



Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 08:47 pm BST

PDB ID : 5BTC
Title : Crystal structure of a topoisomerase II complex
Authors : Blower, T.R.; Williamson, B.H.; Kerns, R.J.; Berger, J.M.
Deposited on : 2015-06-02
Resolution : 2.55 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 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

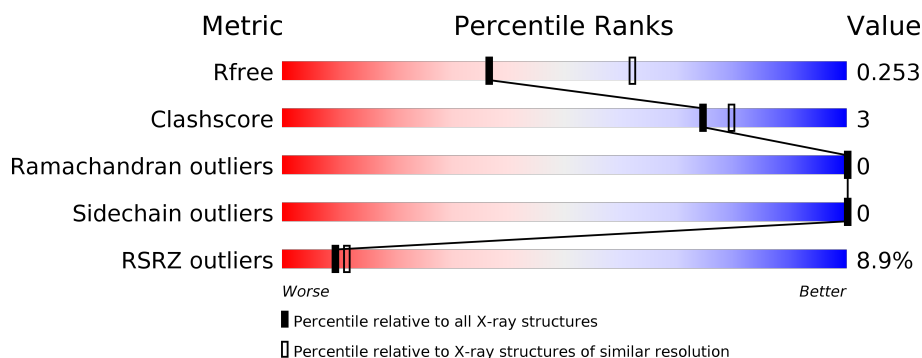
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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 $\leq 5\%$. 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	503	<div> <div>6%</div> <div> <div></div> <div>91%</div> <div>6%</div> <div>•</div> </div> </div>
1	C	503	<div> <div>12%</div> <div> <div></div> <div>91%</div> <div>6%</div> <div>•</div> </div> </div>
2	B	253	<div> <div>7%</div> <div> <div></div> <div>92%</div> <div>5%</div> <div>•</div> </div> </div>
2	D	253	<div> <div>4%</div> <div> <div></div> <div>92%</div> <div>6%</div> <div>•</div> </div> </div>
3	E	24	<div> <div>25%</div> <div> <div></div> <div>71%</div> <div>17%</div> <div>13%</div> </div> </div>
3	H	24	<div> <div>13%</div> <div> <div></div> <div>67%</div> <div>21%</div> <div>13%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	24	<div><div></div><div>17%</div><div>75%</div><div>13%</div><div>13%</div></div>
4	G	24	<div><div></div><div>21%</div><div>67%</div><div>21%</div><div>13%</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 25950 atoms, of which 12501 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 gyrase subunit A.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	487	Total	C	H	N	O	P	S	0	0	0
			7671	2383	3840	701	733	1	13			
1	C	487	Total	C	H	N	O	P	S	0	0	0
			7671	2383	3840	701	733	1	13			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	90	SER	ALA	engineered mutation	UNP P9WG47
A	501	ILE	-	expression tag	UNP P9WG47
A	502	GLY	-	expression tag	UNP P9WG47
A	503	SER	-	expression tag	UNP P9WG47
A	504	GLY	-	expression tag	UNP P9WG47
C	90	SER	ALA	engineered mutation	UNP P9WG47
C	501	ILE	-	expression tag	UNP P9WG47
C	502	GLY	-	expression tag	UNP P9WG47
C	503	SER	-	expression tag	UNP P9WG47
C	504	GLY	-	expression tag	UNP P9WG47

- Molecule 2 is a protein called DNA gyrase subunit B.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
2	B	245	Total	C	H	N	O	S		0	0	0
			3906	1217	1974	348	360	7				
2	D	247	Total	C	H	N	O	S		0	0	0
			3930	1224	1984	351	364	7				

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	423	SER	-	expression tag	UNP P9WG45
B	424	ASN	-	expression tag	UNP P9WG45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	425	ALA	-	expression tag	UNP P9WG45
D	423	SER	-	expression tag	UNP P9WG45
D	424	ASN	-	expression tag	UNP P9WG45
D	425	ALA	-	expression tag	UNP P9WG45

- Molecule 3 is a DNA chain called DNA substrate 24-mer GGTCATGAATGACTATGCAC GTAA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	21	Total	C	H	N	O	P	0	21	0
			621	196	211	74	120	20			
3	H	21	Total	C	H	N	O	P	0	21	0
			617	196	207	74	120	20			

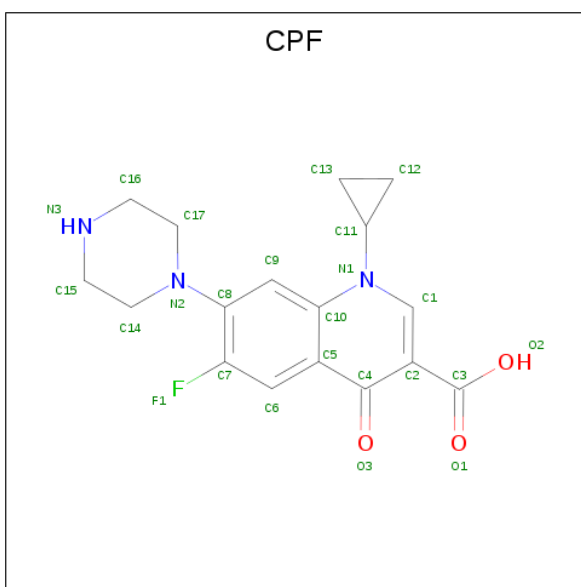
- Molecule 4 is a DNA chain called DNA substrate 24-mer TTACGTGCATAGTCATTTCAT GACC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	F	21	Total	C	H	N	O	P	0	21	0
			618	196	208	74	120	20			
4	G	21	Total	C	H	N	O	P	0	21	0
			613	196	203	74	120	20			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mg	0	0
			1	1		
5	A	1	Total	Mg	0	0
			1	1		
5	D	1	Total	Mg	0	0
			1	1		
5	C	1	Total	Mg	0	0
			1	1		

- Molecule 6 is 1-CYCLOPROPYL-6-FLUORO-4-OXO-7-PIPERAZIN-1-YL-1,4-DIHYDRO QUINOLINE-3-CARBOXYLIC ACID (three-letter code: CPF) (formula: C₁₇H₁₈FN₃O₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	G	1	Total	C	F	H	N	O	
			41	17	1	17	3	3	0
6	G	1	Total	C	F	H	N	O	
			41	17	1	17	3	3	0

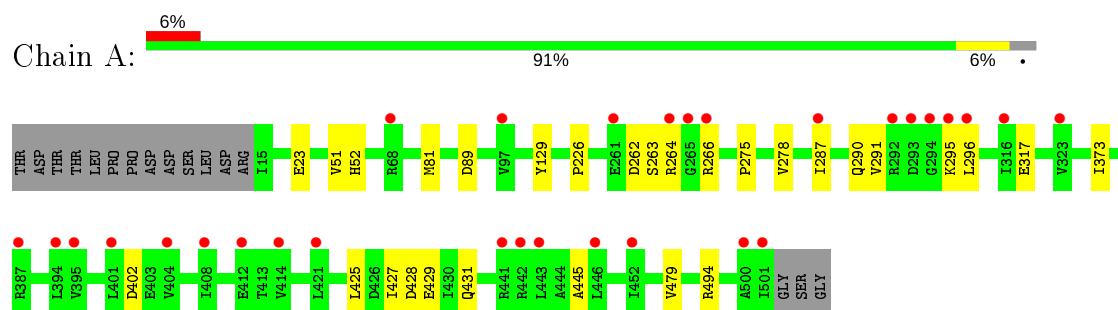
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	76	Total	O		
			76	76	0	0
7	B	31	Total	O		
			31	31	0	0
7	C	65	Total	O		
			65	65	0	0
7	D	26	Total	O		
			26	26	0	0
7	E	2	Total	O		
			2	2	0	0
7	F	8	Total	O		
			8	8	0	0
7	G	8	Total	O		
			8	8	0	0
7	H	1	Total	O		
			1	1	0	0

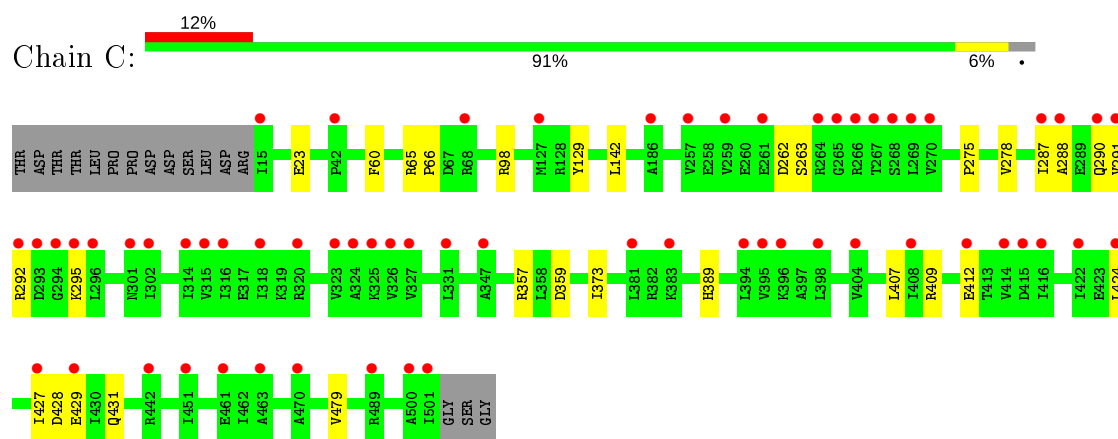
3 Residue-property plots [i](#)

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.

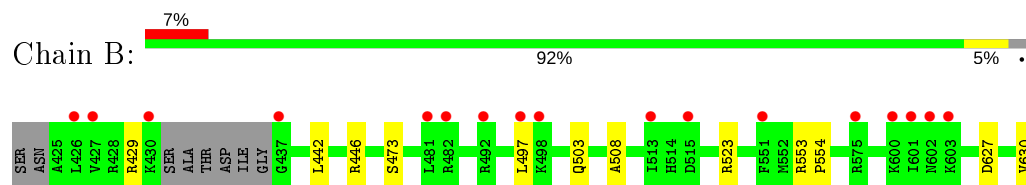
- Molecule 1: DNA gyrase subunit A



- Molecule 1: DNA gyrase subunit A

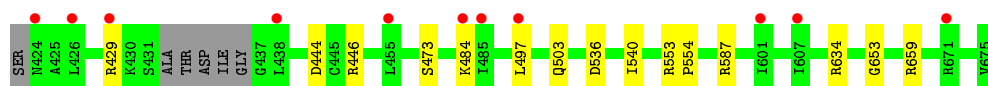


- Molecule 2: DNA gyrase subunit B

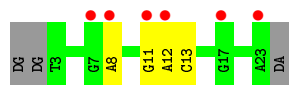


- Molecule 2: DNA gyrase subunit B

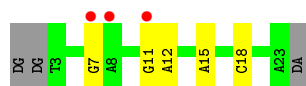




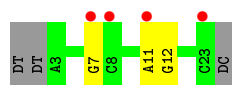
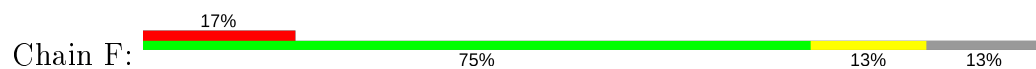
- Molecule 3: DNA substrate 24-mer GGTCATGAATGACTATGCACGTAA



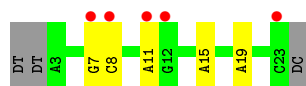
- Molecule 3: DNA substrate 24-mer GGTCATGAATGACTATGCACGTAA



- Molecule 4: DNA substrate 24-mer TTACGTGCATAGTCATTCATGACC



- Molecule 4: DNA substrate 24-mer TTACGTGCATAGTCATTCATGACC



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	108.34Å 82.98Å 129.70Å 90.00° 109.06° 90.00°	Depositor
Resolution (Å)	47.70 – 2.55 47.70 – 2.55	Depositor EDS
% Data completeness (in resolution range)	95.5 (47.70-2.55) 96.1 (47.70-2.55)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.18 (at 2.54Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.224 , 0.250 0.228 , 0.253	Depositor DCC
R_{free} test set	3453 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	51.0	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 52.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	25950	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.37% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, PTR, CPF

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	0.22	0/3872	0.40	0/5243
1	C	0.21	0/3872	0.39	0/5243
2	B	0.21	0/1960	0.39	0/2634
2	D	0.22	0/1974	0.39	0/2653
3	E	0.54	0/458	0.94	0/704
3	H	0.56	0/458	0.96	0/704
4	F	0.53	0/458	0.95	0/704
4	G	0.52	0/458	0.93	0/704
All	All	0.28	0/13510	0.52	0/18589

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 [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	3831	3840	3840	26	0
1	C	3831	3840	3840	22	0
2	B	1932	1974	1973	7	0
2	D	1946	1984	1984	10	0
3	E	410	211	211	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	410	207	207	5	0
4	F	410	208	208	4	0
4	G	410	203	203	5	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	G	48	34	32	2	0
7	A	76	0	0	0	0
7	B	31	0	0	0	0
7	C	65	0	0	0	0
7	D	26	0	0	1	0
7	E	2	0	0	0	0
7	F	8	0	0	0	0
7	G	8	0	0	0	0
7	H	1	0	0	0	0
All	All	13449	12501	12498	67	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 3.

All (67) 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:C:129:PTR:O3P	4:G:11[B]:DA:C5'	1.68	1.40
1:C:129:PTR:O3P	3:E:11[A]:DG:C5'	1.88	1.20
1:A:295:LYS:O	1:A:296:LEU:HD12	1.44	1.14
1:C:290:GLN:HG2	1:C:295:LYS:HD2	1.23	1.13
1:A:129:PTR:O1P	3:H:11[B]:DG:C5'	1.98	1.10
1:A:129:PTR:O1P	4:F:11[A]:DA:C5'	2.06	1.03
1:A:295:LYS:C	1:A:296:LEU:HD12	1.93	0.88
2:D:653:GLY:O	2:D:659:ARG:NH2	2.11	0.83
1:C:290:GLN:O	1:C:295:LYS:HB2	1.78	0.83
1:A:290:GLN:HB3	1:A:296:LEU:HD13	1.60	0.81
1:C:98:ARG:NH1	4:F:7[A]:DG:OP1	2.19	0.74
1:C:287:ILE:O	1:C:291:VAL:HG23	1.88	0.72
1:C:288:ALA:O	1:C:292:ARG:HG3	1.92	0.69
1:C:357:ARG:NH1	1:C:359:ASP:OD1	2.25	0.68
2:D:484:LYS:NZ	2:D:536:ASP:O	2.29	0.66
1:A:262:ASP:HB3	1:A:266:ARG:HB2	1.79	0.64
1:A:287:ILE:O	1:A:291:VAL:HG23	1.99	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:659:ARG:NH1	4:G:19[B]:DA:OP1	2.40	0.54
2:D:446:ARG:NH2	2:D:473:SER:OG	2.40	0.54
1:C:98:ARG:NH1	3:H:7[B]:DG:OP1	2.41	0.54
2:B:429:ARG:NH1	2:B:442:LEU:O	2.41	0.54
3:E:12[A]:DA:N6	4:F:12[A]:DG:O6	2.43	0.52
1:A:427:ILE:CG2	1:A:431:GLN:HB2	2.39	0.52
1:C:407:LEU:HD11	1:C:424:LEU:HD22	1.93	0.51
2:D:497:LEU:O	2:D:503:GLN:NE2	2.42	0.51
2:B:446:ARG:NH1	2:B:473:SER:OG	2.44	0.49
1:C:262:ASP:OD1	1:C:263:SER:N	2.40	0.49
1:A:51:VAL:HG11	3:E:8[A]:DA:H5'	1.94	0.49
2:D:587:ARG:NH2	7:D:804:HOH:O	2.45	0.49
1:A:52:HIS:NE2	4:G:8[B]:DC:OP1	2.43	0.49
1:C:428:ASP:OD1	1:C:429:GLU:N	2.44	0.48
1:A:262:ASP:CB	1:A:266:ARG:HB2	2.44	0.47
1:A:264:ARG:O	1:A:266:ARG:HG3	2.14	0.47
1:A:23:GLU:OE1	2:B:634:ARG:NH1	2.48	0.47
1:A:402:ASP:OD1	1:C:409:ARG:NH1	2.48	0.46
1:C:373:ILE:HD12	1:C:479:VAL:HG22	1.98	0.46
2:B:508:ALA:O	2:B:523:ARG:NE	2.44	0.45
4:G:7[B]:DG:N2	3:H:18[B]:DC:O2	2.42	0.45
1:C:389:HIS:O	1:C:431:GLN:NE2	2.43	0.44
1:A:291:VAL:HG22	1:A:296:LEU:HB2	1.99	0.44
1:A:275:PRO:HG2	1:A:278:VAL:HG21	1.99	0.44
1:C:65:ARG:HB3	1:C:66:PRO:HD2	1.99	0.44
1:A:266:ARG:HD3	1:A:317:GLU:OE2	2.18	0.44
1:C:23:GLU:OE1	2:D:634:ARG:NH1	2.51	0.44
2:B:627:ASP:HB3	2:B:630:VAL:HG22	2.00	0.43
1:C:290:GLN:HG2	1:C:295:LYS:CD	2.17	0.43
1:A:262:ASP:OD1	1:A:263:SER:N	2.48	0.42
1:C:60:PHE:HB2	1:C:142:LEU:HD13	2.01	0.42
3:E:13[A]:DC:N4	4:F:11[A]:DA:N6	2.68	0.42
1:A:373:ILE:HD12	1:A:479:VAL:HG22	2.01	0.42
6:G:102:CPF:H171	3:H:15[B]:DA:N3	2.35	0.42
1:A:290:GLN:HB3	1:A:296:LEU:CD1	2.41	0.42
2:D:553:ARG:N	2:D:554:PRO:CD	2.83	0.42
2:B:553:ARG:N	2:B:554:PRO:CD	2.83	0.41
1:C:275:PRO:HG2	1:C:278:VAL:HG21	2.02	0.41
1:A:81:MET:HE1	1:A:89:ASP:HB3	2.02	0.41
1:A:428:ASP:OD1	1:A:429:GLU:N	2.43	0.41
2:D:429:ARG:HG2	2:D:444:ASP:HB2	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:G:15[B]:DA:N3	6:G:101:CPF:H171	2.36	0.41
1:A:445:ALA:HB3	1:C:412:GLU:HA	2.03	0.41
2:B:497:LEU:O	2:B:503:GLN:NE2	2.53	0.41
1:A:226:PRO:O	1:A:494:ARG:NH2	2.53	0.41
2:D:484:LYS:HZ2	2:D:540:ILE:HG13	1.85	0.41
1:A:425:LEU:HB2	1:A:427:ILE:HD11	2.04	0.40
3:H:11[B]:DG:H2"	3:H:12[B]:DA:C8	2.57	0.40
1:A:427:ILE:HG22	1:A:428:ASP:N	2.37	0.40
1:C:427:ILE:CG2	1:C:431:GLN:HB2	2.51	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	484/503 (96%)	463 (96%)	21 (4%)	0	100	100
1	C	484/503 (96%)	467 (96%)	17 (4%)	0	100	100
2	B	241/253 (95%)	233 (97%)	8 (3%)	0	100	100
2	D	243/253 (96%)	233 (96%)	10 (4%)	0	100	100
All	All	1452/1512 (96%)	1396 (96%)	56 (4%)	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	407/421 (97%)	407 (100%)	0	100	100
1	C	407/421 (97%)	407 (100%)	0	100	100
2	B	205/211 (97%)	205 (100%)	0	100	100
2	D	207/211 (98%)	207 (100%)	0	100	100
All	All	1226/1264 (97%)	1226 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some 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 ⓘ

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	PTR	A	129	1	15,16,17	1.34	1 (6%)	19,22,24	1.33	2 (10%)
1	PTR	C	129	1	15,16,17	1.15	2 (13%)	19,22,24	1.04	1 (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
1	PTR	A	129	1	-	2/10/11/13	0/1/1/1
1	PTR	C	129	1	-	2/10/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	129	PTR	OH-CZ	-2.76	1.34	1.40
1	C	129	PTR	OH-CZ	-2.42	1.35	1.40
1	C	129	PTR	O-C	2.03	1.28	1.19

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	129	PTR	CB-CA-C	-3.96	104.04	111.47
1	C	129	PTR	CB-CA-C	-2.85	106.13	111.47
1	A	129	PTR	O3P-P-O2P	2.28	116.34	107.64

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	129	PTR	CA-CB-CG-CD2
1	C	129	PTR	CA-CB-CG-CD1
1	A	129	PTR	CA-CB-CG-CD2
1	A	129	PTR	CA-CB-CG-CD1

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	129	PTR	2	0
1	C	129	PTR	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates 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
6	CPF	G	102	5	19,27,27	3.17	7 (36%)	20,40,40	1.60	3 (15%)
6	CPF	G	101	5	19,27,27	3.15	7 (36%)	20,40,40	1.64	4 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	CPF	G	102	5	-	4/8/22/22	0/4/4/4
6	CPF	G	101	5	-	4/8/22/22	0/4/4/4

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	101	CPF	C10-N1	10.09	1.51	1.40
6	G	102	CPF	C10-N1	10.00	1.51	1.40
6	G	102	CPF	C1-C2	5.32	1.48	1.39
6	G	101	CPF	C1-C2	5.17	1.47	1.39
6	G	102	CPF	C2-C3	5.04	1.52	1.47
6	G	101	CPF	C2-C3	4.89	1.52	1.47
6	G	102	CPF	C6-C7	3.53	1.40	1.35
6	G	101	CPF	C6-C7	3.41	1.40	1.35
6	G	101	CPF	C8-N2	2.52	1.46	1.41
6	G	102	CPF	C8-N2	2.49	1.46	1.41
6	G	102	CPF	C17-N2	-2.42	1.42	1.46
6	G	101	CPF	C17-N2	-2.35	1.43	1.46
6	G	102	CPF	C14-N2	-2.13	1.43	1.46
6	G	101	CPF	C14-N2	-2.12	1.43	1.46

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	101	CPF	F1-C7-C8	4.99	122.96	118.42
6	G	102	CPF	F1-C7-C8	4.69	122.69	118.42
6	G	102	CPF	C7-C8-N2	3.46	124.58	120.47
6	G	101	CPF	C7-C8-N2	3.43	124.55	120.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	101	CPF	F1-C7-C6	-2.84	116.72	120.47
6	G	102	CPF	F1-C7-C6	-2.69	116.92	120.47
6	G	101	CPF	C4-C5-C10	2.02	120.35	118.59

There are no chirality outliers.

All (8) torsion outliers are listed below:

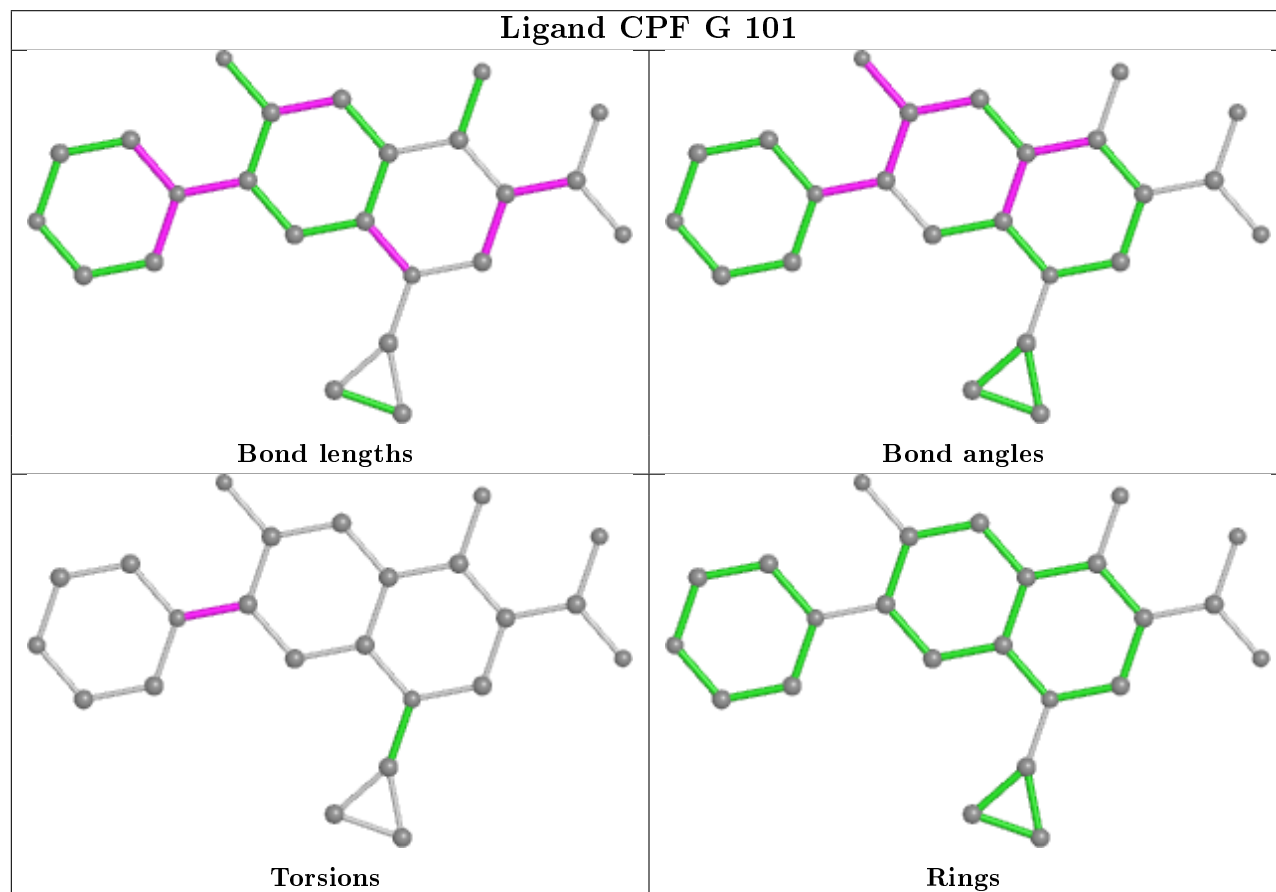
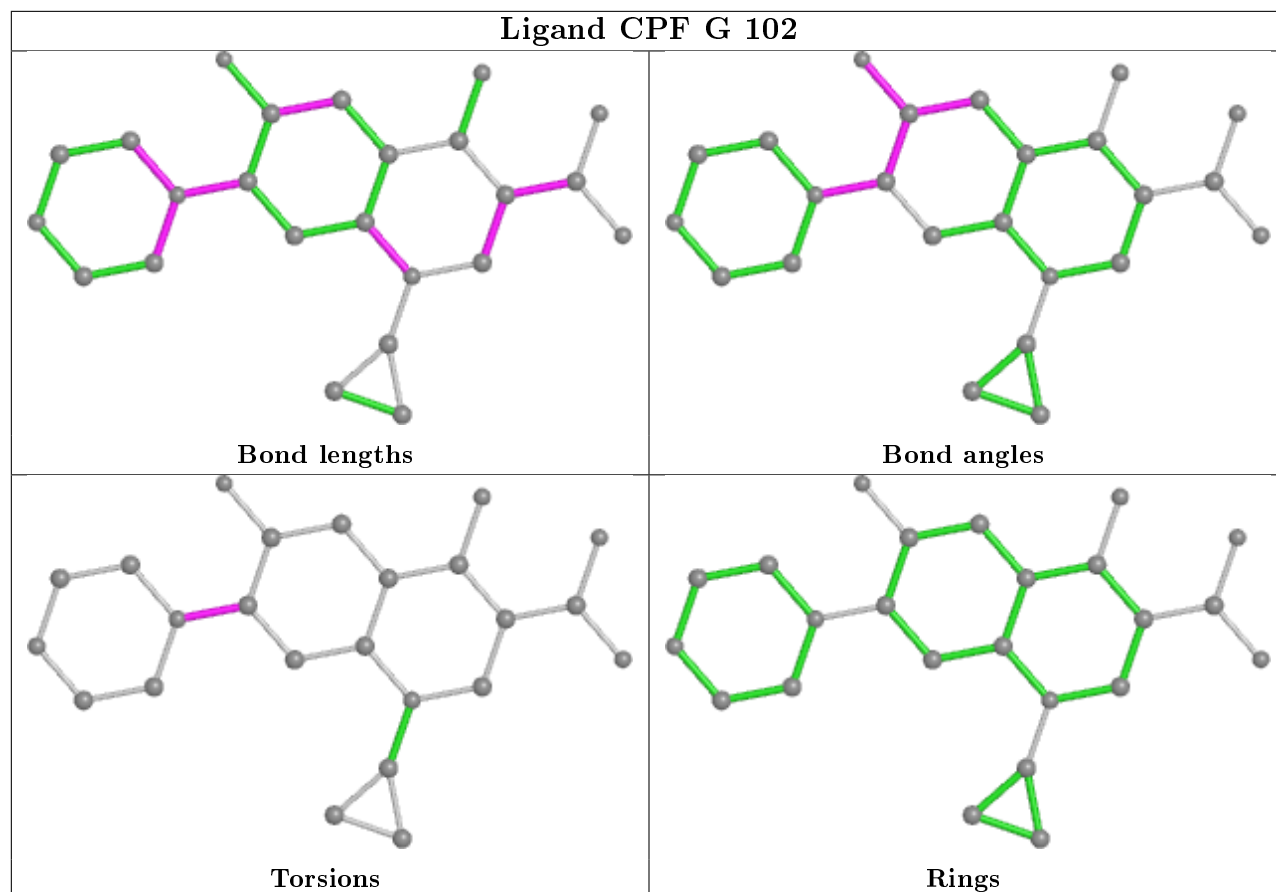
Mol	Chain	Res	Type	Atoms
6	G	101	CPF	C7-C8-N2-C17
6	G	102	CPF	C7-C8-N2-C17
6	G	101	CPF	C9-C8-N2-C14
6	G	102	CPF	C9-C8-N2-C14
6	G	102	CPF	C7-C8-N2-C14
6	G	101	CPF	C7-C8-N2-C14
6	G	101	CPF	C9-C8-N2-C17
6	G	102	CPF	C9-C8-N2-C17

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	G	102	CPF	1	0
6	G	101	CPF	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 than 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.



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(Å ²)	Q<0.9
1	A	486/503 (96%)	0.75	30 (6%) 20 24	33, 59, 120, 179	0
1	C	486/503 (96%)	0.99	62 (12%) 3 4	34, 62, 119, 166	0
2	B	245/253 (96%)	0.79	17 (6%) 16 20	43, 69, 106, 143	0
2	D	247/253 (97%)	0.64	11 (4%) 33 40	44, 65, 112, 143	0
3	E	21/24 (87%)	1.57	6 (28%) 0 0	47, 93, 221, 296	21 (100%)
3	H	21/24 (87%)	1.05	3 (14%) 2 3	17, 54, 148, 216	21 (100%)
4	F	21/24 (87%)	1.16	4 (19%) 1 1	68, 98, 215, 260	21 (100%)
4	G	21/24 (87%)	1.45	5 (23%) 0 0	31, 54, 154, 197	21 (100%)
All	All	1548/1608 (96%)	0.85	138 (8%) 9 11	17, 64, 124, 296	84 (5%)

All (138) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	501	ILE	12.4
1	C	501	ILE	11.2
1	A	292	ARG	9.2
1	C	292	ARG	8.3
3	E	23[A]	DA	6.3
4	G	23[B]	DC	5.8
1	C	265	GLY	5.8
2	B	601	ILE	5.8
1	A	408	ILE	5.4
2	B	437	GLY	5.4
1	A	265	GLY	5.1
4	G	12[B]	DG	5.1
1	C	259	VAL	4.8
1	A	293	ASP	4.8
1	C	500	ALA	4.4
1	C	295	LYS	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	E	12[A]	DA	4.4
2	B	430	LYS	4.4
1	C	422	ILE	4.3
1	A	264	ARG	4.3
1	C	264	ARG	4.3
2	B	426	LEU	4.1
1	C	414	VAL	3.9
1	C	296	LEU	3.8
3	E	11[A]	DG	3.7
1	C	408	ILE	3.7
2	B	603	LYS	3.7
1	A	401	LEU	3.7
1	C	269	LEU	3.7
2	D	426	LEU	3.6
1	C	323	VAL	3.6
1	C	416	ILE	3.5
1	A	412	GLU	3.5
1	A	395	VAL	3.5
1	C	291	VAL	3.4
1	C	301	ASN	3.4
1	C	394	LEU	3.3
2	B	427	VAL	3.3
1	A	387	ARG	3.3
1	C	404	VAL	3.3
1	A	394	LEU	3.2
1	C	412	GLU	3.2
1	C	424	LEU	3.2
1	C	327	VAL	3.2
1	A	421	LEU	3.2
2	D	424	ASN	3.1
1	C	261	GLU	3.1
1	C	396	LYS	3.1
1	C	470	ALA	3.1
2	B	551	PHE	3.1
1	A	266	ARG	3.1
1	A	296	LEU	3.1
1	C	395	VAL	3.0
1	A	294	GLY	3.0
1	A	500	ALA	3.0
2	B	497	LEU	3.0
1	A	446	LEU	3.0
2	B	498	LYS	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	484	LYS	2.9
2	B	515	ASP	2.9
4	G	11[B]	DA	2.9
1	C	266	ARG	2.9
4	F	7[A]	DG	2.9
1	C	294	GLY	2.9
1	C	442	ARG	2.9
2	B	600	LYS	2.9
1	C	324	ALA	2.9
1	C	290	GLN	2.9
1	A	452	ILE	2.9
1	A	261	GLU	2.8
3	H	11[B]	DG	2.8
3	H	7[B]	DG	2.8
4	F	23[A]	DC	2.8
2	B	481	LEU	2.7
1	A	68	ARG	2.7
1	C	429	GLU	2.7
1	C	318	ILE	2.7
3	E	7[A]	DG	2.7
1	C	398	LEU	2.7
2	B	575	ARG	2.7
1	C	293	ASP	2.7
1	C	268	SER	2.6
1	A	323	VAL	2.6
1	C	315	VAL	2.6
1	C	316	ILE	2.6
4	G	7[B]	DG	2.6
1	C	489	ARG	2.5
1	A	414	VAL	2.5
1	C	415	ASP	2.5
1	C	302	ILE	2.5
2	D	601	ILE	2.5
1	C	461	GLU	2.5
1	C	326	VAL	2.4
2	D	438	LEU	2.4
4	F	8[A]	DC	2.4
1	C	15	ILE	2.4
2	D	497	LEU	2.4
1	C	427	ILE	2.4
1	C	463	ALA	2.4
4	F	11[A]	DA	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	E	17[A]	DG	2.3
1	C	320	ARG	2.3
2	D	671	ARG	2.3
1	C	68	ARG	2.3
1	C	267	THR	2.3
1	A	295	LYS	2.3
1	A	441	ARG	2.3
1	C	186	ALA	2.2
1	C	325	LYS	2.2
1	C	287	ILE	2.2
1	C	288	ALA	2.2
3	H	8[B]	DA	2.2
1	A	316	ILE	2.2
1	A	443	LEU	2.2
1	C	314	ILE	2.2
1	C	331	LEU	2.2
1	C	270	VAL	2.2
2	B	482	ARG	2.2
1	C	381	LEU	2.2
4	G	8[B]	DC	2.2
2	B	492	ARG	2.1
3	E	8[A]	DA	2.1
1	C	127	MET	2.1
2	B	602	ASN	2.1
1	C	257	VAL	2.1
1	C	42	PRO	2.1
1	C	451	ILE	2.1
2	D	455	LEU	2.1
1	A	442	ARG	2.1
1	C	383	LYS	2.1
1	A	287	ILE	2.1
2	D	607	ILE	2.1
2	D	429	ARG	2.0
1	A	97	VAL	2.0
2	D	485	ILE	2.0
1	C	347	ALA	2.0
1	A	404	VAL	2.0
2	B	513	ILE	2.0

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(\AA^2)	Q<0.9
1	PTR	A	129	16/17	0.92	0.26	46,53,109,115	0
1	PTR	C	129	16/17	0.93	0.25	50,61,103,104	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

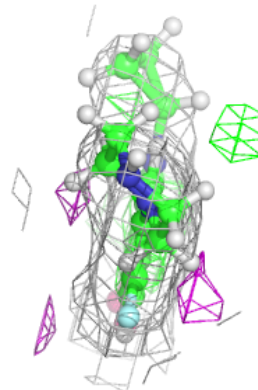
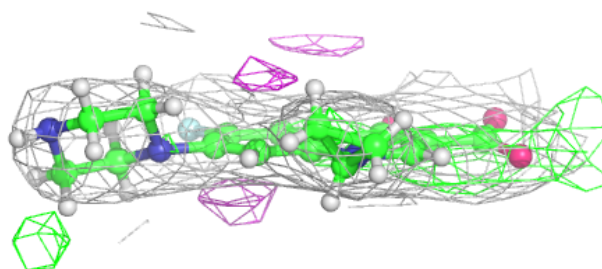
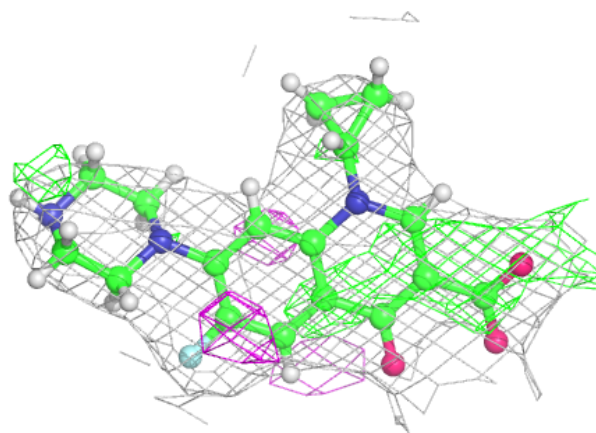
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
6	CPF	G	102	24/24	0.90	0.34	73,84,101,105	0
5	MG	D	701	1/1	0.90	0.23	47,47,47,47	0
6	CPF	G	101	24/24	0.92	0.31	51,64,92,101	0
5	MG	C	601	1/1	0.96	0.20	42,42,42,42	0
5	MG	A	601	1/1	0.97	0.25	58,58,58,58	0
5	MG	B	701	1/1	0.98	0.17	39,39,39,39	0

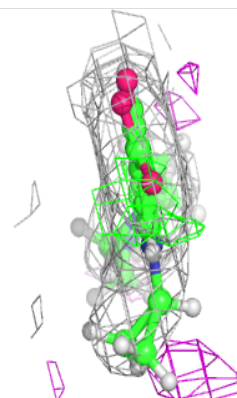
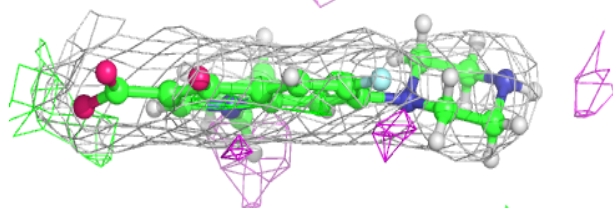
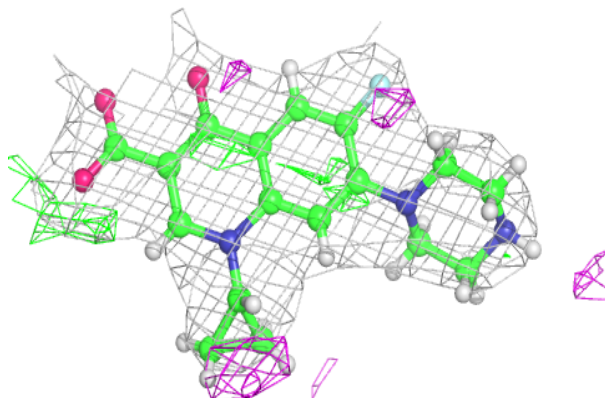
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CPF G 102:

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

**Electron density around CPF G 101:**

$2mF_o-DF_c$ (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.