG2	1.65	0.41
1:A:473:LEU:HB3	1:A:541:VAL:HG12	2.01	0.41
1:B:451:LEU:HD12	1:B:451:LEU:HA	1.87	0.41
1:A:83:GLY:CA	1:B:83:GLY:HA3	2.51	0.41
1:B:205:LEU:HD11	1:B:245:ALA:HB2	2.02	0.41
1:A:7:THR:OG1	1:A:8:GLU:N	2.54	0.41
1:A:590:ASP:OD2	3:D:19:DG:N1	2.33	0.41
1:A:189:LEU:HA	1:A:190:PRO:HD2	1.93	0.41
1:B:348:LEU:HB2	1:B:357:TRP:CE2	2.56	0.41
1:B:649:PHE:HA	1:B:650:PRO:HD3	1.91	0.41
1:A:43:TYR:N	1:A:44:PRO:HD2	2.36	0.41
1:B:45:LEU:HD13	1:B:81:ARG:HD3	2.03	0.41
1:B:640:ARG:HG3	1:B:649:PHE:CD1	2.56	0.40
1:A:54:GLY:O	1:A:57:THR:HG23	2.22	0.40
1:B:297:ILE:HG23	1:B:301:LEU:HD22	2.02	0.40
1:A:51:ARG:NE	4:N:5:DC:OP1	2.50	0.40
1:A:217:LEU:O	1:A:218:PRO:C	2.59	0.40
1:B:254:ILE:HA	1:B:255:PRO:HD3	1.94	0.40
1:B:604:LEU:HD21	1:B:614:PRO:HB2	2.04	0.40
1:A:283:TRP:HZ3	1:A:287:ARG:HH21	1.69	0.40
1:B:117:LEU:HD23	1:B:132:LEU:HD22	2.01	0.40
1:B:123:ARG:NH2	7:B:949:HOH:O	2.53	0.40
1:A:494:VAL:HG22	1:A:500:HIS:HB2	2.04	0.40
2:E:4:DG:H2'	2:E:5:DG:C8	2.57	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	670/685 (98%)	657 (98%)	13 (2%)	0	100	100
1	B	668/685 (98%)	656 (98%)	12 (2%)	0	100	100
All	All	1338/1370 (98%)	1313 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	521/549 (95%)	481 (92%)	40 (8%)	13	20
1	B	520/549 (95%)	474 (91%)	46 (9%)	10	15
All	All	1041/1098 (95%)	955 (92%)	86 (8%)	11	17

All (86) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	46	LEU
1	A	57	THR
1	A	64	LEU
1	A	155	LEU
1	A	165	LEU
1	A	179	LEU
1	A	199	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	210	GLU
1	A	217	LEU
1	A	223	LEU
1	A	234	GLN
1	A	321	LEU
1	A	324	ARG
1	A	335	ARG
1	A	346	LEU
1	A	355	GLN
1	A	366	LEU
1	A	392	ARG
1	A	406	VAL
1	A	426	LEU
1	A	451	LEU
1	A	454	LEU
1	A	475	VAL
1	A	484	SER
1	A	505	LEU
1	A	531	ARG
1	A	543	LEU
1	A	556	LEU
1	A	558	LEU
1	A	570	LEU
1	A	580	ARG
1	A	585	GLN
1	A	604	LEU
1	A	608	ARG
1	A	639	THR
1	A	640	ARG
1	A	641	LEU
1	A	676	GLU
1	A	680	GLU
1	B	3	HIS
1	B	46	LEU
1	B	49	VAL
1	B	52	ARG
1	B	57	THR
1	B	64	LEU
1	B	98	LEU
1	B	155	LEU
1	B	165	LEU
1	B	179	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	199	ARG
1	B	215	LEU
1	B	223	LEU
1	B	264	VAL
1	B	271	HIS
1	B	273	GLU
1	B	277	LEU
1	B	289	ARG
1	B	321	LEU
1	B	324	ARG
1	B	335	ARG
1	B	336	VAL
1	B	346	LEU
1	B	366	LEU
1	B	392	ARG
1	B	406	VAL
1	B	426	LEU
1	B	437	VAL
1	B	440	ARG
1	B	451	LEU
1	B	454	LEU
1	B	475	VAL
1	B	505	LEU
1	B	531	ARG
1	B	543	LEU
1	B	545	ARG
1	B	548	ARG
1	B	558	LEU
1	B	570	LEU
1	B	604	LEU
1	B	622	GLU
1	B	628	LEU
1	B	639	THR
1	B	640	ARG
1	B	641	LEU
1	B	680	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 6 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	DT	A	701	-	22,22,22	0.66	0	33,33,33	0.92	1 (3%)
5	DT	B	701	-	22,22,22	0.60	0	33,33,33	0.82	1 (3%)

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	DT	A	701	-	-	0/10/22/22	0/2/2/2
5	DT	B	701	-	-	3/10/22/22	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	701	DT	O4-C4-C5	-3.13	121.27	124.90
5	B	701	DT	O4-C4-C5	-2.62	121.86	124.90

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	701	DT	C5'-O5'-P-OP3
5	B	701	DT	C5'-O5'-P-OP1
5	B	701	DT	C5'-O5'-P-OP2

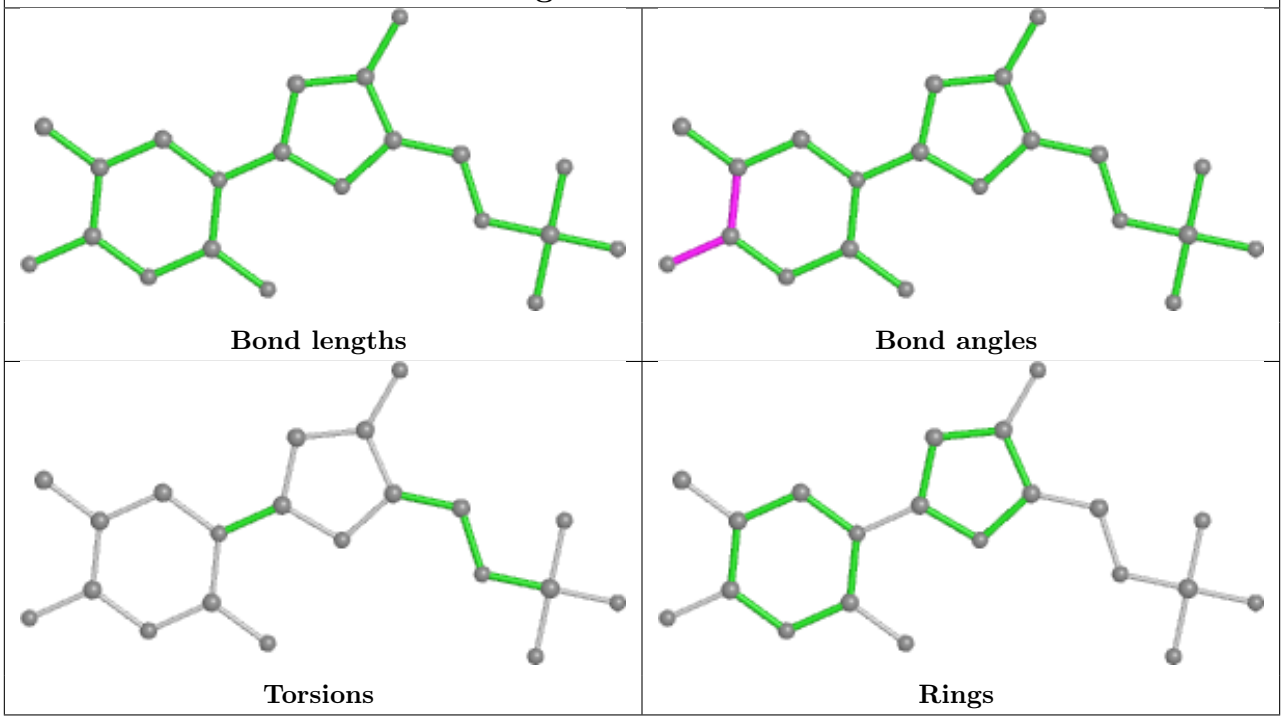
There are no ring outliers.

1 monomer is involved in 1 short contact:

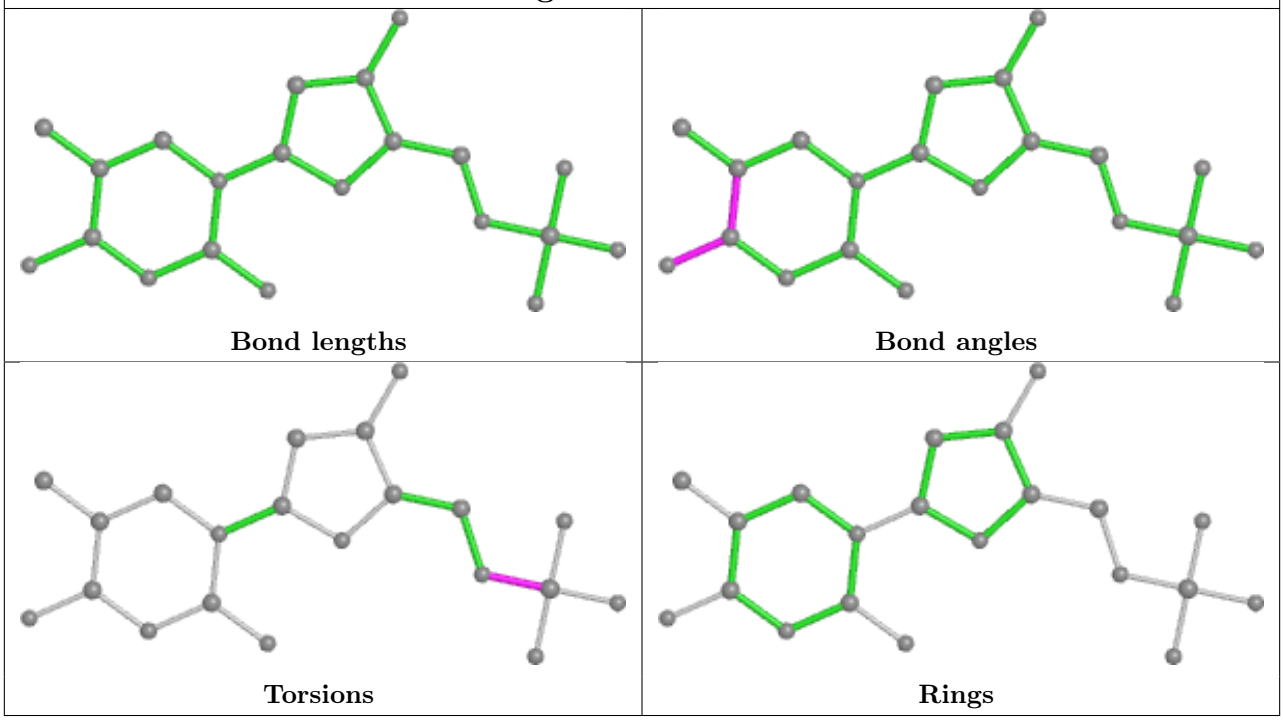
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	701	DT	1	0

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 DT A 701



Ligand DT B 701



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	676/685 (98%)	0.16	53 (7%) 13 11	19, 45, 95, 141	0
1	B	675/685 (98%)	0.18	59 (8%) 10 9	20, 40, 111, 140	0
2	C	16/21 (76%)	-0.66	0 100 100	27, 41, 70, 92	0
2	E	16/21 (76%)	-0.62	0 100 100	32, 38, 60, 62	0
3	D	10/10 (100%)	-0.83	0 100 100	36, 46, 52, 59	0
3	F	10/10 (100%)	-0.65	0 100 100	37, 56, 59, 62	0
4	M	6/9 (66%)	-0.29	0 100 100	39, 53, 76, 79	0
4	N	6/9 (66%)	-0.13	0 100 100	38, 60, 89, 94	0
All	All	1415/1450 (97%)	0.13	112 (7%) 12 11	19, 42, 101, 141	0

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	205	LEU	7.8
1	B	207	LEU	7.7
1	B	189	LEU	7.7
1	A	79	LEU	7.5
1	B	223	LEU	7.1
1	B	183	LEU	7.0
1	A	83	GLY	6.7
1	B	243	TRP	6.4
1	A	81	ARG	6.3
1	B	182	TRP	5.8
1	A	82	MET	5.7
1	B	204	LEU	5.6
1	B	208	GLY	5.3
1	B	216	PRO	5.2
1	B	242	ALA	5.1
1	A	88	TYR	5.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	209	GLU	5.0
1	A	42	VAL	4.9
1	B	271	HIS	4.7
1	B	179	LEU	4.6
1	A	91	TYR	4.6
1	A	46	LEU	4.6
1	B	215	LEU	4.6
1	A	49	VAL	4.5
1	B	212	PRO	4.4
1	B	187	HIS	4.4
1	A	37	PRO	4.4
1	A	32	VAL	4.3
1	A	84	GLN	4.3
1	B	226	TYR	4.2
1	A	28	ARG	4.2
1	A	183	LEU	4.2
1	B	221	LEU	4.1
1	B	244	VAL	4.1
1	B	241	VAL	4.0
1	B	240	ARG	4.0
1	B	255	PRO	4.0
1	B	186	GLY	3.9
1	B	238	GLY	3.8
1	B	261	LEU	3.8
1	B	277	LEU	3.7
1	B	214	GLU	3.7
1	B	246	ASP	3.6
1	B	180	GLU	3.5
1	A	85	THR	3.5
1	B	239	GLY	3.5
1	B	260	LEU	3.5
1	A	189	LEU	3.5
1	B	257	LEU	3.5
1	A	80	ALA	3.4
1	A	36	PRO	3.4
1	B	206	ARG	3.4
1	B	270	LEU	3.4
1	B	200	ARG	3.4
1	A	182	TRP	3.4
1	A	86	TYR	3.3
1	A	96	ARG	3.2
1	B	222	SER	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	30	GLU	3.1
1	A	145	TRP	3.1
1	B	203	GLU	3.0
1	B	145	TRP	3.0
1	B	202	TRP	3.0
1	B	220	GLY	3.0
1	B	245	ALA	3.0
1	B	253	PRO	2.9
1	B	258	THR	2.9
1	B	196	ALA	2.8
1	A	184	ALA	2.8
1	B	191	LYS	2.8
1	A	34	ASP	2.8
1	B	184	ALA	2.7
1	A	33	LEU	2.7
1	B	278	ALA	2.7
1	B	185	GLN	2.7
1	B	232	ARG	2.6
1	B	193	VAL	2.6
1	B	219	GLY	2.6
1	A	43	TYR	2.6
1	A	90	LEU	2.5
1	A	270	LEU	2.5
1	A	185	GLN	2.4
1	B	267	LEU	2.4
1	A	223	LEU	2.4
1	A	60	MET	2.4
1	A	61	GLY	2.4
1	A	212	PRO	2.4
1	A	62	ASP	2.4
1	B	254	ILE	2.4
1	A	93	LYS	2.4
1	B	265	LEU	2.3
1	A	39	ARG	2.3
1	A	38	GLY	2.3
1	B	272	GLU	2.3
1	A	215	LEU	2.3
1	A	179	LEU	2.3
1	A	29	LEU	2.2
1	A	224	LEU	2.2
1	A	269	ASP	2.2
1	A	235	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	209	GLU	2.2
1	A	44	PRO	2.1
1	B	137	ARG	2.1
1	A	143	PRO	2.1
1	B	84	GLN	2.1
1	A	180	GLU	2.1
1	A	240	ARG	2.1
1	A	31	VAL	2.1
1	A	45	LEU	2.0
1	A	100	PRO	2.0
1	A	234	GLN	2.0
1	B	177	MET	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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.

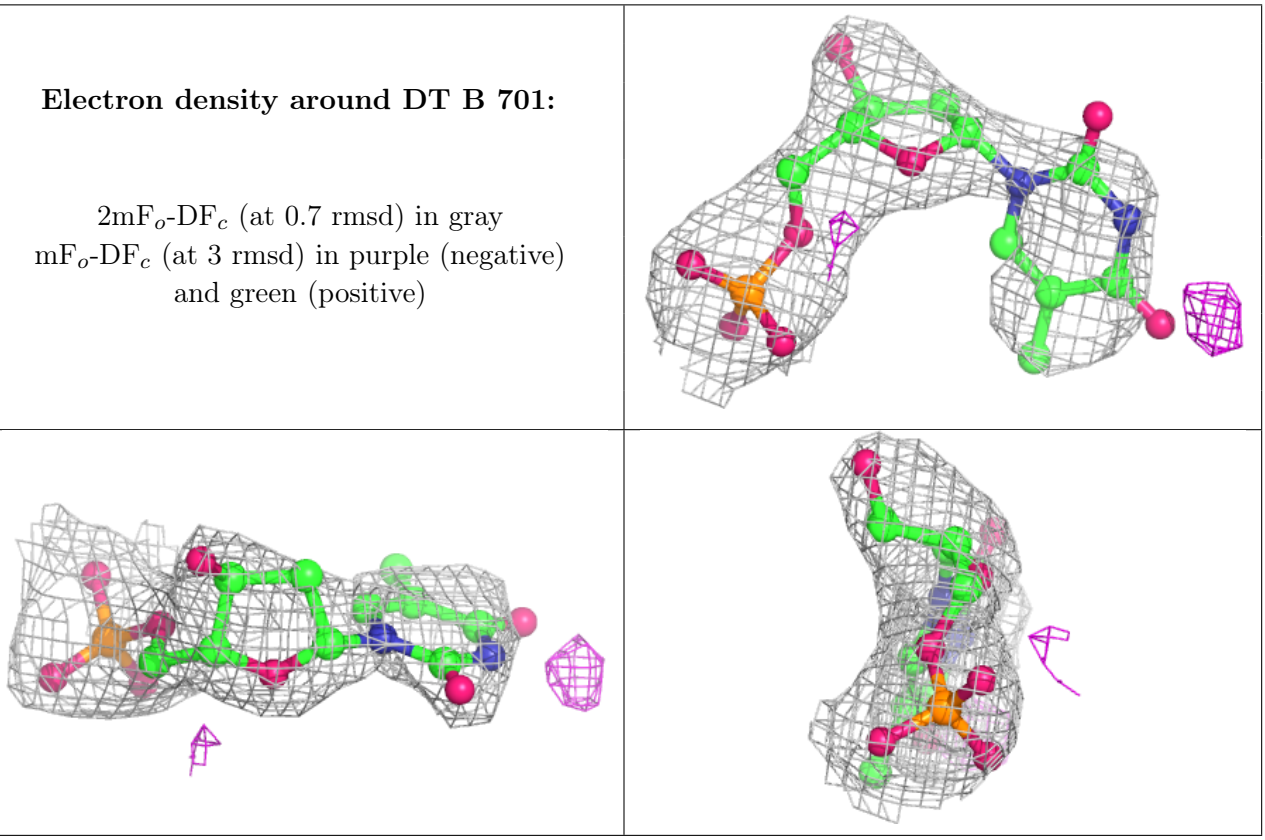
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	DT	B	701	21/21	0.93	0.37	82,102,118,122	0
5	DT	A	701	21/21	0.97	0.20	43,53,84,93	0
6	MN	A	703	1/1	0.99	0.12	32,32,32,32	0
6	MN	B	702	1/1	0.99	0.11	32,32,32,32	0
6	MN	B	703	1/1	0.99	0.09	34,34,34,34	0
6	MN	C	101	1/1	0.99	0.13	28,28,28,28	0
6	MN	A	702	1/1	1.00	0.11	32,32,32,32	0
6	MN	E	101	1/1	1.00	0.13	33,33,33,33	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.

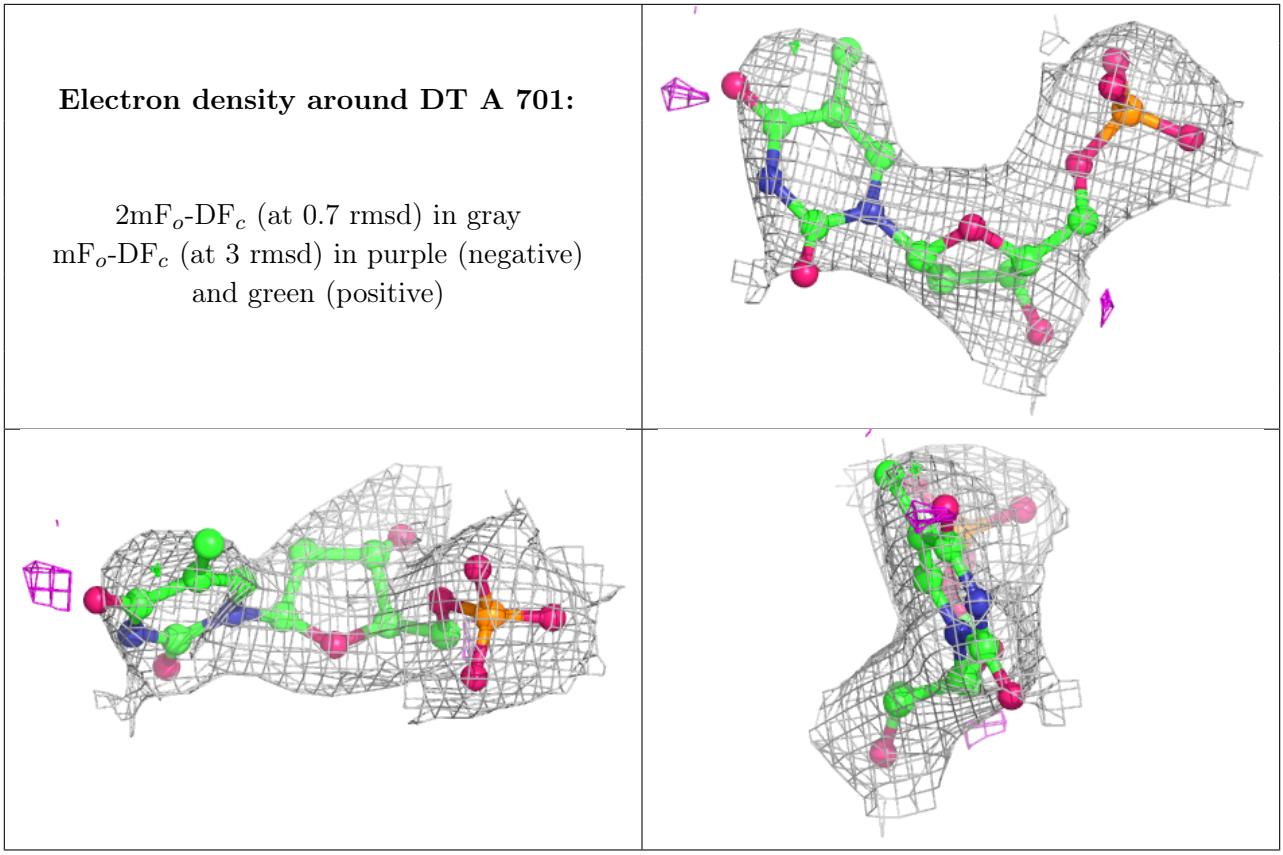
Electron density around DT B 701:

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



Electron density around DT A 701:

$2mF_o$-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.