



wwPDB X-ray Structure Validation Summary Report ⓘ

May 6, 2025 – 12:40 AM EDT

PDB ID : 9N9Y / pdb_00009n9y
Title : Crystal structure of truncated USP1:UAF1 in complex with compound 18
Authors : Whittington, D.A.
Deposited on : 2025-02-11
Resolution : 3.15 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

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 Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.43.1

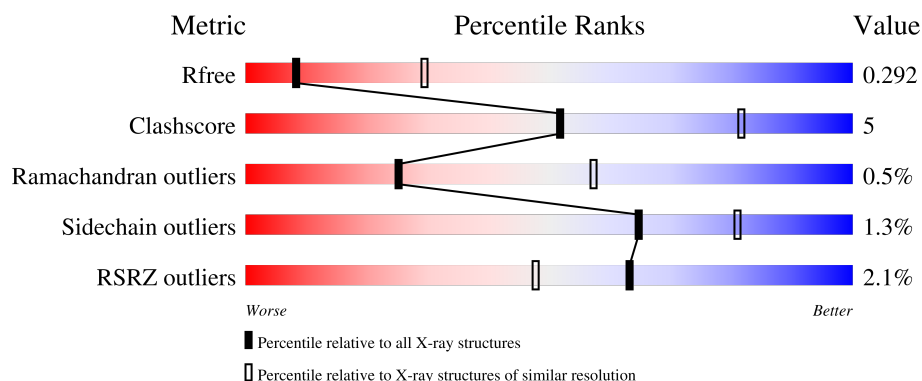
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.15 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}	164625	2168 (3.20-3.12)
Clashscore	180529	2333 (3.20-3.12)
Ramachandran outliers	177936	2266 (3.20-3.12)
Sidechain outliers	177891	2265 (3.20-3.12)
RSRZ outliers	164620	2169 (3.20-3.12)

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 $\leq 5\%$. 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	569	 2% 77% 14% 9%
2	B	381	 % 71% 14% 15%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6739 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 WD repeat-containing protein 48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	519	Total	C	N	O	S	0	0	0
			4080	2576	716	766	22			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	564	GLU	-	expression tag	UNP Q8TAF3
A	565	ASN	-	expression tag	UNP Q8TAF3
A	566	LEU	-	expression tag	UNP Q8TAF3
A	567	TYR	-	expression tag	UNP Q8TAF3
A	568	PHE	-	expression tag	UNP Q8TAF3
A	569	GLN	-	expression tag	UNP Q8TAF3

- Molecule 2 is a protein called Ubiquitin carboxyl-terminal hydrolase 1, N-terminal fragment, Ubiquitin carboxyl-terminal hydrolase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	325	Total	C	N	O	S	0	0	0
			2611	1664	426	506	15			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	84	GLY	-	expression tag	UNP O94782
B	223	GLY	-	linker	UNP O94782
B	224	SER	-	linker	UNP O94782
B	225	GLY	-	linker	UNP O94782
B	414	SER	-	linker	UNP O94782
B	415	GLY	-	linker	UNP O94782
B	416	SER	-	linker	UNP O94782
B	417	GLY	-	linker	UNP O94782
B	418	SER	-	linker	UNP O94782

Continued on next page...

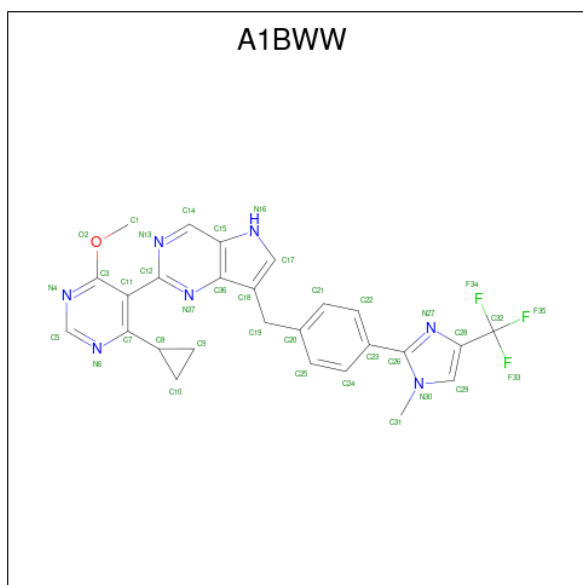
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	419	GLY	-	linker	UNP O94782
B	420	SER	-	linker	UNP O94782

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Zn 1 1	0	0

- Molecule 4 is 2-(4-cyclopropyl-6-methoxypyrimidin-5-yl)-7-({4-[1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl]phenyl}methyl)-5H-pyrrolo[3,2-d]pyrimidine (CCD ID: A1BWW) (formula: C₂₆H₂₂F₃N₇O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C F N O 37 26 3 7 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	7	Total O 7 7	0	0
5	B	3	Total O 3 3	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	123.20Å 161.88Å 160.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.43 – 3.15 49.43 – 3.15	Depositor EDS
% Data completeness (in resolution range)	47.1 (49.43-3.15) 47.1 (49.43-3.15)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.26 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, REFMAC 5.8.0267	Depositor
R, R_{free}	0.237 , 0.291 0.238 , 0.292	Depositor DCC
R_{free} test set	1423 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	87.0	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 49.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.019 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	6739	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: A1BWW, ZN

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	0.08	0/4163	0.28	0/5653
2	B	0.09	0/2655	0.27	0/3579
All	All	0.09	0/6818	0.28	0/9232

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

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	4080	0	4086	43	0
2	B	2611	0	2596	25	0
3	B	1	0	0	0	0
4	B	37	0	0	0	0
5	A	7	0	0	0	0
5	B	3	0	0	0	0
All	All	6739	0	6682	67	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 5.

The worst 5 of 67 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:519:GLU:HB3	1:A:523:ARG:HB3	1.79	0.64
1:A:259:GLN:HB2	1:A:268:TYR:HB2	1.80	0.63
1:A:213:VAL:HA	1:A:229:SER:HA	1.81	0.63
2:B:495:VAL:HG22	2:B:518:SER:HB3	1.82	0.62
1:A:297:VAL:HA	1:A:316:THR:HG22	1.85	0.58

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	513/569 (90%)	480 (94%)	30 (6%)	3 (1%)	22	54
2	B	311/381 (82%)	290 (93%)	20 (6%)	1 (0%)	37	66
All	All	824/950 (87%)	770 (93%)	50 (6%)	4 (0%)	25	57

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	345	PRO
1	A	159	VAL
1	A	344	ASN
2	B	739	SER

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	455/498 (91%)	454 (100%)	1 (0%)	92	96
2	B	300/345 (87%)	291 (97%)	9 (3%)	36	63
All	All	755/843 (90%)	745 (99%)	10 (1%)	65	81

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	517	ARG
2	B	600	VAL
2	B	741	LYS
2	B	186	ARG
2	B	206	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	383	ASN
2	B	454	GLN
2	B	97	GLN
2	B	511	HIS
2	B	201	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	A1BWW	B	802	-	37,42,42	0.91	3 (8%)	47,63,63	2.04	11 (23%)

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
4	A1BWW	B	802	-	-	5/24/26/26	0/6/6/6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	802	A1BWW	F34-C32	2.04	1.40	1.33
4	B	802	A1BWW	F33-C32	2.04	1.40	1.33
4	B	802	A1BWW	F35-C32	2.02	1.40	1.33

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	802	A1BWW	C32-C28-N27	5.27	125.92	119.72
4	B	802	A1BWW	C14-N13-C12	5.14	121.46	115.94
4	B	802	A1BWW	C15-C36-N37	-4.71	119.18	125.12
4	B	802	A1BWW	N6-C5-N4	-4.36	121.98	128.58
4	B	802	A1BWW	C11-C12-N13	3.65	120.03	116.95

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	802	A1BWW	C24-C23-C26-N27
4	B	802	A1BWW	C22-C23-C26-N27
4	B	802	A1BWW	C22-C23-C26-N30

Continued on next page...

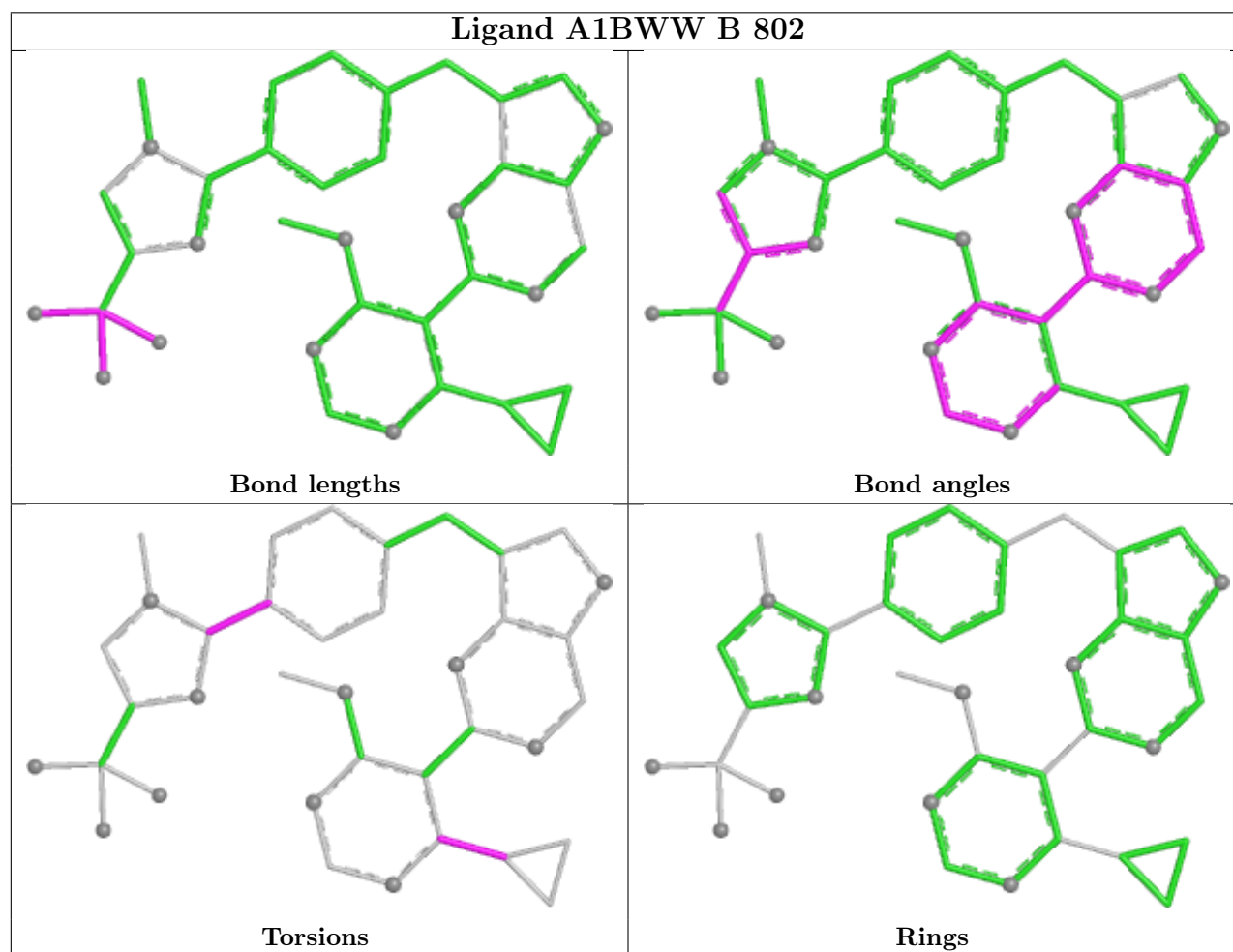
Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	B	802	A1BWW	C11-C7-C8-C9
4	B	802	A1BWW	C24-C23-C26-N30

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 [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, 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(Å ²)	Q<0.9
1	A	519/569 (91%)	0.02	14 (2%) 56 40	55, 76, 110, 138	0
2	B	325/381 (85%)	0.15	4 (1%) 76 60	52, 87, 126, 159	0
All	All	844/950 (88%)	0.07	18 (2%) 63 48	52, 80, 119, 159	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	44	ARG	3.9
1	A	159	VAL	3.7
1	A	162	SER	3.6
1	A	158	THR	3.4
1	A	161	THR	3.3

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](#)

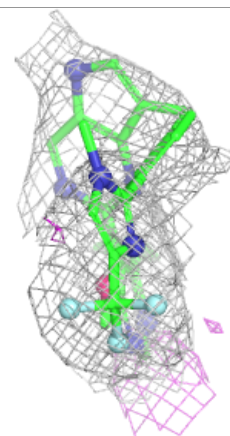
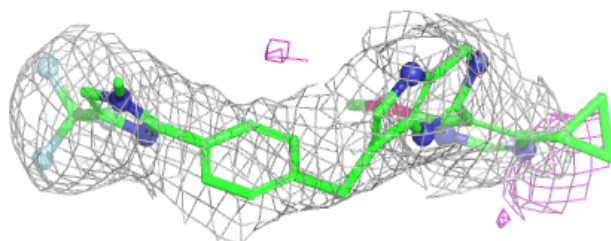
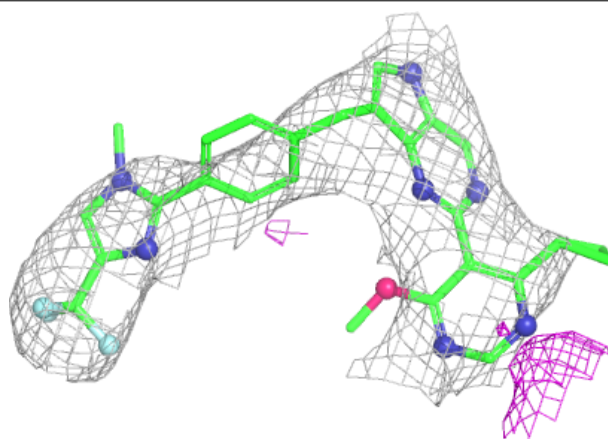
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	A1BWW	B	802	37/37	0.94	0.11	66,80,94,99	0
3	ZN	B	801	1/1	1.00	0.04	70,70,70,70	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 A1BWW B 802:

2mF_o-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers [i](#)

There are no such residues in this entry.