Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 10:11 am BST

PDB ID	:	6GUF
Title	:	CDK2/CyclinA in complex with CGP74514A
Authors	:	Wood, D.J.; Korolchuk, S.; Tatum, N.J.; Wang, L.Z.; Endicott, J.A.; Noble, M.E.M.; Martin, M.P.
Deposited on	:	2018-06-19
Resolution	:	2.65 Å(reported)

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

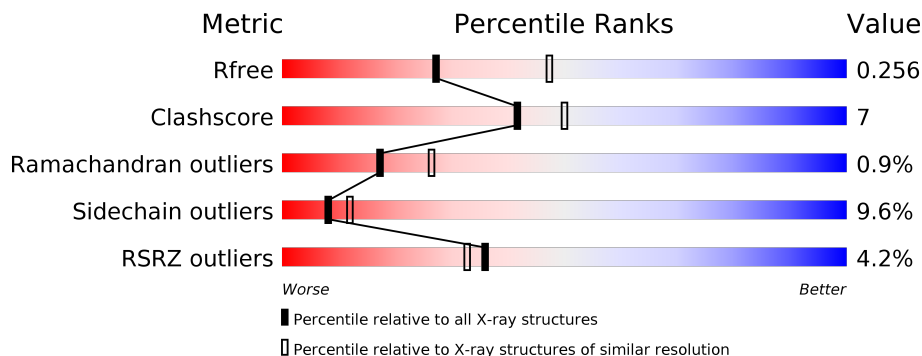|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.65 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		0.256
Clashscore		7
Ramachandran outliers		0.9%
Sidechain outliers		9.6%
RSRZ outliers		4.2%

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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 on 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	302	2% · 81% · 13% · ·
1	C	302	12% · 78% · 16% · 5% ·
2	B	268	82% · 13% · ·
2	D	268	% · 84% · 11% · ·

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	23D	A	301	X	-	-	-
3	23D	C	301	X	-	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9172 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 Cyclin-dependent kinase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	296	Total 2383	C 1549	N 406	O 419	P 1	S 8	0	0	0
1	C	298	Total 2399	C 1556	N 408	O 426	P 1	S 8	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P24941
A	-2	PRO	-	expression tag	UNP P24941
A	-1	GLY	-	expression tag	UNP P24941
A	0	SER	-	expression tag	UNP P24941
C	-3	GLY	-	expression tag	UNP P24941
C	-2	PRO	-	expression tag	UNP P24941
C	-1	GLY	-	expression tag	UNP P24941
C	0	SER	-	expression tag	UNP P24941

- Molecule 2 is a protein called Cyclin-A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	261	Total 2106	C 1364	N 343	O 389	S 10	0	0	0
2	D	261	Total 2106	C 1364	N 343	O 389	S 10	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	171	GLY	-	expression tag	UNP P30274
B	433	HIS	-	expression tag	UNP P30274
B	434	HIS	-	expression tag	UNP P30274
B	435	HIS	-	expression tag	UNP P30274

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	436	HIS	-	expression tag	UNP P30274
B	437	HIS	-	expression tag	UNP P30274
B	438	HIS	-	expression tag	UNP P30274
D	171	GLY	-	expression tag	UNP P30274
D	433	HIS	-	expression tag	UNP P30274
D	434	HIS	-	expression tag	UNP P30274
D	435	HIS	-	expression tag	UNP P30274
D	436	HIS	-	expression tag	UNP P30274
D	437	HIS	-	expression tag	UNP P30274
D	438	HIS	-	expression tag	UNP P30274

- Molecule 3 is N2-[(1R,2S)-2-AMINOCYCLOHEXYL]-N6-(3-CHLOROPHENYL)-9-ETH YL-9H-PURINE-2,6-DIAMINE (three-letter code: 23D) (formula: $C_{19}H_{24}ClN_7$) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 27	C 19	Cl 1	N 7	0	0
3	C	1	Total 27	C 19	Cl 1	N 7	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total 47	O 47	0	0

Continued on next page...
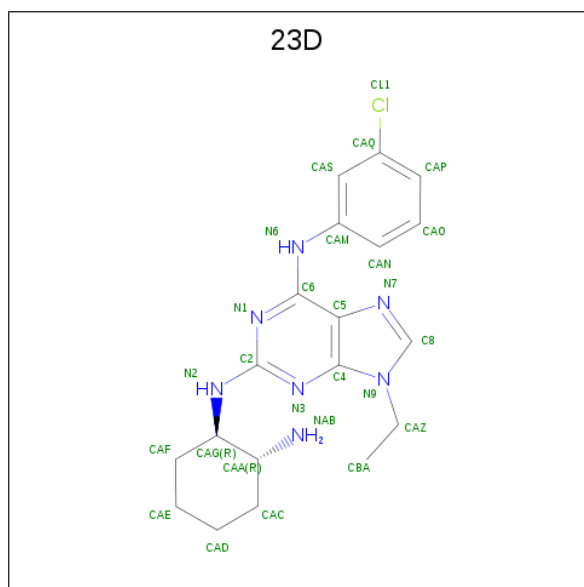
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	36	Total 36	O 36	0	0
4	C	23	Total 23	O 23	0	0
4	D	18	Total 18	O 18	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: Cyclin-dependent kinase 2
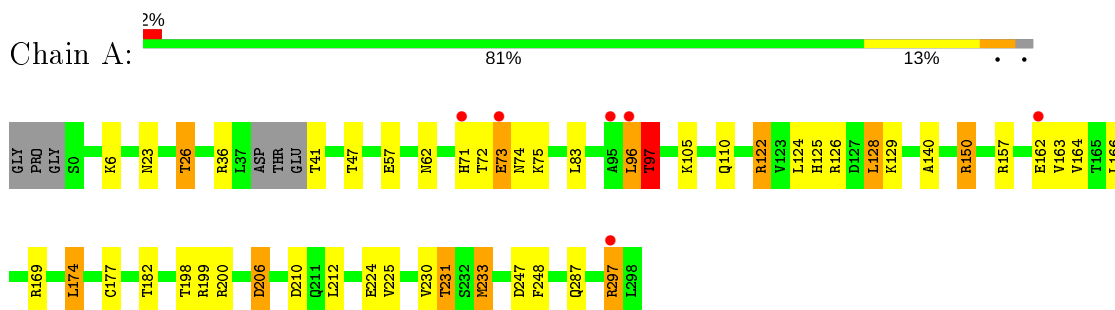
- Molecule 1: Cyclin-dependent kinase 2
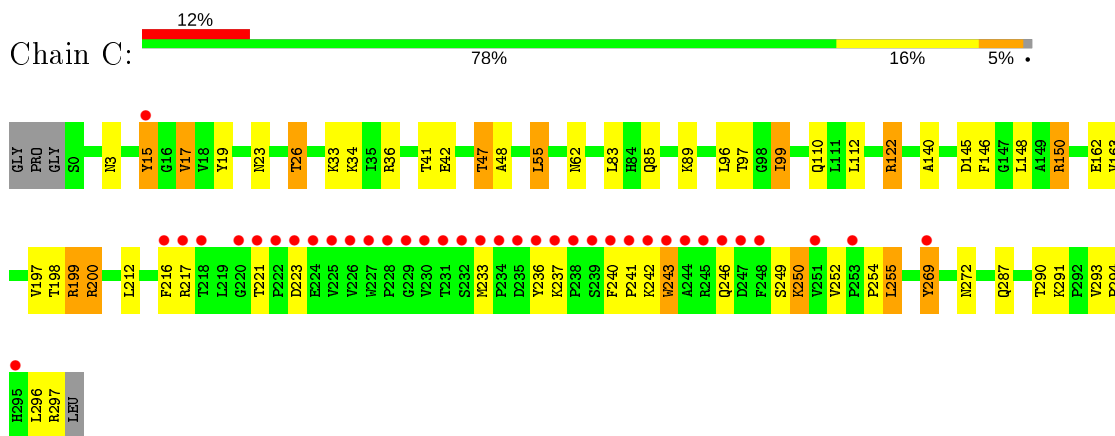
- Molecule 2: Cyclin-A2

- Molecule 2: Cyclin-A2

Chain D:
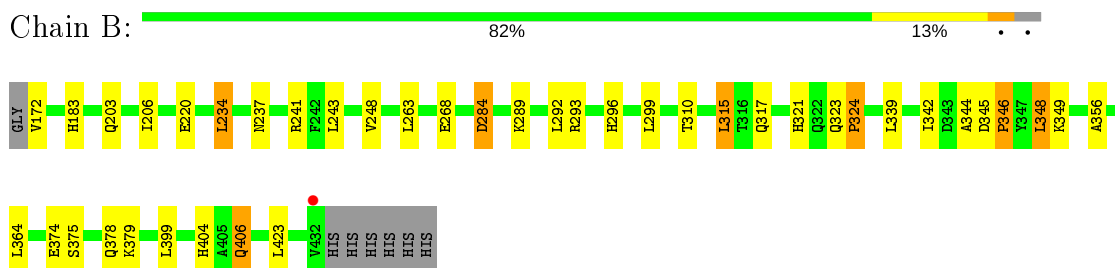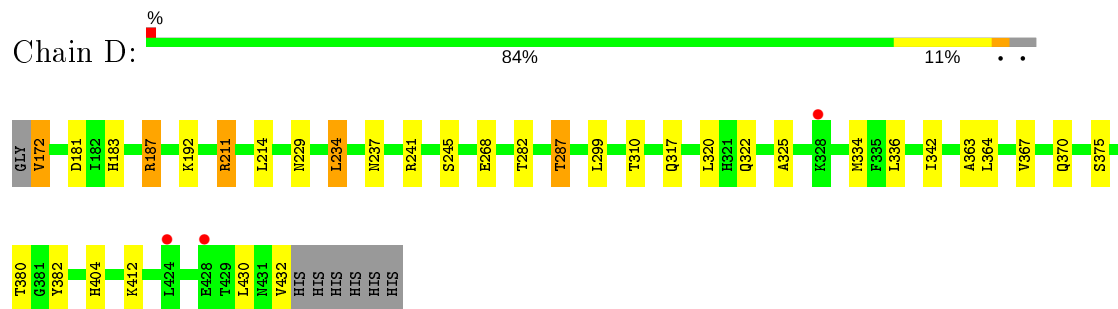
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.32Å 136.69Å 150.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	101.17 – 2.65 101.17 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (101.17-2.65) 99.9 (101.17-2.65)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.203 , 0.255 0.207 , 0.256	Depositor DCC
R_{free} test set	2244 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	40.3	Xtriage
Anisotropy	0.564	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.8	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9172	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: TPO, 23D

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	0.70	0/2432	0.86	0/3297
1	C	0.65	0/2449	0.88	0/3322
2	B	0.67	0/2156	0.83	0/2932
2	D	0.63	0/2156	0.79	0/2932
All	All	0.66	0/9193	0.84	0/12483

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2383	0	2436	33	0
1	C	2399	0	2443	45	0
2	B	2106	0	2125	34	0
2	D	2106	0	2125	24	0
3	A	27	0	24	0	0
3	C	27	0	24	0	0
4	A	47	0	0	2	0
4	B	36	0	0	2	0
4	C	23	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	18	0	0	2	0
All	All	9172	0	9177	128	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (128) 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)
2:B:323:GLN:CB	2:B:324:PRO:HD3	1.54	1.34
1:C:15:TYR:CE2	1:C:47:THR:CG2	2.24	1.20
1:C:15:TYR:CZ	1:C:47:THR:HG23	1.80	1.16
1:C:17:VAL:HG11	1:C:19:TYR:CE1	1.83	1.13
1:C:15:TYR:CD2	1:C:47:THR:CG2	2.33	1.11
2:B:323:GLN:HB2	2:B:324:PRO:HD3	1.20	1.08
1:C:15:TYR:CE2	1:C:47:THR:HG22	1.86	1.08
1:C:17:VAL:CG1	1:C:19:TYR:CE1	2.40	1.04
1:A:125:HIS:CE1	1:A:128:LEU:HD13	1.94	1.02
2:B:323:GLN:CB	2:B:324:PRO:CD	2.37	1.01
2:B:323:GLN:HB2	2:B:324:PRO:CD	1.91	0.99
2:B:323:GLN:HB3	2:B:324:PRO:HD3	1.47	0.96
1:C:17:VAL:HG11	1:C:19:TYR:HE1	1.15	0.95
1:C:17:VAL:CG1	1:C:19:TYR:HE1	1.77	0.94
1:C:15:TYR:CE2	1:C:47:THR:HG23	1.99	0.91
1:A:125:HIS:ND1	1:A:128:LEU:HD13	1.86	0.91
2:D:187:ARG:HH11	2:D:187:ARG:HG2	1.35	0.88
1:C:15:TYR:CD2	1:C:47:THR:HG21	2.09	0.87
1:C:15:TYR:CE1	1:C:47:THR:HG23	2.11	0.85
1:C:249:SER:O	1:C:250:LYS:HG2	1.81	0.81
2:B:237:ASN:HD21	2:B:241:ARG:HH11	1.29	0.80
1:C:15:TYR:CG	1:C:47:THR:HG21	2.16	0.79
2:D:237:ASN:HD21	2:D:241:ARG:HH11	1.29	0.78
1:A:125:HIS:O	1:A:126:ARG:HB2	1.86	0.74
1:C:17:VAL:HG13	1:C:19:TYR:CE1	2.23	0.73
1:A:96:LEU:O	1:A:97:THR:HG23	1.89	0.72
2:D:187:ARG:HG2	2:D:187:ARG:NH1	2.02	0.72
2:D:363:ALA:O	2:D:367:VAL:HG12	1.89	0.72
1:A:26:THR:HG21	4:A:406:HOH:O	1.89	0.72
2:B:323:GLN:OE1	2:B:323:GLN:N	2.24	0.71
2:B:234:LEU:HD13	2:B:310:THR:HG21	1.74	0.70
1:A:297:ARG:HG2	1:A:297:ARG:HH11	1.58	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:15:TYR:CD2	1:C:47:THR:HG22	2.11	0.69
1:A:110:GLN:HE22	1:A:140:ALA:HA	1.56	0.68
2:B:346:PRO:HB2	2:B:349:LYS:NZ	2.09	0.68
1:C:15:TYR:CG	1:C:47:THR:CG2	2.75	0.66
1:C:110:GLN:HE22	1:C:140:ALA:HA	1.60	0.66
2:D:234:LEU:HD13	2:D:310:THR:HG21	1.77	0.66
1:C:197:VAL:HG11	1:C:252:VAL:HG13	1.78	0.65
1:C:15:TYR:CZ	1:C:47:THR:CG2	2.57	0.65
1:A:150:ARG:NH1	2:B:268:GLU:O	2.29	0.64
2:D:187:ARG:NH2	2:D:382:TYR:CZ	2.67	0.63
1:A:125:HIS:CE1	1:A:128:LEU:CD1	2.78	0.62
1:A:224:GLU:OE2	1:A:231:THR:HB	2.00	0.61
1:C:99:ILE:CD1	1:C:199:ARG:HD2	2.30	0.61
1:C:99:ILE:HD13	1:C:199:ARG:HD2	1.83	0.60
1:C:122:ARG:O	1:C:122:ARG:HG2	2.02	0.59
2:B:346:PRO:HB2	2:B:349:LYS:HZ2	1.66	0.59
1:C:150:ARG:NH1	2:D:268:GLU:O	2.34	0.59
1:A:122:ARG:O	1:A:122:ARG:HG2	2.03	0.59
1:A:177:CYS:SG	1:A:233:MET:HG2	2.43	0.59
1:A:71:HIS:O	1:A:73:GLU:N	2.36	0.58
1:C:198:THR:HG22	1:C:200:ARG:HG2	1.85	0.58
2:D:237:ASN:ND2	2:D:241:ARG:HH11	2.00	0.58
1:A:125:HIS:CG	1:A:128:LEU:HD13	2.38	0.58
2:B:203:GLN:HG2	2:B:206:ILE:HG12	1.87	0.56
2:B:234:LEU:HD13	2:B:310:THR:CG2	2.35	0.56
2:D:183:HIS:HB2	2:D:317:GLN:HE22	1.72	0.55
1:C:23:ASN:HB3	1:C:26:THR:HG22	1.89	0.55
1:A:62:ASN:HD21	1:A:110:GLN:HE21	1.53	0.55
2:B:172:VAL:N	4:B:501:HOH:O	2.39	0.55
2:B:237:ASN:ND2	2:B:241:ARG:HH11	2.01	0.54
1:C:62:ASN:HD21	1:C:110:GLN:HE21	1.56	0.54
2:D:187:ARG:NH2	2:D:382:TYR:CE1	2.76	0.54
1:A:57:GLU:OE1	1:A:122:ARG:NH2	2.35	0.53
2:D:234:LEU:HD13	2:D:310:THR:CG2	2.38	0.53
2:D:380:THR:HG21	2:D:382:TYR:CD2	2.45	0.52
2:B:234:LEU:CD1	2:B:310:THR:CG2	2.87	0.52
1:C:17:VAL:HG13	1:C:19:TYR:CD1	2.45	0.52
1:C:17:VAL:HG12	1:C:34:LYS:HB3	1.93	0.51
2:B:203:GLN:HG2	2:B:206:ILE:CG1	2.41	0.51
2:D:404:HIS:CD2	4:D:511:HOH:O	2.64	0.51
2:D:364:LEU:HD22	2:D:370:GLN:HB2	1.94	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:71:HIS:CD2	2:B:296:HIS:NE2	2.80	0.49
1:C:252:VAL:HG11	1:C:255:LEU:HD22	1.94	0.49
1:A:126:ARG:HD2	1:A:163:VAL:HG11	1.94	0.49
2:B:183:HIS:HB2	2:B:317:GLN:HE22	1.77	0.49
1:A:71:HIS:CD2	2:B:296:HIS:CE1	3.01	0.48
1:C:15:TYR:OH	1:C:48:ALA:HA	2.13	0.48
1:A:162:GLU:HA	4:A:403:HOH:O	2.13	0.48
1:C:216:PHE:HE1	1:C:269:TYR:HH	1.62	0.48
2:B:345:ASP:HA	2:B:346:PRO:HA	1.54	0.47
2:D:234:LEU:CD1	2:D:310:THR:CG2	2.92	0.47
1:A:23:ASN:HB3	1:A:26:THR:HG22	1.97	0.47
1:A:206:ASP:OD1	1:A:210:ASP:OD2	2.31	0.47
2:B:293:ARG:NH2	1:C:3:ASN:ND2	2.62	0.47
2:D:287:THR:HG21	4:D:517:HOH:O	2.14	0.47
2:D:211:ARG:HE	2:D:211:ARG:HB3	1.42	0.46
1:A:157:ARG:NH1	2:B:268:GLU:OE2	2.46	0.46
2:B:404:HIS:CE1	2:B:406:GLN:HG2	2.51	0.46
2:D:380:THR:CG2	2:D:382:TYR:CD2	2.97	0.46
1:A:297:ARG:NH1	1:A:297:ARG:HG2	2.29	0.46
2:B:203:GLN:HE21	2:B:248:VAL:H	1.63	0.46
1:A:36:ARG:HD3	1:A:75:LYS:HG2	1.99	0.45
1:A:169:ARG:HD2	1:A:174:LEU:HD13	1.98	0.45
1:A:297:ARG:CG	1:A:297:ARG:HH11	2.26	0.45
2:B:315:LEU:HD13	2:B:356:ALA:HB1	1.98	0.44
2:D:432:VAL:O	2:D:432:VAL:HG12	2.17	0.44
1:A:124:LEU:HD21	1:A:182:THR:HA	1.98	0.44
1:C:197:VAL:HG11	1:C:252:VAL:CG1	2.46	0.44
1:C:272:ASN:ND2	2:D:172:VAL:HG22	2.32	0.44
2:B:284:ASP:OD1	2:B:284:ASP:N	2.50	0.44
1:C:55:LEU:HD11	1:C:146:PHE:CD1	2.53	0.44
1:C:293:VAL:HG12	1:C:294:PRO:O	2.18	0.44
1:C:15:TYR:CD1	1:C:47:THR:CG2	3.01	0.44
2:D:322:GLN:HG3	2:D:325:ALA:HA	2.00	0.43
2:D:234:LEU:HD12	2:D:234:LEU:HA	1.78	0.43
2:B:378:GLN:HG3	4:B:515:HOH:O	2.18	0.43
2:D:187:ARG:NH1	2:D:187:ARG:CG	2.73	0.43
1:A:230:VAL:HA	1:A:233:MET:SD	2.58	0.43
1:C:237:LYS:HB3	1:C:240:PHE:CD1	2.54	0.43
1:A:125:HIS:O	1:A:126:ARG:CB	2.61	0.43
2:B:344:ALA:HB1	2:B:348:LEU:HD22	2.00	0.43
2:B:234:LEU:HD12	2:B:234:LEU:HA	1.78	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:221:THR:HG21	1:C:241:PRO:HB2	1.99	0.42
2:D:229:ASN:HD21	2:D:412:LYS:NZ	2.17	0.42
1:C:290:THR:HG22	1:C:291:LYS:H	1.84	0.42
1:A:71:HIS:NE2	2:B:296:HIS:NE2	2.67	0.42
1:C:249:SER:O	1:C:250:LYS:CG	2.62	0.42
2:B:346:PRO:CB	2:B:349:LYS:HZ1	2.32	0.42
1:C:216:PHE:HB3	1:C:221:THR:HG22	2.01	0.42
1:C:216:PHE:HE1	1:C:269:TYR:OH	2.03	0.42
2:B:315:LEU:HA	2:B:315:LEU:HD12	1.83	0.41
1:C:221:THR:OG1	1:C:243:TRP:N	2.49	0.41
1:A:297:ARG:NH1	1:A:297:ARG:CG	2.83	0.41
1:A:198:THR:HG22	1:A:200:ARG:HG3	2.02	0.40
2:B:321:HIS:HD2	2:B:375:SER:OG	2.03	0.40
1:C:290:THR:HG22	1:C:291:LYS:N	2.36	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	291/302 (96%)	275 (94%)	13 (4%)	3 (1%)	15	23
1	C	295/302 (98%)	280 (95%)	10 (3%)	5 (2%)	9	13
2	B	259/268 (97%)	252 (97%)	5 (2%)	2 (1%)	19	29
2	D	259/268 (97%)	255 (98%)	4 (2%)	0	100	100
All	All	1104/1140 (97%)	1062 (96%)	32 (3%)	10 (1%)	17	26

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	324	PRO
1	A	164	VAL
1	C	163	VAL
1	C	250	LYS
1	C	254	PRO
1	C	145	ASP
1	C	255	LEU
1	A	97	THR
2	B	346	PRO

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	260/264 (98%)	234 (90%)	26 (10%)	7	10
1	C	262/264 (99%)	228 (87%)	34 (13%)	4	5
2	B	234/240 (98%)	216 (92%)	18 (8%)	13	20
2	D	234/240 (98%)	217 (93%)	17 (7%)	14	21
All	All	990/1008 (98%)	895 (90%)	95 (10%)	8	12

All (95) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	LYS
1	A	26	THR
1	A	41	THR
1	A	47	THR
1	A	73	GLU
1	A	74	ASN
1	A	83	LEU
1	A	96	LEU
1	A	97	THR
1	A	105	LYS
1	A	122	ARG
1	A	128	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	129	LYS
1	A	150	ARG
1	A	166	LEU
1	A	174	LEU
1	A	199	ARG
1	A	206	ASP
1	A	212	LEU
1	A	225	VAL
1	A	231	THR
1	A	233	MET
1	A	247	ASP
1	A	248	PHE
1	A	287	GLN
1	A	297	ARG
2	B	220	GLU
2	B	234	LEU
2	B	243	LEU
2	B	263	LEU
2	B	284	ASP
2	B	289	LYS
2	B	292	LEU
2	B	299	LEU
2	B	315	LEU
2	B	339	LEU
2	B	342	ILE
2	B	348	LEU
2	B	364	LEU
2	B	374	GLU
2	B	379	LYS
2	B	399	LEU
2	B	406	GLN
2	B	423	LEU
1	C	15	TYR
1	C	17	VAL
1	C	26	THR
1	C	33	LYS
1	C	36	ARG
1	C	41	THR
1	C	42	GLU
1	C	47	THR
1	C	55	LEU
1	C	83	LEU

Continued from previous page...

Mol	Chain	Res	Type
1	C	85	GLN
1	C	89	LYS
1	C	96	LEU
1	C	97	THR
1	C	99	ILE
1	C	112	LEU
1	C	122	ARG
1	C	148	LEU
1	C	150	ARG
1	C	162	GLU
1	C	199	ARG
1	C	200	ARG
1	C	212	LEU
1	C	217	ARG
1	C	223	ASP
1	C	233	MET
1	C	236	TYR
1	C	242	LYS
1	C	243	TRP
1	C	246	GLN
1	C	269	TYR
1	C	287	GLN
1	C	296	LEU
1	C	297	ARG
2	D	172	VAL
2	D	181	ASP
2	D	187	ARG
2	D	192	LYS
2	D	211	ARG
2	D	214	LEU
2	D	234	LEU
2	D	245	SER
2	D	282	THR
2	D	287	THR
2	D	299	LEU
2	D	320	LEU
2	D	334	MET
2	D	336	LEU
2	D	342	ILE
2	D	375	SER
2	D	430	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (25) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	59	ASN
1	A	74	ASN
1	A	110	GLN
1	A	287	GLN
2	B	179	HIS
2	B	208	ASN
2	B	233	HIS
2	B	237	ASN
2	B	317	GLN
2	B	321	HIS
2	B	395	HIS
1	C	3	ASN
1	C	85	GLN
1	C	110	GLN
1	C	272	ASN
2	D	208	ASN
2	D	229	ASN
2	D	233	HIS
2	D	237	ASN
2	D	296	HIS
2	D	312	ASN
2	D	317	GLN
2	D	321	HIS
2	D	404	HIS
2	D	431	ASN

5.3.3   RNA (i)

There are no RNA molecules in this entry.

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

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	#$\|Z\|>2$
1	TPO	A	160	1	8,10,11	0.87	0	10,14,16	1.66	1 (10%)
1	TPO	C	160	1	8,10,11	0.85	0	10,14,16	1.54	1 (10%)

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	TPO	A	160	1	-	1/9/11/13	-
1	TPO	C	160	1	-	1/9/11/13	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	160	TPO	P-OG1-CB	-3.86	111.55	123.21
1	C	160	TPO	P-OG1-CB	-3.21	113.52	123.21

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	160	TPO	O-C-CA-CB
1	A	160	TPO	O-C-CA-CB

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

2 ligands 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	#$|Z| > 2$
3	23D	A	301	-	26,30,30	0.91	2 (7%)	29,42,42	1.49	4 (13%)
3	23D	C	301	-	26,30,30	0.92	1 (3%)	29,42,42	1.10	2 (6%)

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
3	23D	A	301	-	1/1/2/4	1/10/21/21	0/4/4/4
3	23D	C	301	-	1/1/2/4	0/10/21/21	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	301	23D	CAZ-N9	-3.60	1.46	1.49
3	A	301	23D	CAZ-N9	-3.04	1.46	1.49
3	A	301	23D	C6-N1	2.29	1.36	1.32

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	301	23D	N6-C6-N1	5.04	123.39	118.66
3	A	301	23D	C5-C6-N1	-3.61	117.81	120.81
3	C	301	23D	C5-C6-N1	-3.26	118.10	120.81
3	A	301	23D	C2-N3-C4	-2.53	112.41	115.28
3	A	301	23D	CAS-CAQ-CL1	-2.21	116.39	119.15
3	C	301	23D	C2-N3-C4	-2.19	112.79	115.28

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	301	23D	CAA
3	C	301	23D	CAA

All (1) torsion outliers are listed below:

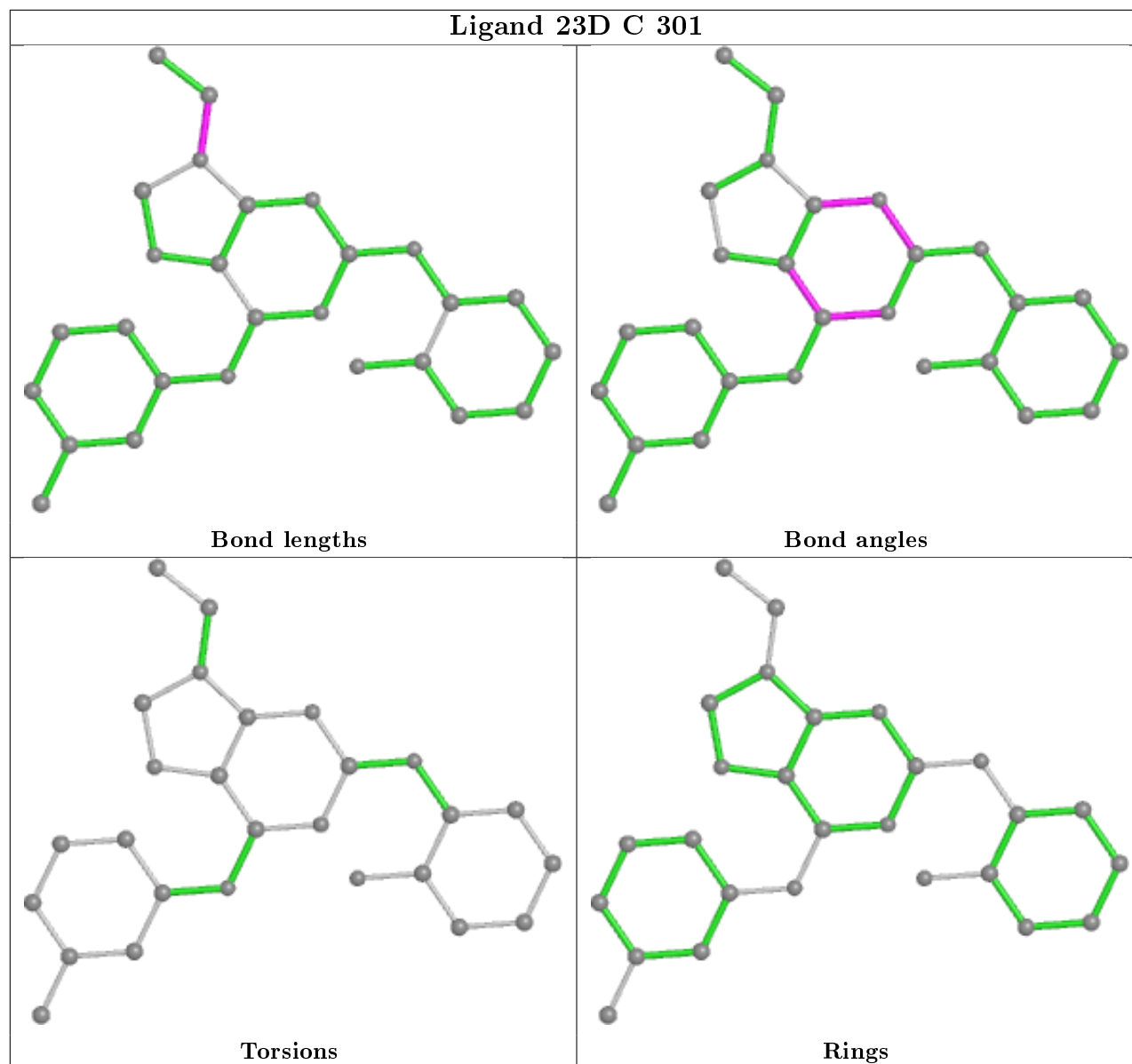| Mol	Chain	Res	Type	Atoms
3	A	301	23D	CBA-CAZ-N9-C4

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 23D A 301

Bond lengths

Bond angles

Torsions

Rings

Ligand 23D C 301

Bond lengths

Bond angles

Torsions

Rings
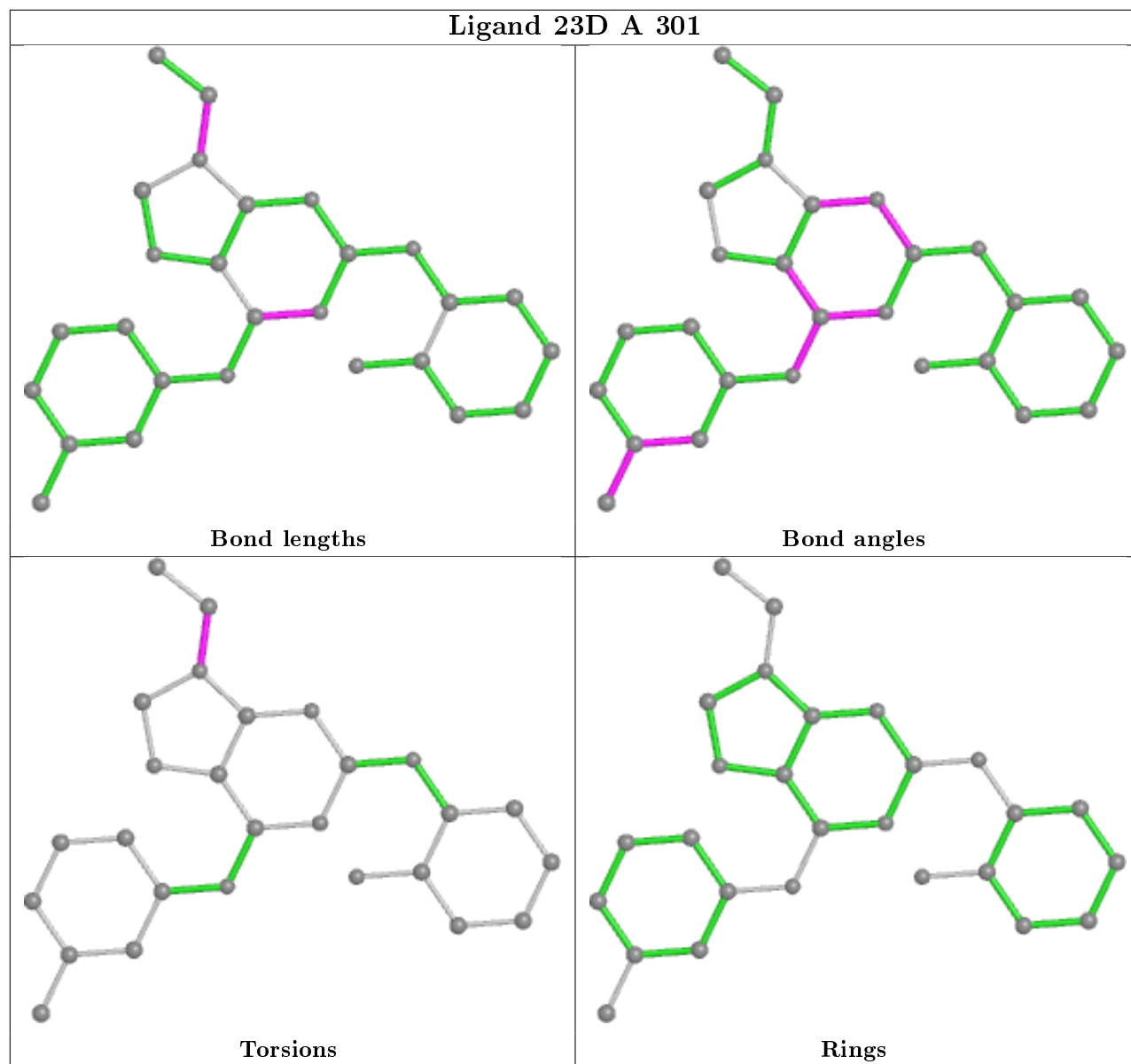
5.7 Other polymers (i)

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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	295/302 (97%)	-0.12	6 (2%)	65	60	20, 33, 72, 141	0
1	C	297/302 (98%)	0.71	37 (12%)	3	2	28, 52, 180, 223	0
2	B	261/268 (97%)	-0.19	1 (0%)	92	93	23, 38, 63, 90	0
2	D	261/268 (97%)	0.15	3 (1%)	80	79	26, 53, 88, 114	0
All	All	1114/1140 (97%)	0.15	47 (4%)	36	33	20, 44, 93, 223	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	PRO	14.1
1	C	241	PRO	10.1
1	C	228	PRO	10.0
1	C	223	ASP	8.9
1	C	234	PRO	8.7
1	C	242	LYS	8.7
1	C	227	TRP	8.3
1	C	233	MET	7.8
1	C	232	SER	7.6
1	C	221	THR	7.4
1	C	243	TRP	7.2
1	C	240	PHE	7.1
1	C	229	GLY	7.0
1	C	231	THR	6.4
1	C	244	ALA	6.2
1	C	230	VAL	5.3
1	C	225	VAL	5.1
1	C	226	VAL	5.1
1	C	245	ARG	4.8
1	C	15	TYR	4.8
1	C	217	ARG	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	238	PRO	4.3
1	C	224	GLU	4.2
1	A	95	ALA	4.1
1	C	236	TYR	4.0
1	C	248	PHE	3.9
1	C	269	TYR	3.7
1	C	235	ASP	3.6
1	C	237	LYS	3.5
1	C	239	SER	3.3
1	C	218	THR	3.3
2	B	432	VAL	3.0
1	C	251	VAL	2.9
1	A	297	ARG	2.9
1	C	253	PRO	2.7
1	A	96	LEU	2.7
1	A	71	HIS	2.6
1	C	246	GLN	2.5
1	C	247	ASP	2.5
1	C	295	HIS	2.3
2	D	424	LEU	2.3
2	D	428	GLU	2.3
2	D	328	LYS	2.3
1	C	216	PHE	2.2
1	A	73	GLU	2.1
1	C	220	GLY	2.1
1	A	162	GLU	2.1

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

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

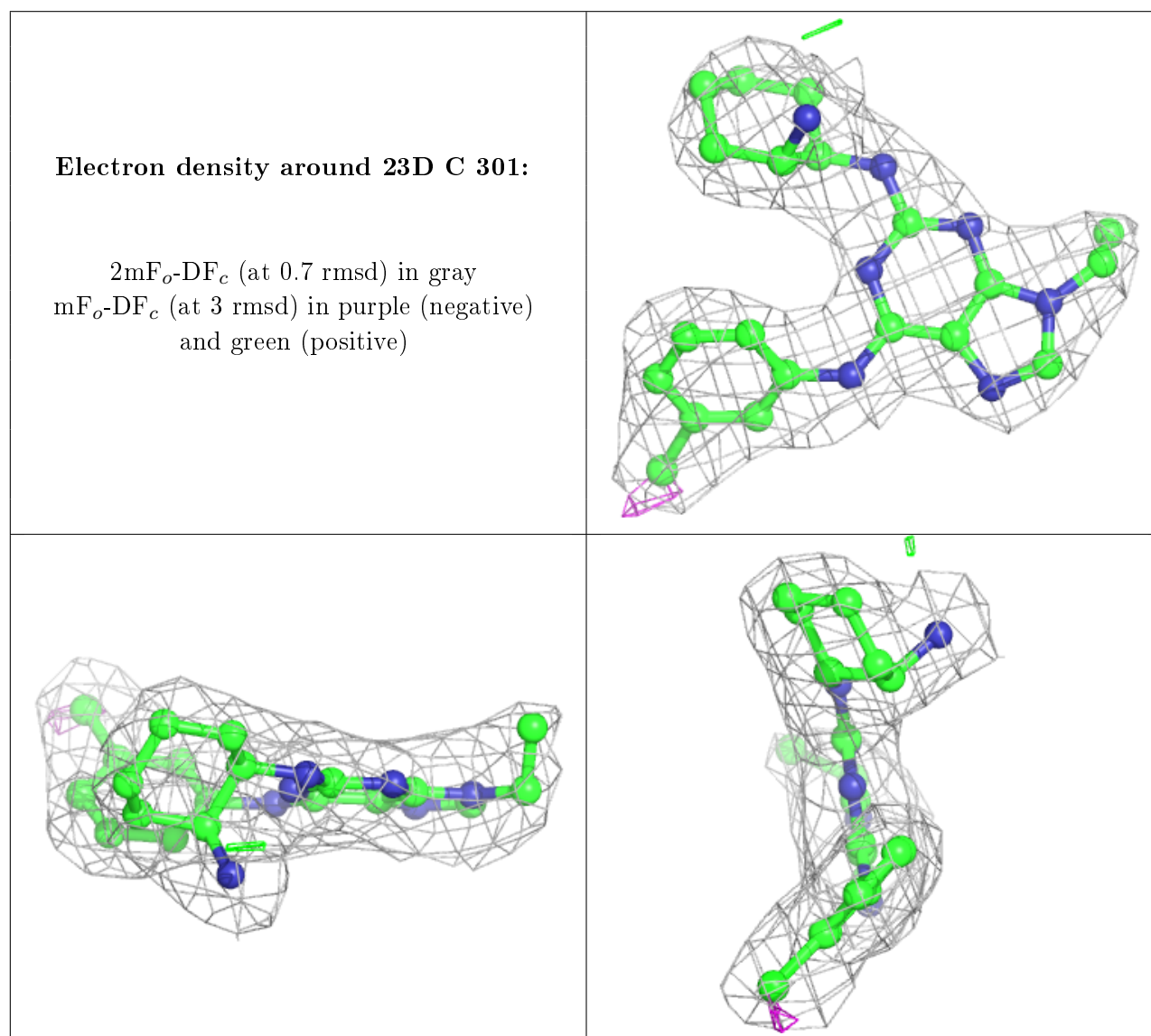| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	TPO	A	160	11/12	0.98	0.15	23,26,28,29	0
1	TPO	C	160	11/12	0.98	0.15	37,44,52,54	0

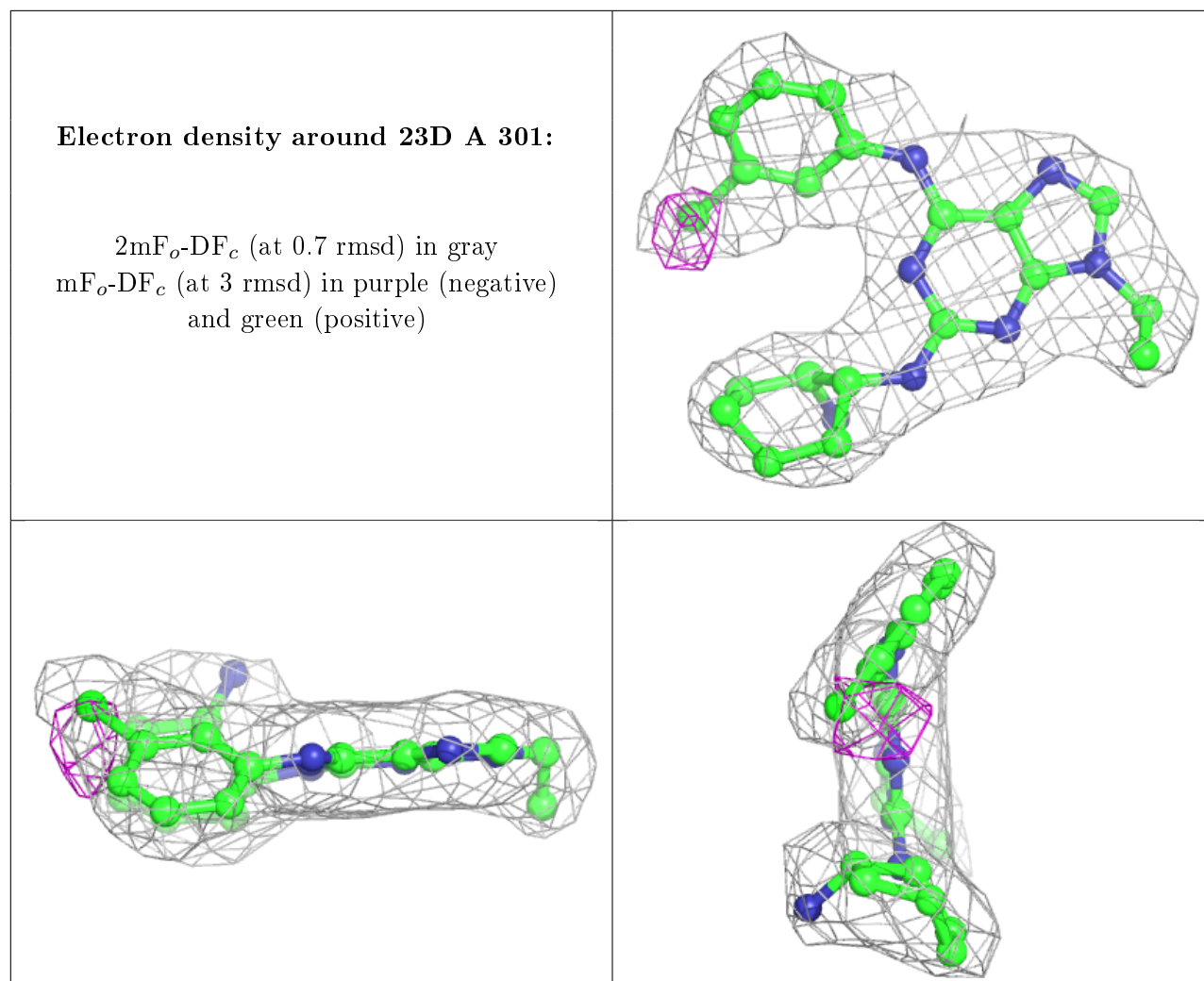## 6.3 Carbohydrates (i)

There are no carbohydrates 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, 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
3	23D	C	301	27/27	0.94	0.18	36,42,50,64	0
3	23D	A	301	27/27	0.97	0.16	25,33,41,61	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 23D C 301:

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

Electron density around 23D A 301:

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

6.5 Other polymers (i)

There are no such residues in this entry.