Full wwPDB X-ray Structure Validation Report (i)

Jan 5, 2022 – 04:22 PM EST

PDB ID : 7RIL
Title : Crystal structure of hairpin polyamide Py-Im 1 bound to 5' CCTGACCAGG
Authors : Oh, J.; Dervan, P.B.; Wang, D.
Deposited on : 2021-07-20
Resolution : 1.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:
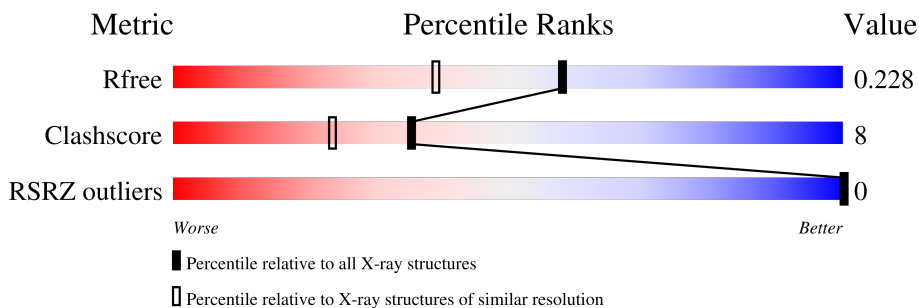
1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.80 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		0.228
Clashscore		8
RSRZ outliers		0

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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	10	70% 30%
2	B	10	10% 70% 20%

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ACT	B	102	-	-	-	X

2   Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 524 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a DNA chain called non-template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	10	201	96	39	57	9	0	0	0

- Molecule 2 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	10	203	97	38	59	9	0	0	0

- Molecule 3 is 3-({3-[(3-{[4-({4-[(4-{[4-({(2R)-2-amino-4-[(1-methyl-4-{[1-methyl-4-({1-methyl-4-[(1-methyl-1H-imidazole-2-carbonyl)amino]-1H-imidazole-2-carbonyl}amino)-1H-pyrrole-2-carbonyl]amino}-1H-pyrrole-2-carbonyl)amino]butanoyl}amino)-1-methyl-1H-imidazole-2-carbonyl]amino}-1-methyl-1H-pyrrole-2-carbonyl)amino]-1-methyl-1H-pyrrole-2-carbonyl}amino)-1-methyl-1H-pyrrole-2-carbonyl]amino}propyl)(methyl)amino]propyl}carbamoyl)benzoic acid (three-letter code: 5N0) (formula: $C_{64}H_{75}N_{23}O_{12}$) (labeled as "Ligand of Interest" by depositor).
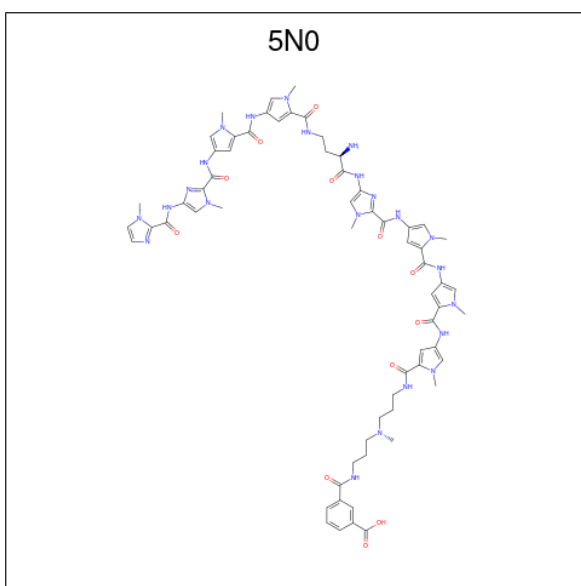
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total 99	C 64	N 23	O 12	0	0

• Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 4	C 2	O 2	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total 6	O 6	0	0
5	B	11	Total 11	O 11	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: non-template DNA

Chain A:

70% 30%

• Molecule 2: template DNA

Chain B:

10% 70% 20%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	33.91Å 33.91Å 47.46Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.37 – 1.80 29.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.37-1.80) 99.8 (29.37-1.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 1.80Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.203 , 0.228 0.203 , 0.228	Depositor DCC
R_{free} test set	307 reflections (5.44%)	wwPDB-VP
Wilson B-factor (Å^2)	47.5	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 47.3	EDS
L-test for twinning[2]	$< |L| > = 0.52, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l 0.093 for h,-h-k,-l 0.072 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	524	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.29% 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: 5N0, ACT

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	1.60	2/225 (0.9%)	1.31	1/345 (0.3%)
2	B	1.67	1/227 (0.4%)	1.42	4/349 (1.1%)
All	All	1.64	3/452 (0.7%)	1.36	5/694 (0.7%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	5	DG	C3'-O3'	-5.73	1.36	1.44
1	A	5	DA	N9-C4	5.31	1.41	1.37
1	A	9	DG	C3'-O3'	-5.16	1.37	1.44

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	3	DT	O4'-C1'-N1	7.93	113.55	108.00
2	B	7	DC	O4'-C1'-N1	-7.81	102.53	108.00
2	B	9	DG	O4'-C1'-N9	-5.71	104.00	108.00
2	B	3	DT	OP1-P-OP2	5.27	127.51	119.60
2	B	2	DC	O4'-C1'-N1	-5.15	104.40	108.00

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	201	0	113	0	0
2	B	203	0	114	6	0
3	B	99	0	0	2	0
4	B	4	0	3	1	0
5	A	6	0	0	0	0
5	B	11	0	0	2	0
All	All	524	0	230	6	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 8.

All (6) 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:1:DC:O5'	5:B:201:HOH:O	2.11	0.68
2:B:4:DG:H2"	2:B:5:DG:C8	2.39	0.56
2:B:10:DG:N3	5:B:202:HOH:O	2.34	0.51
2:B:6:DT:H4'	3:B:101:5N0:C8	2.42	0.49
2:B:7:DC:H4'	3:B:101:5N0:C13	2.47	0.45
2:B:10:DG:N7	4:B:102:ACT:H2	2.35	0.42

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

There are no protein molecules in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein molecules in this entry.

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	$\#|Z| > 2$
4	ACT	B	102	-	1,3,3	5.96	1 (100%)	0,3,3	-	-
3	5N0	B	101	-	91,107,107	2.58	29 (31%)	91,153,153	2.04	25 (27%)

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	5N0	B	101	-	-	12/47/92/92	0/9/9/9

All (30) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	101	5N0	C59-C63	9.78	1.56	1.47
3	B	101	5N0	C26-N14	7.06	1.51	1.35
3	B	101	5N0	C56-N25	6.86	1.48	1.33
3	B	101	5N0	C30-C31	6.14	1.58	1.53
4	B	102	ACT	CH3-C	5.96	1.56	1.48
3	B	101	5N0	C49-N23	5.73	1.46	1.33
3	B	101	5N0	C22-N10	5.57	1.46	1.33
3	B	101	5N0	C5-N3	4.91	1.48	1.35
3	B	101	5N0	C10-N6	4.67	1.48	1.35
3	B	101	5N0	C31-N17	4.48	1.47	1.35
3	B	101	5N0	C43-N21	4.41	1.47	1.35
3	B	101	5N0	C29-N14	4.39	1.50	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	101	5N0	C37-N19	4.18	1.46	1.35
3	B	101	5N0	C16-N8	3.69	1.45	1.35
3	B	101	5N0	C36-N17	3.39	1.48	1.41
3	B	101	5N0	C9-N3	3.22	1.47	1.40
3	B	101	5N0	C57-C56	2.98	1.56	1.50
3	B	101	5N0	C21-N8	2.75	1.47	1.41
3	B	101	5N0	C13-C14	2.71	1.41	1.38
3	B	101	5N0	C6-C10	2.68	1.55	1.53
3	B	101	5N0	C34-C36	2.57	1.41	1.38
3	B	101	5N0	O3-C16	-2.52	1.18	1.23
3	B	101	5N0	O7-C37	-2.40	1.18	1.23
3	B	101	5N0	O9-C49	-2.26	1.18	1.23
3	B	101	5N0	C46-C48	2.25	1.40	1.38
3	B	101	5N0	C38-C42	2.14	1.41	1.39
3	B	101	5N0	C11-N7	-2.14	1.32	1.36
3	B	101	5N0	C42-N19	2.12	1.45	1.41
3	B	101	5N0	C19-C21	2.11	1.40	1.38
3	B	101	5N0	O5-C26	-2.11	1.19	1.23

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	101	5N0	C30-C31-N17	7.49	121.82	113.69
3	B	101	5N0	C24-C25-C26	-6.63	96.50	110.85
3	B	101	5N0	C29-N14-C26	-6.23	118.71	127.55
3	B	101	5N0	C46-C48-C47	4.27	108.14	106.05
3	B	101	5N0	C57-C56-N25	3.82	125.29	117.09
3	B	101	5N0	C6-C10-N6	3.80	117.82	113.69
3	B	101	5N0	C21-N8-C16	-3.64	117.13	126.58
3	B	101	5N0	C15-C14-C13	3.38	107.70	106.05
3	B	101	5N0	C4-C5-N3	3.23	117.20	113.69
3	B	101	5N0	O4-C22-N10	-3.17	116.29	122.61
3	B	101	5N0	C15-C11-C16	-3.08	118.94	128.55
3	B	101	5N0	C23-N10-C22	-2.96	115.33	122.08
3	B	101	5N0	C14-N6-C10	-2.93	118.96	126.58
3	B	101	5N0	C47-C44-C49	-2.77	119.91	128.55
3	B	101	5N0	O5-C26-C25	-2.57	114.73	120.18
3	B	101	5N0	O10-C56-N25	-2.49	117.65	122.61
3	B	101	5N0	C41-N20-C39	2.34	111.13	108.65
3	B	101	5N0	C24-C25-N11	-2.29	104.16	110.17
3	B	101	5N0	O10-C56-C57	-2.17	117.06	120.94
3	B	101	5N0	O6-C31-N17	-2.17	118.76	123.71

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	101	5N0	C24-C23-N10	2.16	118.11	111.87
3	B	101	5N0	C19-C21-C20	2.09	107.07	106.05
3	B	101	5N0	C9-N3-C5	-2.08	122.31	128.07
3	B	101	5N0	C20-C17-C22	-2.06	122.13	128.55
3	B	101	5N0	C55-N25-C56	2.02	126.68	122.08

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	101	5N0	C23-C24-C25-C26
3	B	101	5N0	C51-C52-N24-C53
3	B	101	5N0	O10-C56-N25-C55
3	B	101	5N0	C57-C56-N25-C55
3	B	101	5N0	O10-C56-C57-C58
3	B	101	5N0	N25-C56-C57-C58
3	B	101	5N0	O10-C56-C57-C62
3	B	101	5N0	N25-C56-C57-C62
3	B	101	5N0	N24-C53-C54-C55
3	B	101	5N0	N23-C50-C51-C52
3	B	101	5N0	C51-C52-N24-C64
3	B	101	5N0	N10-C23-C24-C25

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	102	ACT	1	0
3	B	101	5N0	2	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 5N0 B 101

Bond lengths	Bond angles
Torsions	Rings
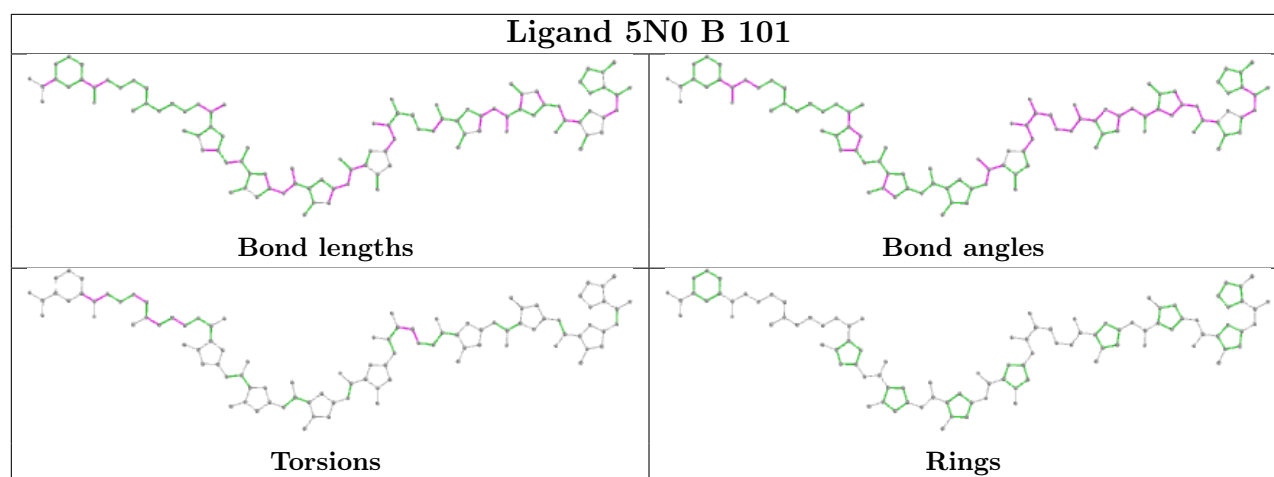
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	10/10 (100%)	-0.08	0	100	100	59, 63, 71, 73	0
2	B	10/10 (100%)	0.00	0	100	100	49, 57, 60, 66	0
All	All	20/20 (100%)	-0.04	0	100	100	49, 60, 71, 73	0

There are no RSRZ outliers to report.

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

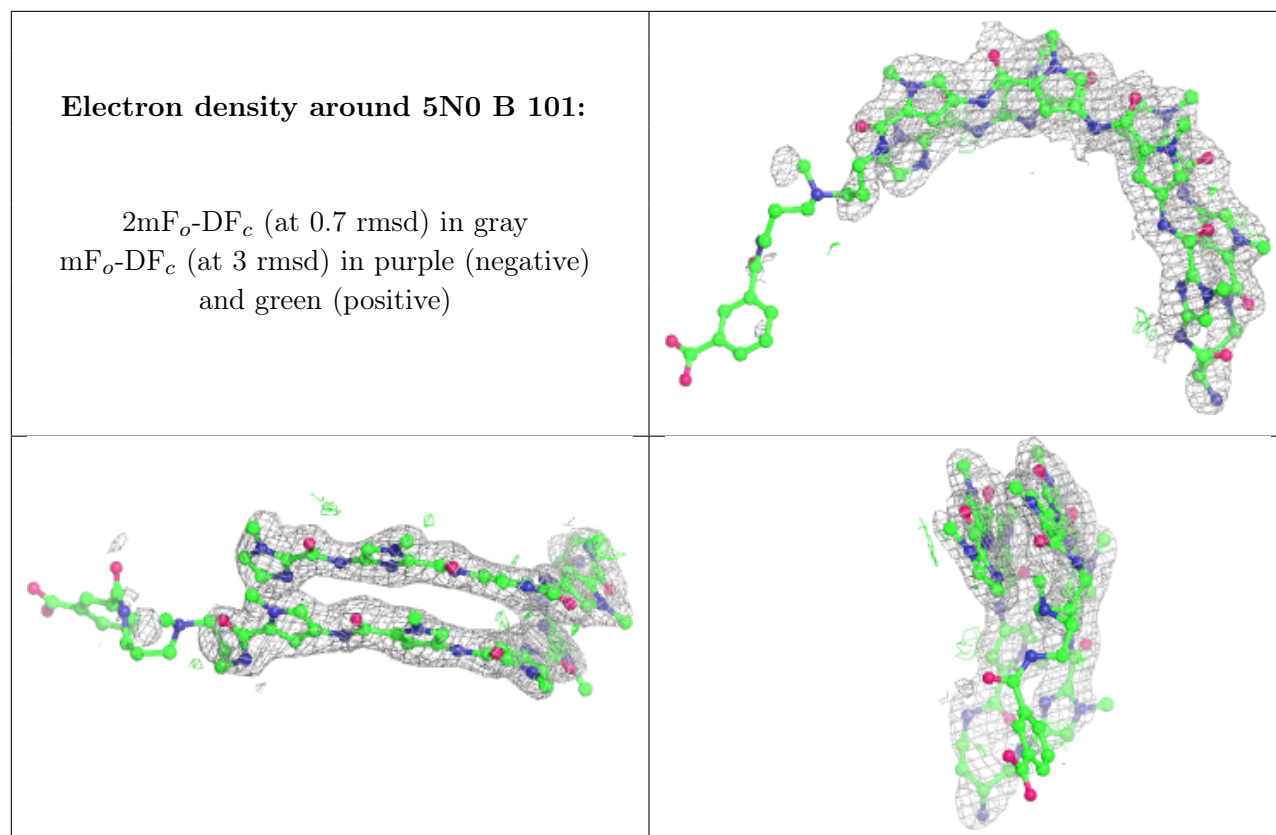There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
4	ACT	B	102	4/4	-0.06	0.45	69,74,79,80	0
3	5N0	B	101	99/99	0.87	0.19	41,62,133,141	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 5N0 B 101:

$2mF_o\text{-}DF_c$ (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.