Full wwPDB X-ray Structure Validation Report (i)

Mar 20, 2023 – 06:21 PM EDT

PDB ID	:	7FVC
Title	:	PanDDA analysis group deposition – PHIP in complex with Z495704106
Authors	:	Grosjean, H.; Tomlinson, C.; Bradshaw, W.J.; Koekemoer, L.; Krojer, T.; Fearon, D.; Biggin, P.C.; von Delft, F.
Deposited on	:	2023-03-09
Resolution	:	1.18 Å(reported)

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

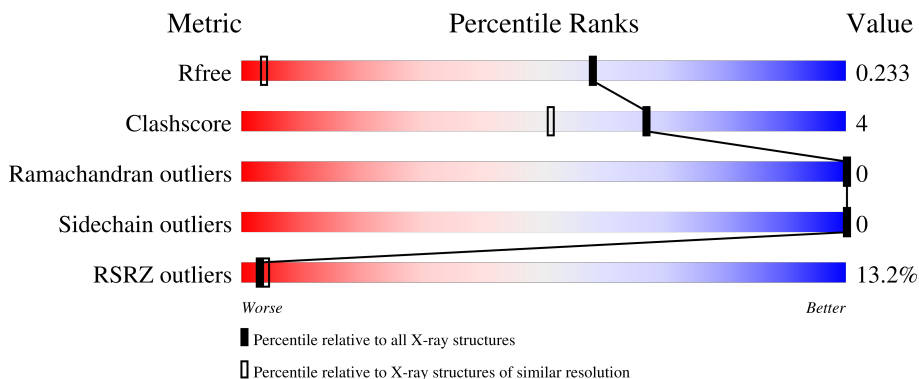|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

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 1.18 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		0.233
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		13.2%

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	1123 (1.20-1.16)
Clashscore	141614	1182 (1.20-1.16)
Ramachandran outliers	138981	1134 (1.20-1.16)
Sidechain outliers	138945	1134 (1.20-1.16)
RSRZ outliers	127900	1102 (1.20-1.16)

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	149	11% / 74% / 7% / 19%

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
2	ZM6	A	1901	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1269 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 PH-interacting protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	121	1033	652	172	202	7	0	5	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1292	MET	-	initiating methionine	UNP Q8WWQ0
A	1293	HIS	-	expression tag	UNP Q8WWQ0
A	1294	HIS	-	expression tag	UNP Q8WWQ0
A	1295	HIS	-	expression tag	UNP Q8WWQ0
A	1296	HIS	-	expression tag	UNP Q8WWQ0
A	1297	HIS	-	expression tag	UNP Q8WWQ0
A	1298	HIS	-	expression tag	UNP Q8WWQ0
A	1299	SER	-	expression tag	UNP Q8WWQ0
A	1300	SER	-	expression tag	UNP Q8WWQ0
A	1301	GLY	-	expression tag	UNP Q8WWQ0
A	1302	VAL	-	expression tag	UNP Q8WWQ0
A	1303	ASP	-	expression tag	UNP Q8WWQ0
A	1304	LEU	-	expression tag	UNP Q8WWQ0
A	1305	GLY	-	expression tag	UNP Q8WWQ0
A	1306	THR	-	expression tag	UNP Q8WWQ0
A	1307	GLU	-	expression tag	UNP Q8WWQ0
A	1308	ASN	-	expression tag	UNP Q8WWQ0
A	1309	LEU	-	expression tag	UNP Q8WWQ0
A	1310	TYR	-	expression tag	UNP Q8WWQ0
A	1311	PHE	-	expression tag	UNP Q8WWQ0
A	1312	GLN	-	expression tag	UNP Q8WWQ0
A	1313	SER	-	expression tag	UNP Q8WWQ0
A	1314	MET	-	expression tag	UNP Q8WWQ0

- Molecule 2 is N-(cyclopropylmethyl)-4-(3,4-dichlorobenzoyl)piperazine-1-carboxamide (three-letter code: ZM6) (formula: $C_{16}H_{19}Cl_2N_3O_2$) (labeled as "Ligand of Interest" by depositor).
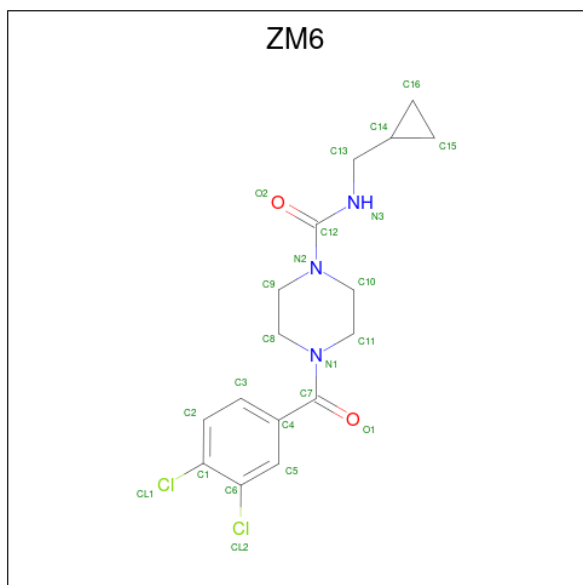
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	0	0
			23	16	2	3	2		

• Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	213	Total	O	0	0
			213	213		

3 Residue-property plots ⓘ

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: PH-interacting protein
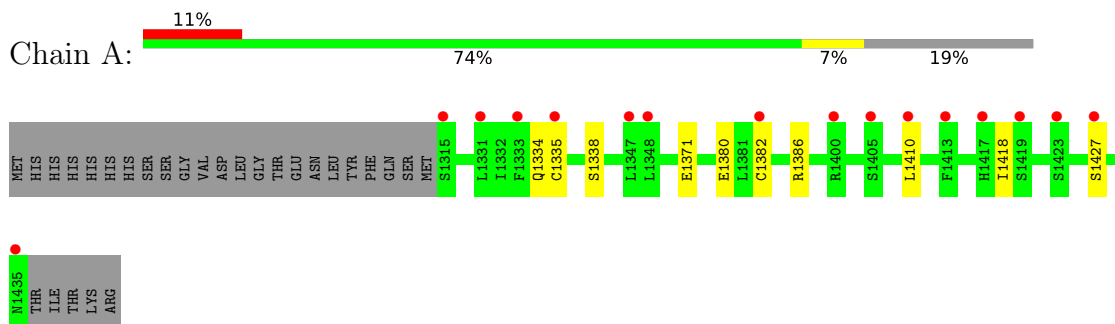
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	81.90Å 27.34Å 56.35Å 90.00° 100.05° 90.00°	Depositor
Resolution (Å)	27.76 – 1.18 27.74 – 1.18	Depositor EDS
% Data completeness (in resolution range)	81.9 (27.76-1.18) 81.9 (27.74-1.18)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 1.18Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.173 , 0.196 0.211 , 0.233	Depositor DCC
R_{free} test set	1662 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	13.6	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 46.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1269	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.02% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZM6

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	0.87	1/1056 (0.1%)	0.84	0/1424

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1380	GLU	CD-OE1	-5.28	1.19	1.25

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	1033	0	991	8	0
2	A	23	0	0	0	0
3	A	213	0	0	4	2
All	All	1269	0	991	8	2

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 4.

All (8) 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:1335:CYS:SG	3:A:2149:HOH:O	2.11	1.07
1:A:1382[A]:CYS:SG	1:A:1386:ARG:CZ	2.65	0.85
1:A:1371:GLU:HG2	3:A:2123:HOH:O	1.85	0.76
1:A:1338:SER:HA	1:A:1410:LEU:HD11	1.78	0.65
1:A:1382[B]:CYS:SG	1:A:1418:ILE:HG12	2.40	0.62
1:A:1427[B]:SER:HB2	3:A:2008:HOH:O	2.12	0.49
1:A:1334:GLN:CG	3:A:2210:HOH:O	2.63	0.47
1:A:1382[A]:CYS:SG	1:A:1386:ARG:NH2	2.90	0.43

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:2053:HOH:O	3:A:2155:HOH:O[4_446]	1.95	0.25
3:A:2168:HOH:O	3:A:2199:HOH:O[4_445]	2.08	0.12

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

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	124/149 (83%)	124 (100%)	0	0	100 100

There are no Ramachandran outliers to report.

5.3.2　Protein sidechains ⓘ

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	118/140 (84%)	118 (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 ⓘ

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

5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6　Ligand geometry ⓘ

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	#$|Z| > 2$
2	ZM6	A	1901	-	25,25,25	0.29	0	33,35,35	0.99	1 (3%)

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ZM6	A	1901	-	-	7/17/29/29	0/3/3/3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1901	ZM6	N3-C12-N2	-4.73	115.39	117.67

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1901	ZM6	N3-C12-N2-C10
2	A	1901	ZM6	N3-C12-N2-C9
2	A	1901	ZM6	O2-C12-N2-C10
2	A	1901	ZM6	C4-C7-N1-C8
2	A	1901	ZM6	O1-C7-N1-C8
2	A	1901	ZM6	O2-C12-N2-C9
2	A	1901	ZM6	C4-C7-N1-C11

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.

Ligand ZM6 A 1901
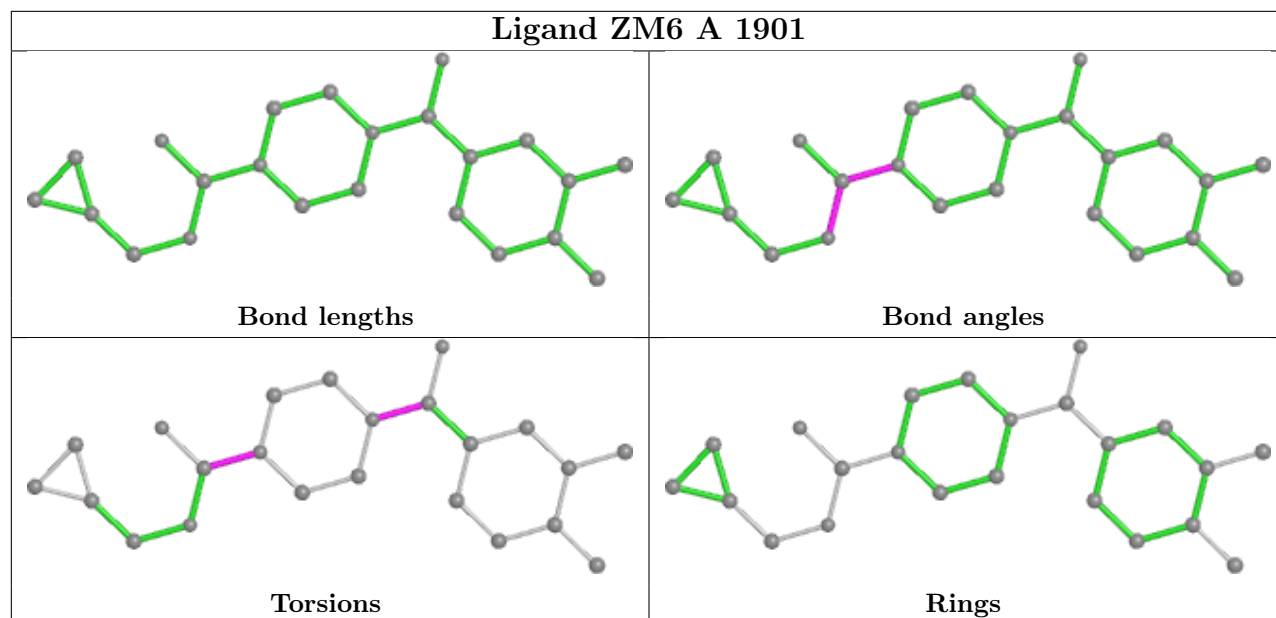
5.7 Other polymers ⓘ

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	121/149 (81%)	0.88	16 (13%)	3	4	11, 14, 27, 32	5 (4%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1382[A]	CYS	9.8
1	A	1315	SER	7.6
1	A	1435	ASN	6.6
1	A	1405[A]	SER	5.8
1	A	1427[A]	SER	5.1
1	A	1423[A]	SER	4.9
1	A	1410	LEU	4.5
1	A	1335	CYS	4.4
1	A	1419[A]	SER	4.3
1	A	1348	LEU	4.2
1	A	1347	LEU	3.1
1	A	1413	PHE	2.7
1	A	1417	HIS	2.5
1	A	1400	ARG	2.1
1	A	1331	LEU	2.1
1	A	1333	PHE	2.1

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

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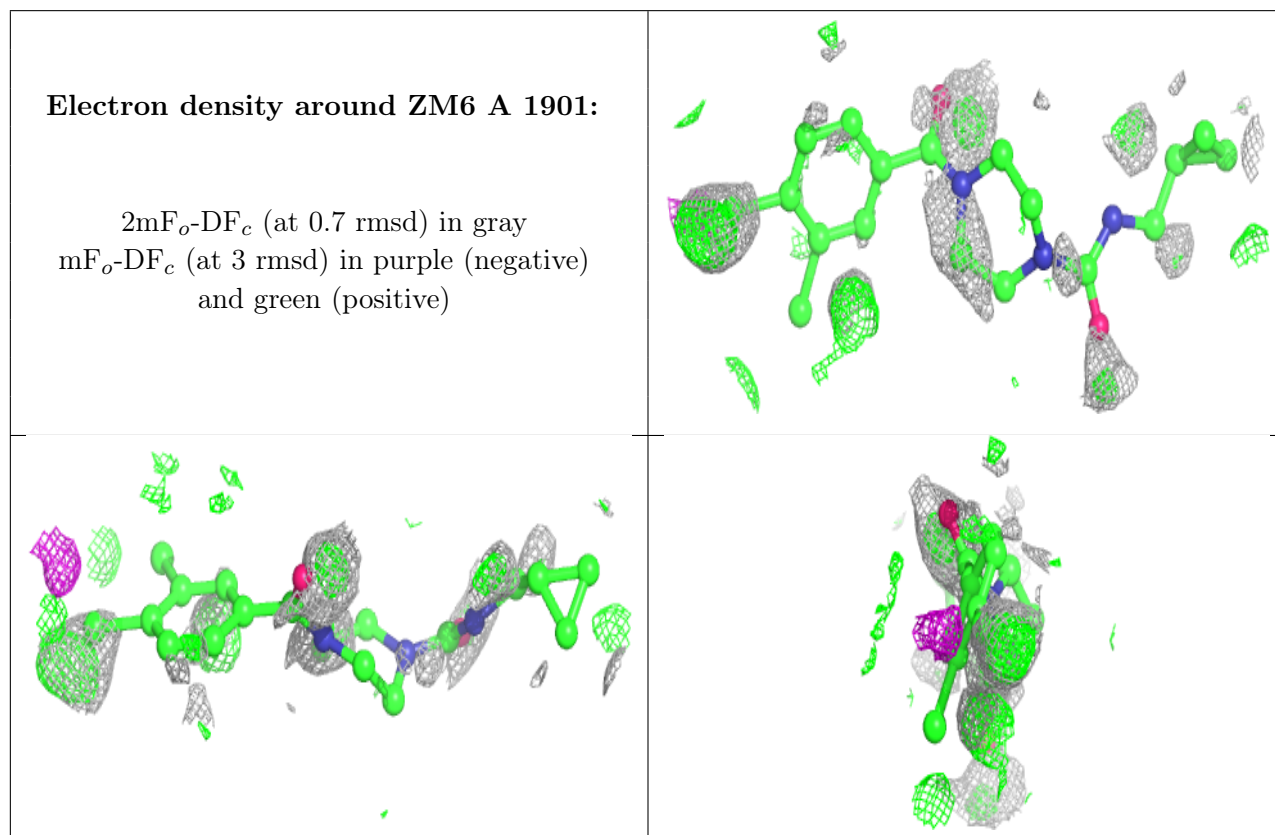
6.3   Carbohydrates ⓘ

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(Å^2)	Q<0.9
2	ZM6	A	1901	23/23	0.20	0.69	59,64,73,74	23

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 ZM6 A 1901:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.