



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 12:23 PM UTC

PDB ID : 9IBY / pdb\_00009iby  
Title : Crystal structure of Zika Virus NS2B-NS3 protease in complex with compound 2  
Authors : Ontario, J.M.; Torrente, E.  
Deposited on : 2025-02-14  
Resolution : 2.29 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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

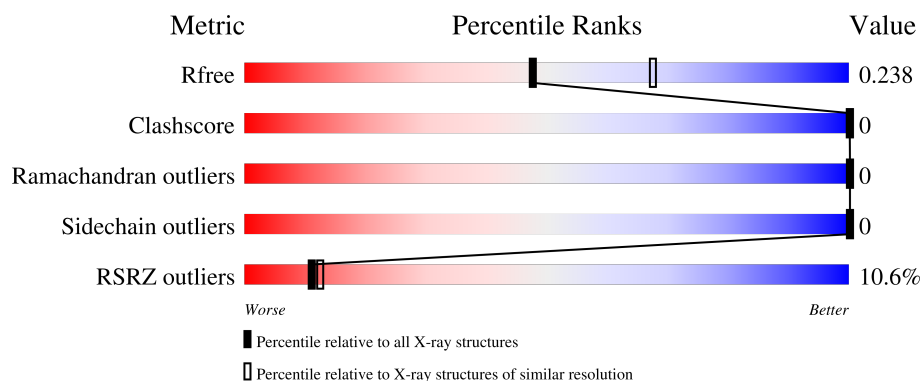
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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))
$R_{free}$	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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  $\geq 3$ , 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	230	 7% 69% 31%
1	B	230	 7% 71% 28%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2608 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 Genome polypeptide.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	159	Total	As	C	N	O	S	44	1	0
			1215	1	765	209	234	6			
1	B	165	Total	As	C	N	O	S	46	1	0
			1259	1	792	218	241	7			

There are 30 discrepancies between the modelled and reference sequences:

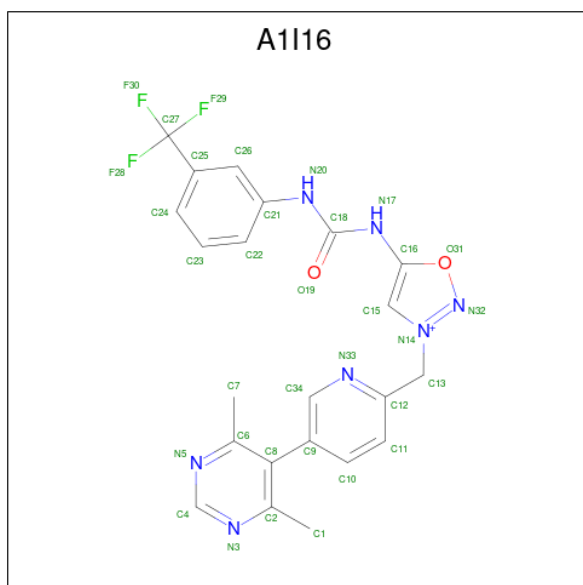
Chain	Residue	Modelled	Actual	Comment	Reference
A	45	GLY	-	expression tag	UNP A0A140DLX4
A	46	SER	-	expression tag	UNP A0A140DLX4
A	47	HIS	-	expression tag	UNP A0A140DLX4
A	48	MET	-	expression tag	UNP A0A140DLX4
A	991	ALA	-	linker	UNP A0A140DLX4
A	992	GLY	-	linker	UNP A0A140DLX4
A	993	GLY	-	linker	UNP A0A140DLX4
A	994	GLY	-	linker	UNP A0A140DLX4
A	995	GLY	-	linker	UNP A0A140DLX4
A	996	SER	-	linker	UNP A0A140DLX4
A	997	GLY	-	linker	UNP A0A140DLX4
A	998	GLY	-	linker	UNP A0A140DLX4
A	999	GLY	-	linker	UNP A0A140DLX4
A	1000	GLY	-	linker	UNP A0A140DLX4
A	1029	GLY	ARG	conflict	UNP A0A140DLX4
B	45	GLY	-	expression tag	UNP A0A140DLX4
B	46	SER	-	expression tag	UNP A0A140DLX4
B	47	HIS	-	expression tag	UNP A0A140DLX4
B	48	MET	-	expression tag	UNP A0A140DLX4
B	991	ALA	-	linker	UNP A0A140DLX4
B	992	GLY	-	linker	UNP A0A140DLX4
B	993	GLY	-	linker	UNP A0A140DLX4
B	994	GLY	-	linker	UNP A0A140DLX4
B	995	GLY	-	linker	UNP A0A140DLX4
B	996	SER	-	linker	UNP A0A140DLX4

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	997	GLY	-	linker	UNP A0A140DLX4
B	998	GLY	-	linker	UNP A0A140DLX4
B	999	GLY	-	linker	UNP A0A140DLX4
B	1000	GLY	-	linker	UNP A0A140DLX4
B	1029	GLY	ARG	conflict	UNP A0A140DLX4

- Molecule 2 is 1-[3-[[5-(4,6-dimethylpyrimidin-5-yl)pyridin-2-yl]methyl]-1-oxa-2,3<sup>l</sup>{4}-diazacyclopenta-2,4-dien-5-yl]-3-[3-(trifluoromethyl)phenyl]urea (CCD ID: A1I16) (formula: C<sub>22</sub>H<sub>19</sub>F<sub>3</sub>N<sub>7</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			34	22	3	7	2		
2	B	1	Total	C	F	N	O	0	0
			34	22	3	7	2		

- Molecule 3 is ACETIC ACID (CCD ID: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).



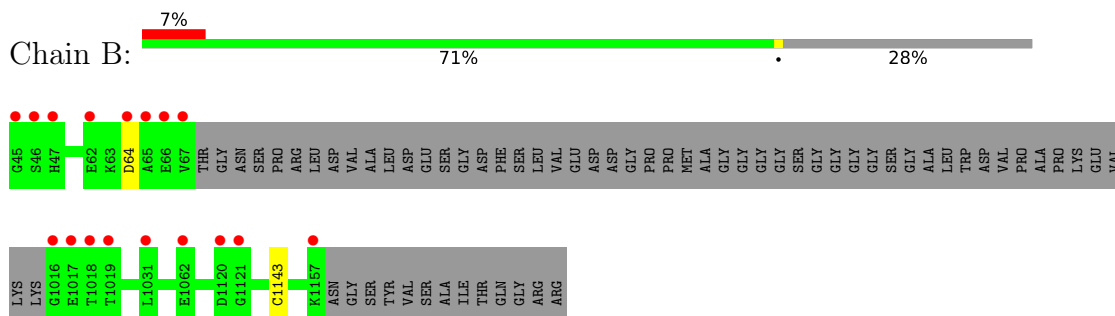
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	23	Total	O	0	0
			23	23		
4	B	31	Total	O	0	0
			31	31		

i

- Molecule 1: Genome polypeptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.79Å 57.47Å 72.45Å 90.00° 96.41° 90.00°	Depositor
Resolution (Å)	44.96 – 2.29 44.96 – 2.29	Depositor EDS
% Data completeness (in resolution range)	62.6 (44.96-2.29) 62.6 (44.96-2.29)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, $R_{free}$	0.189 , 0.236 0.195 , 0.238	Depositor DCC
$R_{free}$ test set	733 reflections (3.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.9	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 32.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2608	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.31% 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: A1I16, ACY, CAF

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	0.54	0/1230	0.92	0/1667
1	B	0.53	0/1275	0.91	1/1726 (0.1%)
All	All	0.54	0/2505	0.91	1/3393 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	64	ASP	CA-CB-CG	7.00	119.60	112.60

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	1215	0	1177	0	0
1	B	1259	0	1223	0	0
2	A	34	0	0	0	0
2	B	34	0	0	0	0
3	A	4	0	3	0	0
3	B	8	0	6	0	0
4	A	23	0	0	0	0
4	B	31	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2608	0	2409	0	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 0.

There are no clashes within the asymmetric unit.

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	155/230 (67%)	150 (97%)	5 (3%)	0	100	100
1	B	161/230 (70%)	157 (98%)	4 (2%)	0	100	100
All	All	316/460 (69%)	307 (97%)	9 (3%)	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	125/175 (71%)	125 (100%)	0	100	100
1	B	130/175 (74%)	130 (100%)	0	100	100
All	All	255/350 (73%)	255 (100%)	0	100	100

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

Sometimes 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	CAF	B	1143	1	5,9,10	0.73	0	4,12,14	1.25	1 (25%)
1	CAF	A	1143	1	5,9,10	0.55	0	4,12,14	0.72	0

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	CAF	B	1143	1	-	0/0/8/10	-
1	CAF	A	1143	1	-	0/0/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1143	CAF	CB-CA-C	-2.32	104.50	110.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

5 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	$\# Z  > 2$
2	A1I16	A	1201	-	34,37,37	2.13	1 (2%)	40,53,53	2.10	6 (15%)
2	A1I16	B	1201	-	34,37,37	2.26	1 (2%)	40,53,53	2.16	9 (22%)
3	ACY	A	1202	-	3,3,3	0.99	0	3,3,3	0.97	0
3	ACY	B	1202	-	3,3,3	1.00	0	3,3,3	0.85	0
3	ACY	B	1203	-	3,3,3	1.08	0	3,3,3	0.78	0

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	A1I16	B	1201	-	-	2/20/22/22	0/4/4/4
2	A1I16	A	1201	-	-	2/20/22/22	0/4/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1201	A1I16	O31-N32	-12.29	1.22	1.37
2	A	1201	A1I16	O31-N32	-11.61	1.23	1.37

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1201	A1I16	C13-N14-N32	7.33	126.34	118.91
2	B	1201	A1I16	C13-N14-N32	7.21	126.22	118.91
2	B	1201	A1I16	C15-N14-N32	-5.10	107.17	113.56
2	B	1201	A1I16	C12-C13-N14	-5.05	105.24	111.68
2	A	1201	A1I16	C15-N14-N32	-4.83	107.51	113.56
2	A	1201	A1I16	N5-C4-N3	-4.56	121.67	128.58
2	A	1201	A1I16	C34-N33-C12	4.51	123.74	118.00
2	B	1201	A1I16	N5-C4-N3	-4.48	121.80	128.58
2	A	1201	A1I16	C12-C13-N14	-4.19	106.34	111.68
2	B	1201	A1I16	C34-N33-C12	3.50	122.45	118.00
2	A	1201	A1I16	C9-C34-N33	-2.73	119.94	124.28
2	B	1201	A1I16	F30-C27-C25	-2.68	107.16	112.90
2	B	1201	A1I16	C22-C21-C26	-2.48	116.65	119.66
2	B	1201	A1I16	C9-C34-N33	-2.33	120.58	124.28
2	B	1201	A1I16	C23-C24-C25	-2.01	118.71	120.75

There are no chirality outliers.

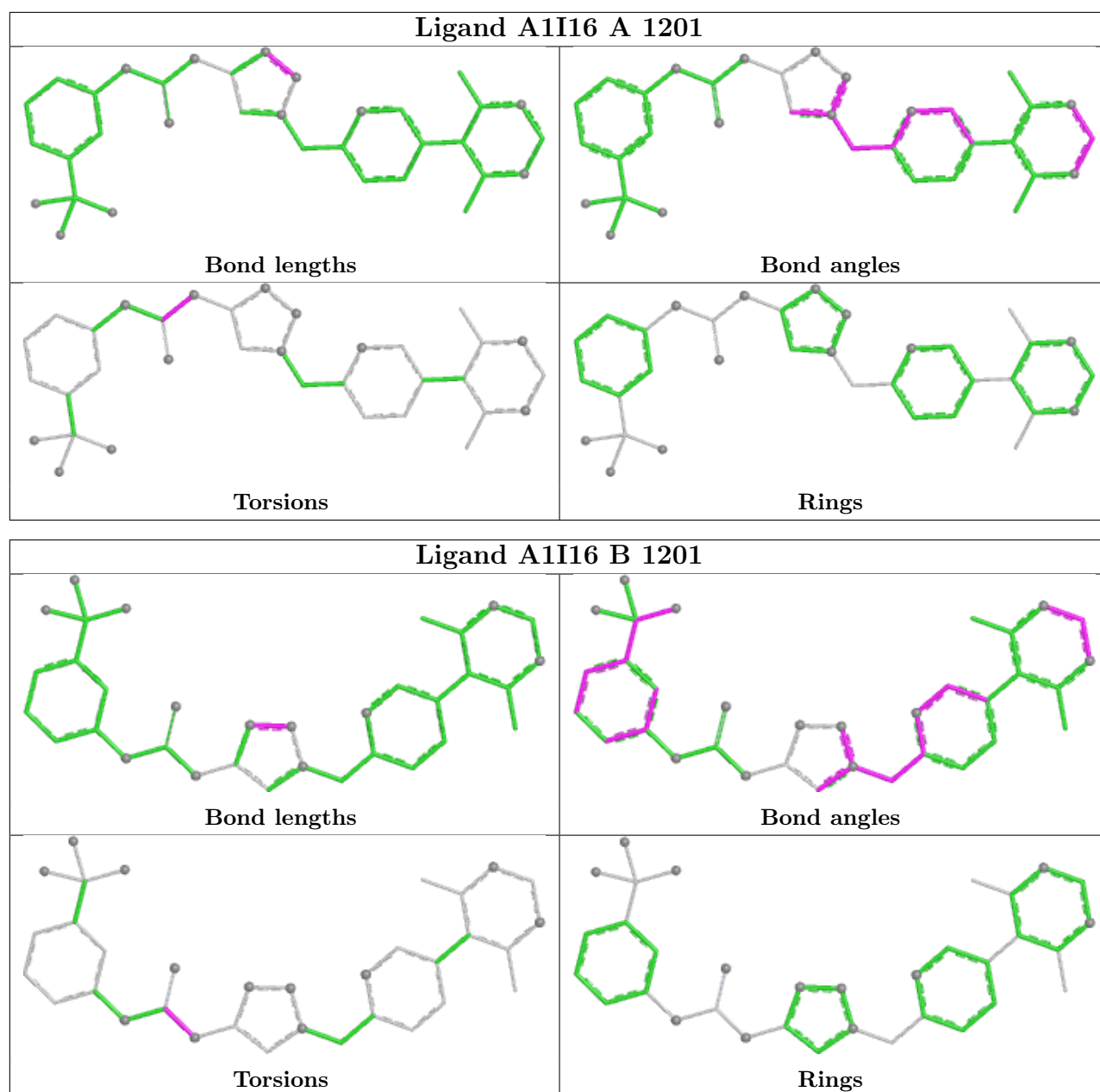
All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1201	A1I16	O19-C18-N17-C16
2	B	1201	A1I16	O19-C18-N17-C16
2	A	1201	A1I16	N20-C18-N17-C16
2	B	1201	A1I16	N20-C18-N17-C16

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 [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

### 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, 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	158/230 (68%)	0.49	17 (10%)	11 12	19, 48, 78, 112	17 (10%)
1	B	164/230 (71%)	0.40	17 (10%)	11 13	19, 46, 84, 112	16 (9%)
All	All	322/460 (70%)	0.44	34 (10%)	11 12	19, 47, 82, 112	33 (10%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	67	VAL	6.1
1	B	67	VAL	5.9
1	A	1017	GLU	5.4
1	A	1018	THR	5.4
1	B	66	GLU	5.2
1	B	65	ALA	5.2
1	B	1120	ASP	5.2
1	A	66	GLU	5.1
1	A	1156	ILE	5.0
1	B	1157	LYS	4.6
1	A	1016	GLY	4.4
1	A	1062	GLU	4.3
1	A	1120	ASP	4.3
1	B	1016	GLY	4.2
1	B	1017	GLU	3.9
1	A	63	LYS	3.8
1	B	1018	THR	3.7
1	B	64	ASP	3.4
1	A	1019	THR	3.0
1	B	45	GLY	3.0
1	B	1031	LEU	3.0
1	A	64	ASP	2.9
1	A	65	ALA	2.8
1	B	46	SER	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1119	LYS	2.6
1	B	47	HIS	2.4
1	A	1064	ARG	2.4
1	B	1019	THR	2.3
1	B	62	GLU	2.2
1	A	1061	GLY	2.1
1	B	1062	GLU	2.1
1	A	1063	GLY	2.1
1	A	1059	ARG	2.0
1	B	1121	GLY	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, 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
1	CAF	A	1143	10/11	0.97	0.10	41,54,86,88	0
1	CAF	B	1143	10/11	0.97	0.11	41,46,76,78	0

## 6.3 Carbohydrates [i](#)

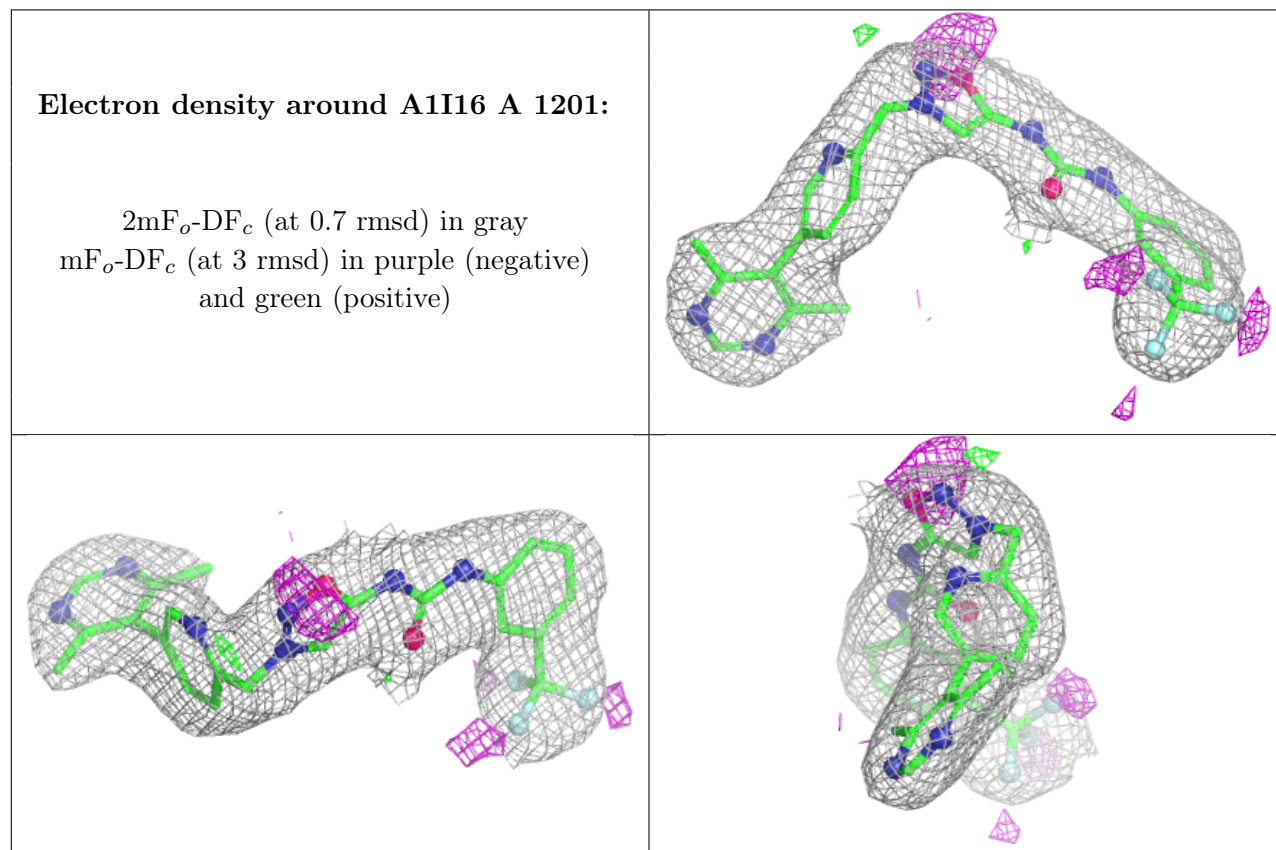
There are no oligosaccharides in this entry.

## 6.4 Ligands [i](#)

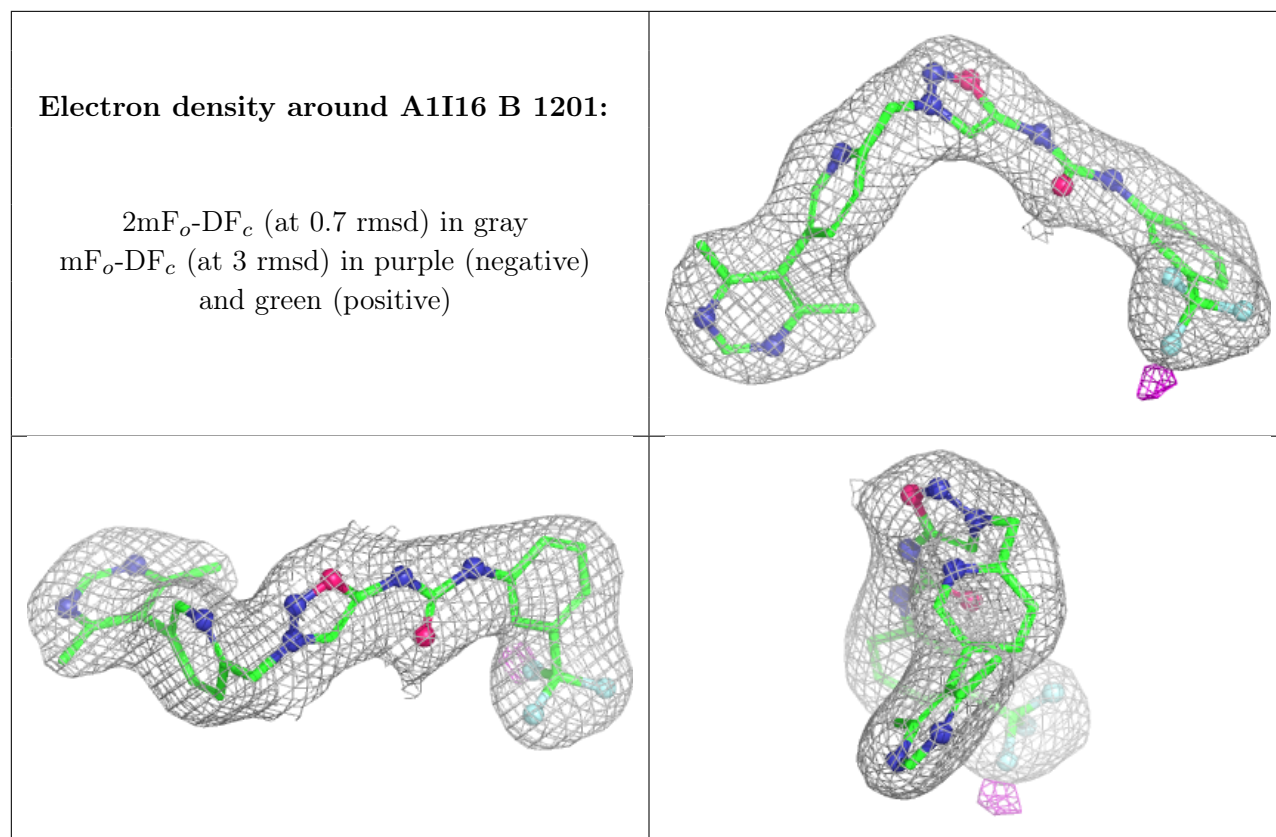
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	ACY	A	1202	4/4	0.78	0.23	70,72,72,73	0
3	ACY	B	1203	4/4	0.86	0.19	67,68,69,69	0
3	ACY	B	1202	4/4	0.88	0.19	68,68,72,73	0
2	A1I16	A	1201	34/34	0.95	0.08	39,45,50,51	0
2	A1I16	B	1201	34/34	0.97	0.07	34,44,48,49	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.