



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 02:35 am BST

PDB ID : 5ENP  
Title : MBX2931 bound structure of bacterial efflux pump.  
Authors : Sjuts, H.; Ornik, A.R.; Pos, K.M.  
Deposited on : 2015-11-09  
Resolution : 1.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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

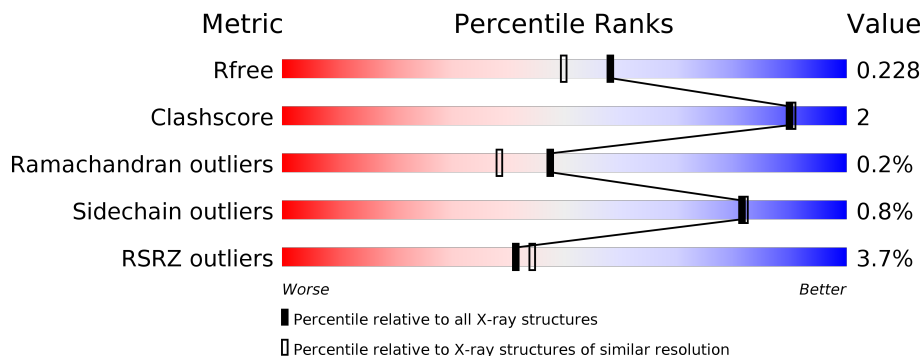
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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  $\geq 3$ , 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	609	 4% 90% 5% 5%
1	B	609	 2% 89% 5% 5%
1	C	609	 3% 88% 6% 5%
2	D	169	 4% 92% 5% 5%
2	E	169	 7% 90% 8% 5%
2	F	169	 7% 87% 9% 5%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 18206 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 Multidrug efflux pump subunit AcrB, Multidrug efflux pump subunit AcrB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	580	4431	2783	749	877	22	0	3	0
1	B	578	4411	2770	746	873	22	0	2	0
1	C	578	4404	2766	745	871	22	0	1	0

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	552	GLY	-	linker	UNP P31224
A	553	GLY	-	linker	UNP P31224
A	554	SER	-	linker	UNP P31224
A	555	GLY	-	linker	UNP P31224
A	556	GLY	-	linker	UNP P31224
A	557	SER	-	linker	UNP P31224
A	558	GLY	-	linker	UNP P31224
A	559	GLY	-	linker	UNP P31224
A	560	SER	-	linker	UNP P31224
B	552	GLY	-	linker	UNP P31224
B	553	GLY	-	linker	UNP P31224
B	554	SER	-	linker	UNP P31224
B	555	GLY	-	linker	UNP P31224
B	556	GLY	-	linker	UNP P31224
B	557	SER	-	linker	UNP P31224
B	558	GLY	-	linker	UNP P31224
B	559	GLY	-	linker	UNP P31224
B	560	SER	-	linker	UNP P31224
C	552	GLY	-	linker	UNP P31224
C	553	GLY	-	linker	UNP P31224
C	554	SER	-	linker	UNP P31224
C	555	GLY	-	linker	UNP P31224

*Continued on next page...*

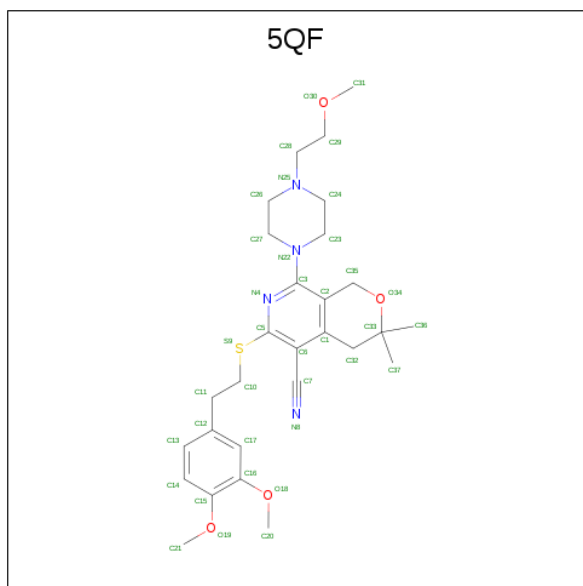
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	556	GLY	-	linker	UNP P31224
C	557	SER	-	linker	UNP P31224
C	558	GLY	-	linker	UNP P31224
C	559	GLY	-	linker	UNP P31224
C	560	SER	-	linker	UNP P31224

- Molecule 2 is a protein called DARPin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	161	Total	C	N	O	S	0	0	0
			1227	771	221	234	1			
2	E	156	Total	C	N	O	S	0	0	0
			1177	741	206	229	1			
2	F	162	Total	C	N	O	S	0	0	0
			1237	777	224	235	1			

- Molecule 3 is 6-[2-(3,4-dimethoxyphenyl)ethylsulfanyl]-8-[4-(2-methoxyethyl)piperazin-1-yl]-3,3-dimethyl-1,4-dihydropyranopyridine-5-carbonitrile (three-letter code: 5QF) (formula: C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	C	1	Total	C	N	O	S	0	0
			37	28	4	4	1		

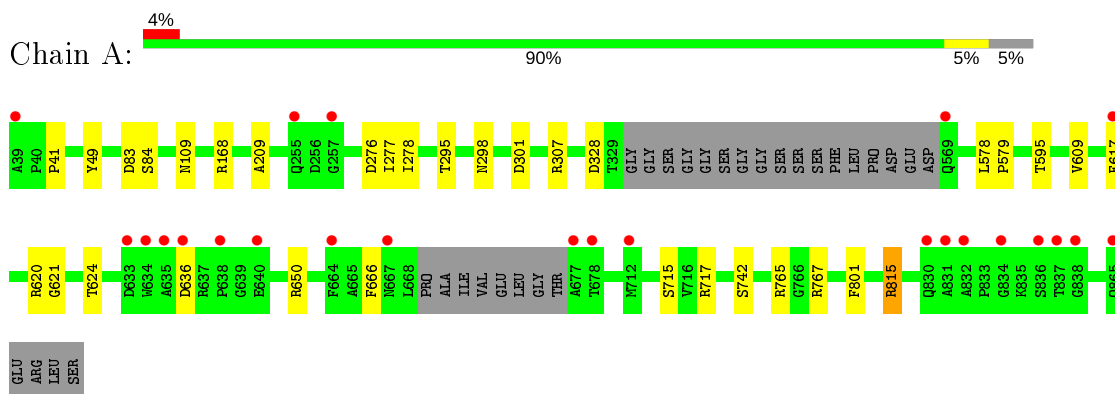
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	385	Total 385	O 385	0	0
4	B	385	Total 385	O 385	0	0
4	C	354	Total 354	O 354	0	0
4	D	64	Total 64	O 64	0	0
4	E	45	Total 45	O 45	0	0
4	F	49	Total 49	O 49	0	0

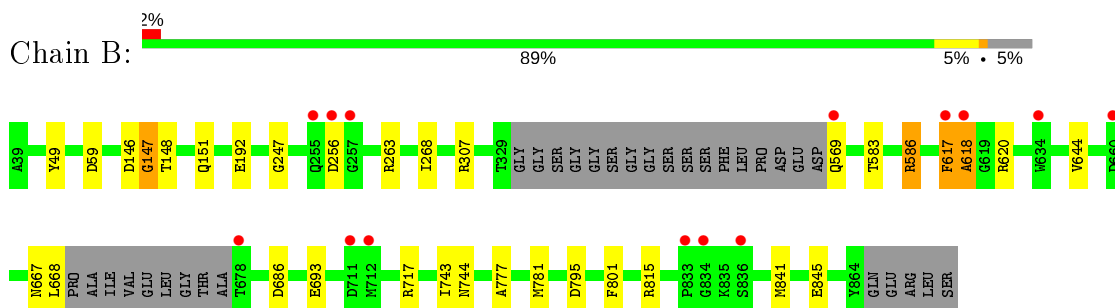
### 3 Residue-property plots [i](#)

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.

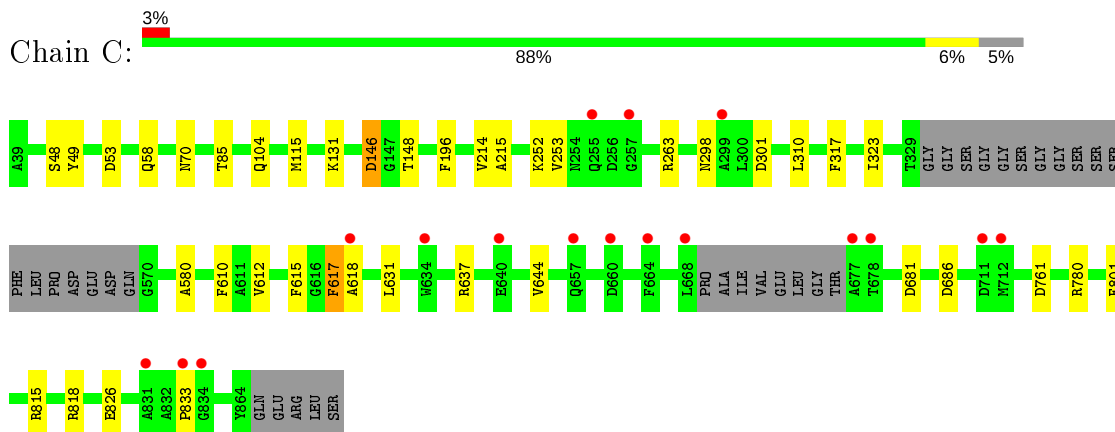
- Molecule 1: Multidrug efflux pump subunit AcrB, Multidrug efflux pump subunit AcrB



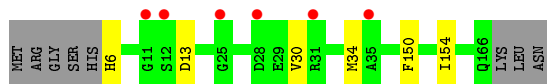
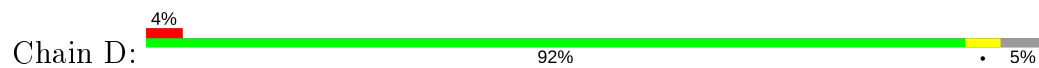
- Molecule 1: Multidrug efflux pump subunit AcrB, Multidrug efflux pump subunit AcrB



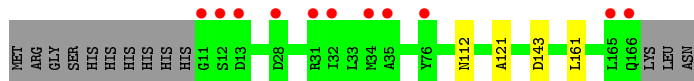
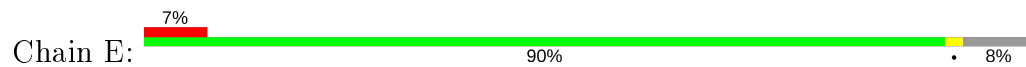
- Molecule 1: Multidrug efflux pump subunit AcrB, Multidrug efflux pump subunit AcrB



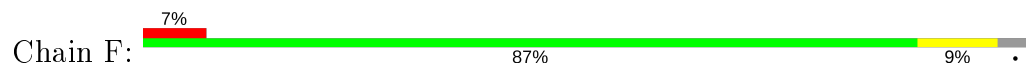
- Molecule 2: DARPin



- Molecule 2: DARPin



- Molecule 2: DARPin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.65Å 145.34Å 174.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 49.64 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-1.90) 99.8 (49.64-1.90)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 1.90Å)	Xtrriage
Refinement program	REFMAC 5.8.0131	Depositor
R, $R_{free}$	0.182 , 0.222 0.191 , 0.228	Depositor DCC
$R_{free}$ test set	10661 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.5	Xtrriage
Anisotropy	0.234	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 36.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	18206	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: 5QF

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	0.96	2/4515 (0.0%)	0.93	10/6119 (0.2%)
1	B	0.93	0/4492	0.96	10/6086 (0.2%)
1	C	0.90	1/4482 (0.0%)	0.94	15/6073 (0.2%)
2	D	0.88	0/1251	0.89	1/1701 (0.1%)
2	E	0.79	0/1196	0.90	2/1626 (0.1%)
2	F	0.83	1/1262 (0.1%)	0.88	2/1716 (0.1%)
All	All	0.91	4/17198 (0.0%)	0.93	40/23321 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	84	SER	CB-OG	8.54	1.53	1.42
1	C	48	SER	CB-OG	-5.42	1.35	1.42
1	A	765	ARG	CZ-NH2	5.26	1.39	1.33
2	F	57	TRP	CB-CG	5.08	1.59	1.50

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	815	ARG	NE-CZ-NH1	9.42	125.01	120.30
1	A	717	ARG	NE-CZ-NH1	8.12	124.36	120.30
1	C	761	ASP	CB-CG-OD1	7.69	125.22	118.30
1	C	263	ARG	NE-CZ-NH1	6.89	123.74	120.30
1	A	767	ARG	NE-CZ-NH2	6.86	123.73	120.30

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	4431	0	4377	14	0
1	B	4411	0	4353	17	0
1	C	4404	0	4345	19	0
2	D	1227	0	1194	4	0
2	E	1177	0	1159	2	0
2	F	1237	0	1201	5	0
3	C	37	0	0	0	0
4	A	385	0	0	4	0
4	B	385	0	0	6	0
4	C	354	0	0	5	0
4	D	64	0	0	1	0
4	E	45	0	0	1	0
4	F	49	0	0	0	0
All	All	18206	0	16629	60	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 2.

The worst 5 of 60 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:298:ASN:HD22	1:A:301:ASP:H	1.37	0.73
2:D:150:PHE:CE2	2:D:154:ILE:HD11	2.24	0.72
1:A:620:ARG:HD3	4:A:902:HOH:O	1.90	0.71
1:C:58:GLN:OE1	4:C:1001:HOH:O	2.10	0.70
1:B:620:ARG:HD3	4:B:902:HOH:O	1.91	0.70

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	577/609 (95%)	562 (97%)	15 (3%)	0	100	100
1	B	574/609 (94%)	561 (98%)	11 (2%)	2 (0%)	41	31
1	C	573/609 (94%)	555 (97%)	17 (3%)	1 (0%)	47	38
2	D	159/169 (94%)	155 (98%)	4 (2%)	0	100	100
2	E	154/169 (91%)	150 (97%)	4 (3%)	0	100	100
2	F	160/169 (95%)	156 (98%)	3 (2%)	1 (1%)	25	15
All	All	2197/2334 (94%)	2139 (97%)	54 (2%)	4 (0%)	47	38

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	11	GLY
1	B	618	ALA
1	C	833	PRO
1	B	147	GLY

### 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	474/492 (96%)	468 (99%)	6 (1%)	69	68
1	B	472/492 (96%)	467 (99%)	5 (1%)	73	73
1	C	470/492 (96%)	467 (99%)	3 (1%)	86	87
2	D	125/132 (95%)	125 (100%)	0	100	100
2	E	120/132 (91%)	120 (100%)	0	100	100
2	F	126/132 (96%)	125 (99%)	1 (1%)	81	82
All	All	1787/1872 (96%)	1772 (99%)	15 (1%)	81	82

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	49	TYR
1	B	151	GLN
1	C	610	PHE
1	A	801	PHE
1	C	49	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	231	ASN
1	B	255	GLN
1	C	298	ASN
1	B	112	GLN
1	C	284	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is 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	5QF	C	901	-	40,40,40	1.87	7 (17%)	49,56,56	2.29	16 (32%)

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	5QF	C	901	-	-	7/19/41/41	0/4/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	901	5QF	C6-C7	-7.50	1.30	1.44
3	C	901	5QF	C5-S9	-4.88	1.70	1.76
3	C	901	5QF	O34-C33	-3.59	1.42	1.45
3	C	901	5QF	O34-C35	-2.63	1.40	1.43
3	C	901	5QF	C6-C5	-2.62	1.38	1.41

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	901	5QF	C21-O19-C15	7.39	128.68	117.53
3	C	901	5QF	O19-C15-C16	6.09	123.89	115.41
3	C	901	5QF	C10-S9-C5	4.51	108.92	101.82
3	C	901	5QF	C27-C26-N25	3.91	118.66	110.64
3	C	901	5QF	C2-C3-N4	-3.45	117.29	122.61

There are no chirality outliers.

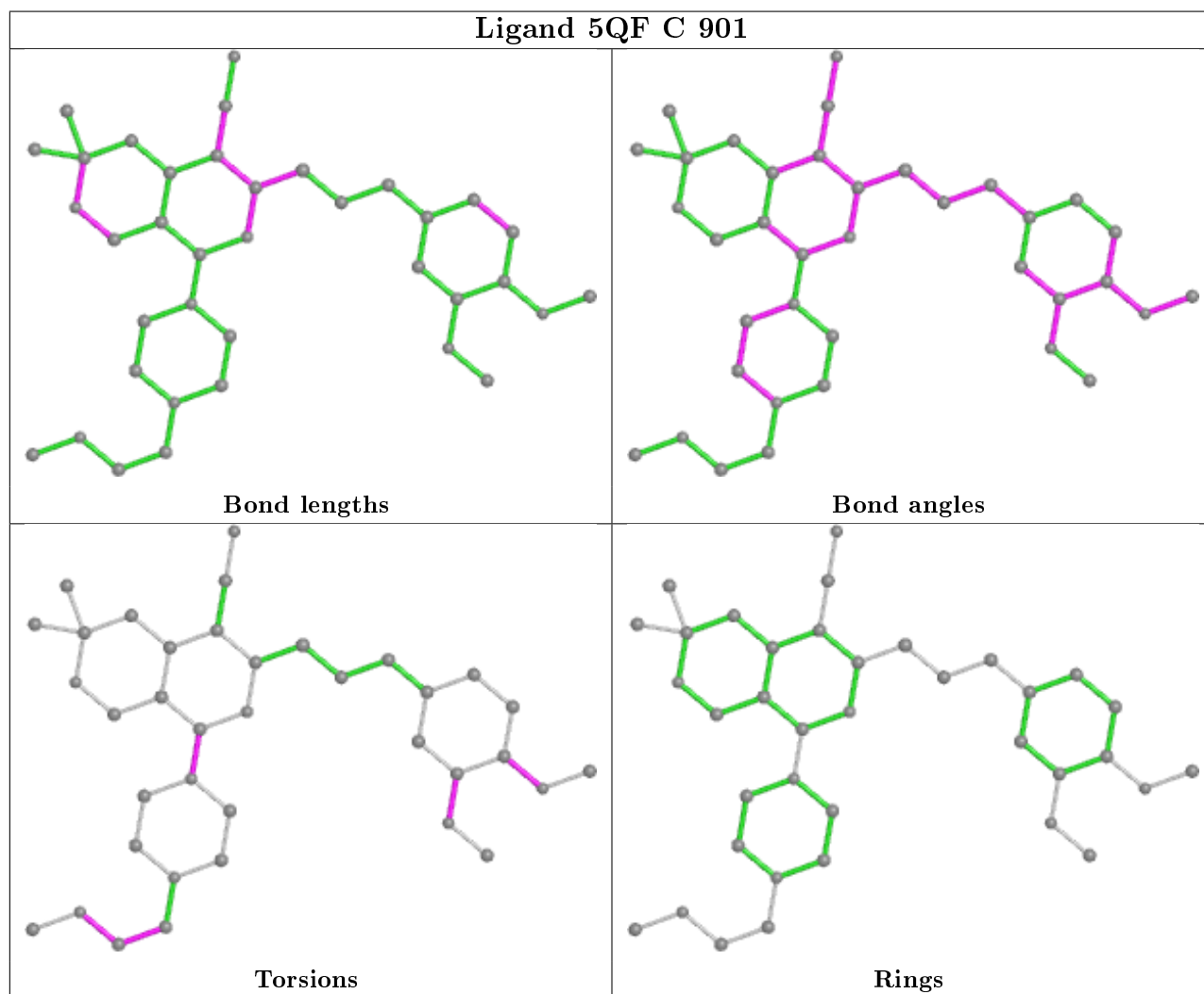
5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	901	5QF	C15-C16-O18-C20
3	C	901	5QF	C17-C16-O18-C20
3	C	901	5QF	C14-C15-O19-C21
3	C	901	5QF	N25-C28-C29-O30
3	C	901	5QF	C16-C15-O19-C21

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 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	580/609 (95%)	0.17	24 (4%) 37 40	18, 29, 56, 92	0
1	B	578/609 (94%)	0.04	14 (2%) 59 62	18, 28, 53, 104	0
1	C	578/609 (94%)	0.14	17 (2%) 51 54	17, 31, 61, 98	0
2	D	161/169 (95%)	0.11	6 (3%) 41 44	24, 34, 61, 79	0
2	E	156/169 (92%)	0.37	11 (7%) 16 17	27, 40, 70, 125	0
2	F	162/169 (95%)	0.26	11 (6%) 17 19	26, 37, 71, 104	0
All	All	2215/2334 (94%)	0.14	83 (3%) 41 44	17, 31, 61, 125	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	11	GLY	16.0
1	A	677	ALA	6.1
1	C	677	ALA	5.9
2	E	12	SER	5.8
1	A	634	TRP	4.7

### 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 carbohydrates in this entry.

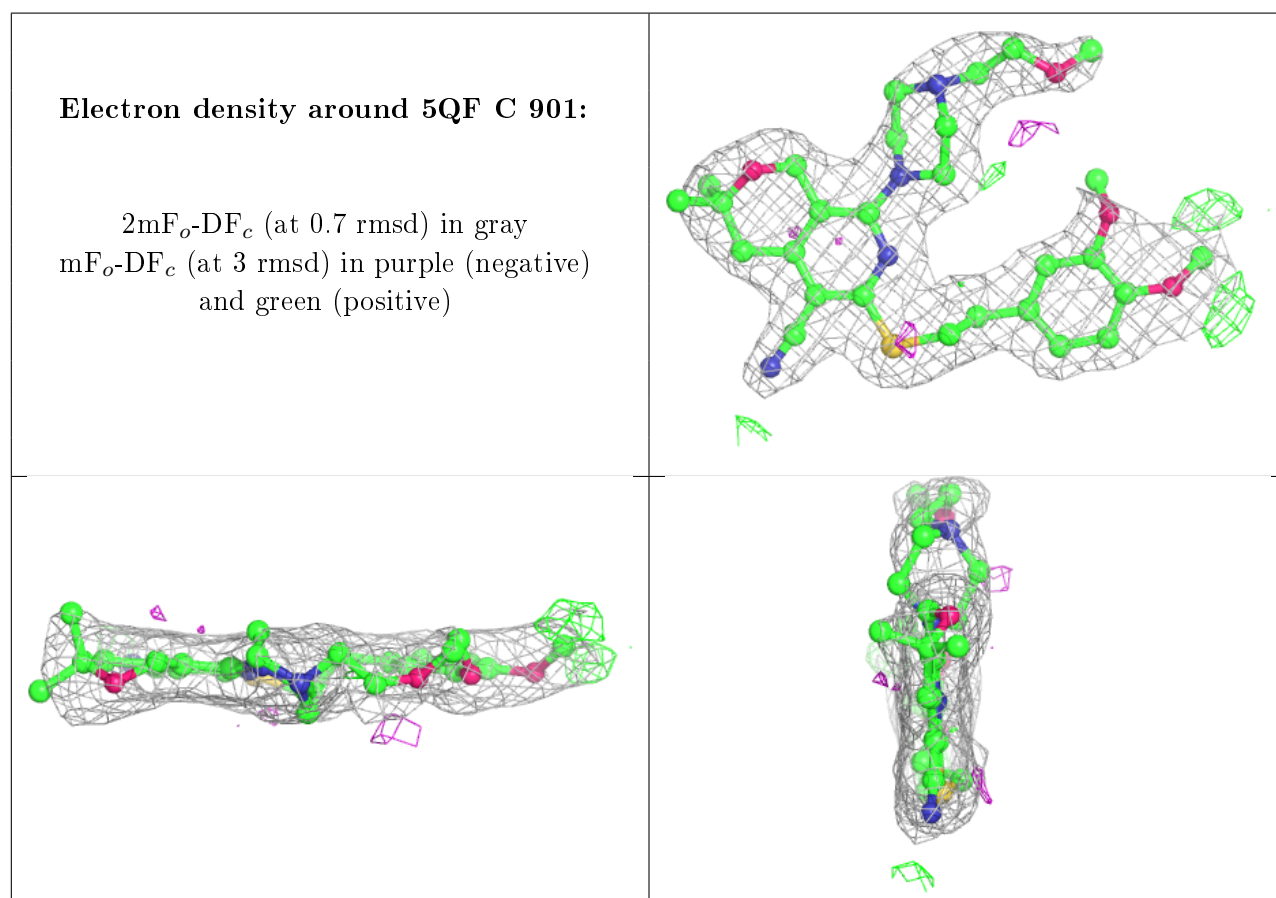


## 6.4 Ligands [i](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	5QF	C	901	37/37	0.93	0.13	37,47,69,71	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.