

# Full wwPDB X-ray Structure Validation Report (i)

#### May 5, 2025 - 03:46 pm BST

PDB ID	:	$9\mathrm{F78}\/\mathrm{pdb}\_00009\mathrm{f78}$
Title	:	BAZ2A bromodomain in complex with acetylpyrrole derivative compound 28-
		TN23
Authors	:	Dalle Vedove, A.; Cazzanelli, G.; Caflisch, A.; Lolli, G.
Deposited on	:	2024-05-03
Resolution	:	1.70  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 (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

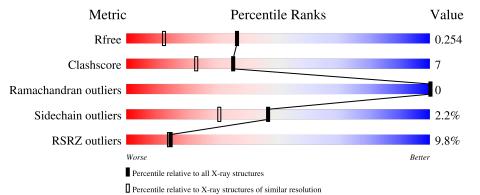
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as $541$ be (2020)
Xtriage (Phenix)	:	2.0rc1
$\mathrm{EDS}$	:	3.0
buster-report	:	1.1.7(2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
		9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY\;DIFFRACTION$ 

The reported resolution of this entry is 1.70 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	164625	5161(1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (1.70-1.70)

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	А	105	8%	16% ••
1	В	105	88%	10% •



## 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2049 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 Bromodomain adjacent to zinc finger domain protein 2A.

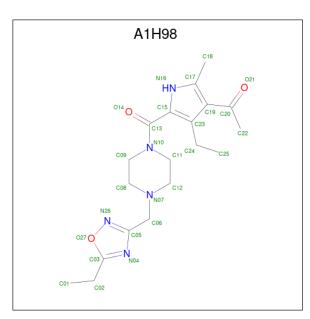
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Λ	102	Total	С	Ν	0	S	0	3	0
		102	855	541	142	164	8	0		
1	В	102	Total	С	Ν	0	S	0	6	0
	ГВ	103	884	560	151	166	7	0		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	1794	SER	-	expression tag	UNP Q9UIF9
А	1795	MET	-	expression tag	UNP Q9UIF9
А	1848	SER	LEU	engineered mutation	UNP Q9UIF9
В	1794	SER	-	expression tag	UNP Q9UIF9
В	1795	MET	-	expression tag	UNP Q9UIF9
В	1848	SER	LEU	engineered mutation	UNP Q9UIF9

• Molecule 2 is 1-[4-ethyl-5-[4-[(5-ethyl-1,2,4-oxadiazol-3-yl)methyl]piperazin-1-yl]carbonyl-2 -methyl-1 {H}-pyrrol-3-yl]ethanone (CCD ID: A1H98) (formula: C<sub>19</sub>H<sub>27</sub>N<sub>5</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	А	1	Total 27			0	0
2	В	1	Total 27			0	0

• Molecule 3 is water.

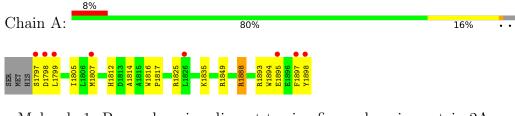
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	124	Total O 124 124	0	0
3	В	132	Total         O           132         132	0	0



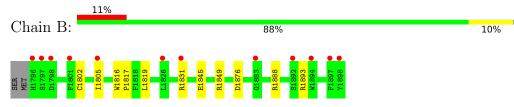
# 3 Residue-property plots (i)

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: Bromodomain adjacent to zinc finger domain protein 2A



• Molecule 1: Bromodomain adjacent to zinc finger domain protein 2A





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	43.02Å 54.36Å 90.16Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	38.83 - 1.70	Depositor
Resolution (A)	38.83 - 1.70	EDS
% Data completeness	95.8 (38.83-1.70)	Depositor
(in resolution range)	95.8 (38.83-1.70)	EDS
R <sub>merge</sub>	0.07	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.58 (at 1.70 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
D D	0.209 , $0.252$	Depositor
$R, R_{free}$	0.211 , $0.254$	DCC
$R_{free}$ test set	1208 reflections $(5.05\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	21.8	Xtriage
Anisotropy	0.708	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.33 , $45.5$	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.42, < L^2>=0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2049	wwPDB-VP
Average B, all atoms $(Å^2)$	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 12.61% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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:  $\rm A1H98$ 

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		lengths	Bond angles		
		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.32	0/885	0.49	0/1191	
1	В	0.29	0/924	0.46	0/1243	
All	All	0.30	0/1809	0.48	0/2434	

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	А	855	0	805	14	0
1	В	884	0	845	11	0
2	А	27	0	0	0	0
2	В	27	0	0	0	0
3	А	124	0	0	3	0
3	В	132	0	0	6	0
All	All	2049	0	1650	25	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 (25) 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:1812:HIS:HD2	1:A:1814:ALA:H	1.17	0.92
1:A:1888:ARG:HH11	1:A:1888:ARG:HG2	1.52	0.73
1:B:1831:ARG:NE	3:B:2001:HOH:O	2.23	0.71
1:A:1812:HIS:CD2	1:A:1814:ALA:H	2.06	0.66
1:A:1807[B]:MET:SD	3:A:2024:HOH:O	2.55	0.65
1:B:1849[A]:ARG:NH1	3:B:2002:HOH:O	2.34	0.59
1:A:1888:ARG:HG2	1:A:1888:ARG:NH1	2.19	0.57
1:A:1849:ARG:NH2	3:A:2007:HOH:O	2.41	0.54
1:B:1849[B]:ARG:NH1	3:B:2002:HOH:O	2.36	0.53
1:B:1849[B]:ARG:NH1	3:B:2009:HOH:O	2.42	0.53
1:B:1816:TRP:CG	1:B:1817:PRO:HD3	2.46	0.51
1:B:1845:GLU:OE1	1:B:1849[B]:ARG:NH2	2.47	0.47
1:B:1888[A]:ARG:HG3	3:B:2057:HOH:O	2.13	0.47
1:B:1888[B]:ARG:HG3	3:B:2057:HOH:O	2.16	0.46
1:A:1805:ILE:HD11	1:A:1897:PHE:HD2	1.81	0.45
1:A:1797:SER:N	3:A:2011:HOH:O	2.50	0.45
1:B:1893:ARG:HD2	1:B:1893:ARG:HA	1.77	0.44
1:A:1895:GLU:N	1:A:1895:GLU:OE1	2.51	0.44
1:A:1893:ARG:HA	1:A:1893:ARG:HD2	1.87	0.43
1:A:1894:TRP:HB3	1:A:1895:GLU:OE1	2.19	0.42
1:B:1876:ASP:OD1	1:B:1876:ASP:N	2.53	0.42
1:A:1799:LEU:HD13	1:A:1799:LEU:HA	1.88	0.42
1:B:1802:CYS:HA	1:B:1805:ILE:HG22	2.02	0.42
1:A:1816:TRP:CG	1:A:1817:PRO:HD3	2.55	0.41
1:A:1798:ASP:O	1:A:1898:TYR:OH	2.38	0.41

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	Percer	ntiles
1	А	103/105~(98%)	102 (99%)	1 (1%)	0	100	100
1	В	107/105~(102%)	106 (99%)	1 (1%)	0	100	100
All	All	210/210~(100%)	208~(99%)	2(1%)	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 side chain 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	А	94/94~(100%)	91~(97%)	3~(3%)	34 17
1	В	98/94 (104%)	97~(99%)	1 (1%)	73 64
All	All	192/188~(102%)	188~(98%)	4 (2%)	47 32

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	1825	ARG
1	А	1835	LYS
1	А	1888	ARG
1	В	1819	LEU

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such side chains are listed below:

Mol	Chain	Res	Type
1	А	1812	HIS
1	А	1870	GLN
1	В	1870	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 oligosaccharides 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	Mol Type Chain		in Res Link		Bo	Bond lengths			Bond angles		
	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2	
2	A1H98	А	1901	-	21,29,29	<mark>3.63</mark>	9 (42%)	22,41,41	1.26	3 (13%)	
2	A1H98	В	1901	-	21,29,29	3.62	8 (38%)	22,41,41	1.39	3 (13%)	

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
2	A1H98	А	1901	-	-	2/10/30/30	0/3/3/3
2	A1H98	В	1901	-	-	1/10/30/30	0/3/3/3

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$\mathrm{Ideal}(\mathrm{\AA})$
2	А	1901	A1H98	C06-N07	-10.85	1.26	1.47
2	В	1901	A1H98	C06-N07	-10.63	1.27	1.47
2	В	1901	A1H98	C02-C03	7.29	1.54	1.49
2	А	1901	A1H98	C02-C03	6.76	1.53	1.49
2	В	1901	A1H98	C13-N10	6.34	1.48	1.34
2	А	1901	A1H98	C13-N10	6.21	1.48	1.34

Continued on next page...



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	В	1901	A1H98	C06-C05	5.06	1.55	1.50
2	А	1901	A1H98	C06-C05	4.52	1.55	1.50
2	А	1901	A1H98	C11-N10	-3.34	1.41	1.47
2	В	1901	A1H98	C09-N10	-2.84	1.42	1.47
2	А	1901	A1H98	C09-N10	-2.71	1.42	1.47
2	А	1901	A1H98	C12-N07	-2.61	1.39	1.46
2	А	1901	A1H98	C08-N07	-2.58	1.39	1.46
2	А	1901	A1H98	O14-C13	-2.48	1.17	1.22
2	В	1901	A1H98	C11-N10	-2.46	1.42	1.47
2	В	1901	A1H98	C19-C20	2.36	1.55	1.49
2	В	1901	A1H98	C12-N07	-2.28	1.40	1.46

Continued from previous page...

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	В	1901	A1H98	C12-N07-C08	3.28	116.20	108.83
2	А	1901	A1H98	C12-N07-C08	3.05	115.69	108.83
2	А	1901	A1H98	C06-C05-N04	2.84	127.15	123.09
2	В	1901	A1H98	C06-C05-N26	2.28	126.35	123.09
2	В	1901	A1H98	C06-C05-N04	2.16	126.18	123.09
2	А	1901	A1H98	C17-C19-C23	2.13	107.97	104.34

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	А	1901	A1H98	C15-C23-C24-C25
2	А	1901	A1H98	C19-C23-C24-C25
2	В	1901	A1H98	C19-C23-C24-C25

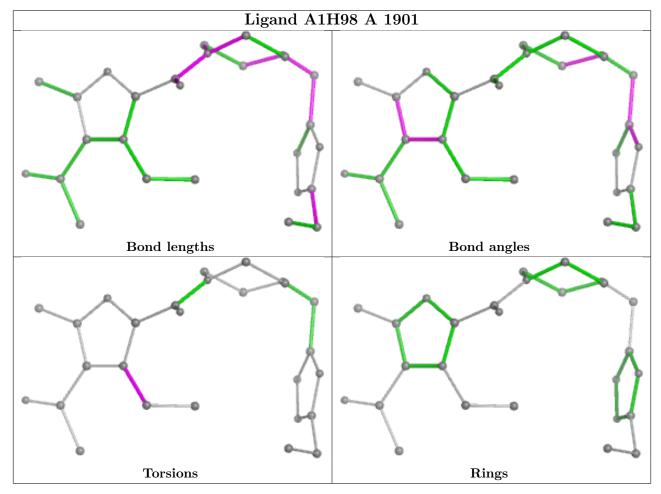
There are no ring outliers.

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

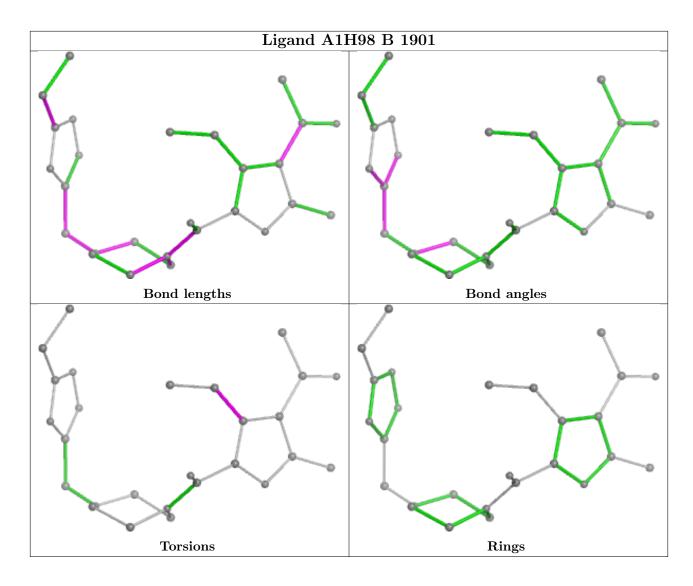


any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.









## 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,  $95^{th}$  percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	102/105~(97%)	0.60	8 (7%) 20 21	16, 27, 55, 66	3(2%)
1	В	103/105~(98%)	0.75	12 (11%) 10 10	17, 26, 57, 75	6 (5%)
All	All	205/210~(97%)	0.68	20 (9%) 14 14	16, 27, 57, 75	9 (4%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	1801	PHE	4.6
1	А	1898	TYR	3.9
1	В	1897	PHE	3.3
1	А	1897	PHE	3.1
1	В	1805	ILE	2.9
1	В	1898	TYR	2.9
1	В	1883	GLY	2.7
1	В	1796	HIS	2.7
1	А	1798	ASP	2.5
1	В	1797	SER	2.5
1	В	1894	TRP	2.4
1	А	1895	GLU	2.4
1	А	1799	LEU	2.3
1	А	1797	SER	2.2
1	В	1798	ASP	2.1
1	В	1826	LEU	2.1
1	В	1892	SER	2.0
1	А	1826	LEU	2.0
1	В	1831	ARG	2.0
1	А	1807[A]	MET	2.0



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

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

### 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

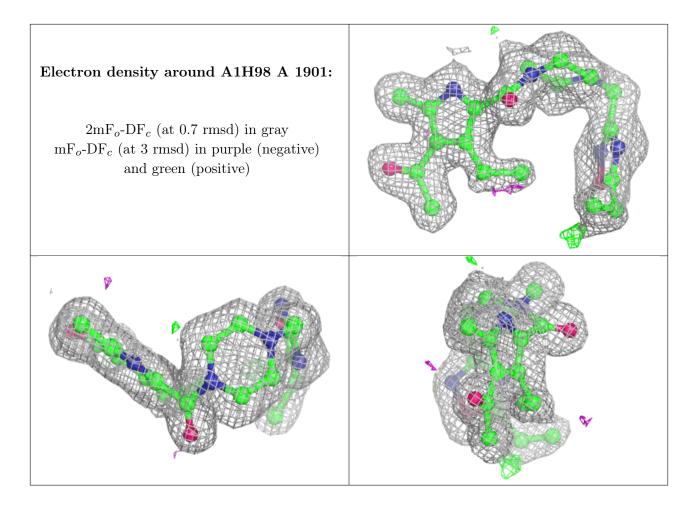
### 6.4 Ligands (i)

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

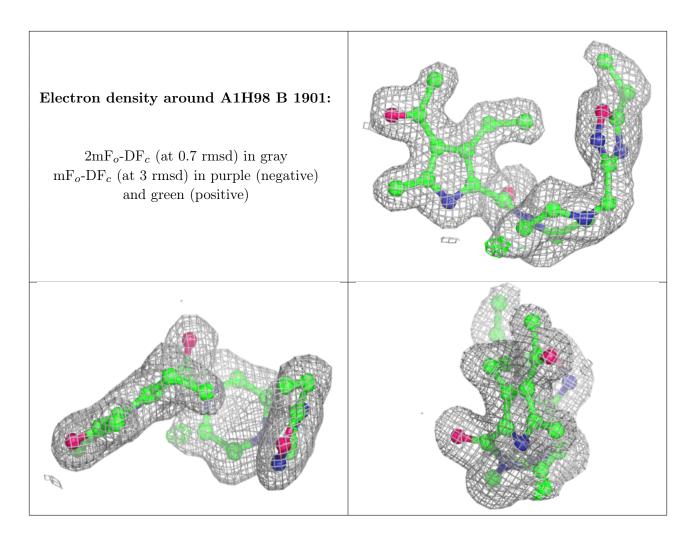
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
2	A1H98	A	1901	27/27	0.92	0.09	13,24,30,35	0
2	A1H98	В	1901	27/27	0.92	0.09	16,25,32,39	0

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.









## 6.5 Other polymers (i)

There are no such residues in this entry.

