



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 29, 2024 – 01:55 pm GMT

PDB ID : 9FIS  
Title : Structure-guided discovery of selective USP7 inhibitors with in vivo activity  
Authors : Baker, L.M.; Murray, J.; Hubbard, R.E.; Whitehead, N.  
Deposited on : 2024-05-29  
Resolution : 2.77 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 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

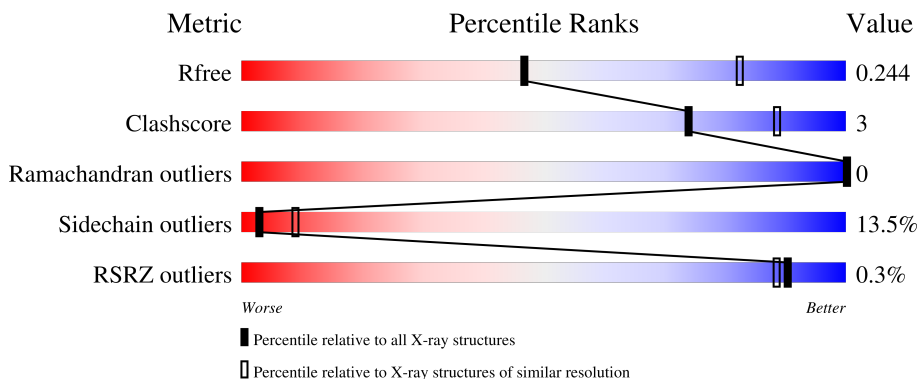
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.77 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	4924 (2.80-2.76)
Clashscore	180529	5458 (2.80-2.76)
Ramachandran outliers	177936	5386 (2.80-2.76)
Sidechain outliers	177891	5388 (2.80-2.76)
RSRZ outliers	164620	4926 (2.80-2.76)

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	360	 77% 14% • 7%
1	B	360	 74% 16% • 8%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 10952 atoms, of which 5388 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.

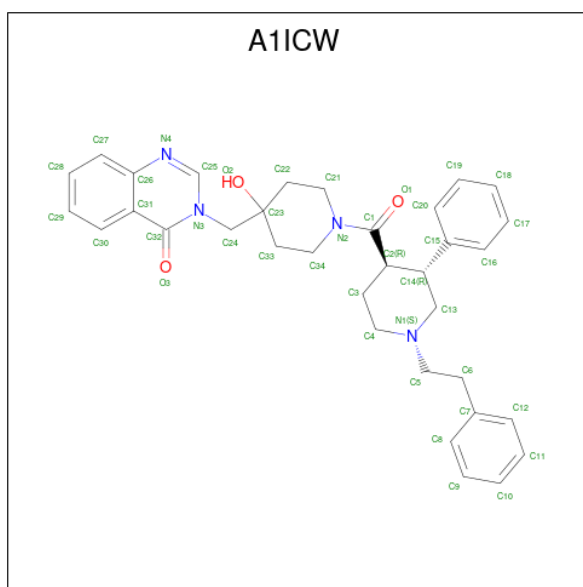
- Molecule 1 is a protein called Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	335	5384	1719	2668	459	522	16	75	0	0
1	B	333	5343	1708	2644	457	518	16	77	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	MET	-	initiating methionine	UNP Q93009
A	561	HIS	-	expression tag	UNP Q93009
A	562	HIS	-	expression tag	UNP Q93009
A	563	HIS	-	expression tag	UNP Q93009
A	564	HIS	-	expression tag	UNP Q93009
A	565	HIS	-	expression tag	UNP Q93009
A	566	HIS	-	expression tag	UNP Q93009
B	207	MET	-	initiating methionine	UNP Q93009
B	561	HIS	-	expression tag	UNP Q93009
B	562	HIS	-	expression tag	UNP Q93009
B	563	HIS	-	expression tag	UNP Q93009
B	564	HIS	-	expression tag	UNP Q93009
B	565	HIS	-	expression tag	UNP Q93009
B	566	HIS	-	expression tag	UNP Q93009

- Molecule 2 is 3-[[4-oxidanyl-1-[(3 {R},4 {R})-3-phenyl-1-(2-phenylethyl)piperidin-4-yl]carbonyl-piperidin-4-yl]methyl]quinazolin-4-one (three-letter code: A1ICW) (formula: C<sub>34</sub>H<sub>38</sub>N<sub>4</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	H	N			O
2	A	1	79	34	38	4	3	1	0
2	B	1	79	34	38	4	3	1	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	28	Total O 28 28	0	0
3	B	39	Total O 39 39	0	0



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.84Å 70.58Å 78.59Å 90.00° 92.96° 90.00°	Depositor
Resolution (Å)	52.78 – 2.77 52.78 – 2.77	Depositor EDS
% Data completeness (in resolution range)	97.6 (52.78-2.77) 97.6 (52.78-2.77)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.185 , 0.243 0.186 , 0.244	Depositor DCC
$R_{free}$ test set	1052 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	70.7	Xtriage
Anisotropy	0.190	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for l,k,-h 0.024 for h,-k,-l 0.018 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10952	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.48% 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: A1ICW

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	0.41	0/2772	0.83	1/3739 (0.0%)
1	B	0.41	0/2755	0.86	6/3717 (0.2%)
All	All	0.41	0/5527	0.85	7/7456 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	343	ARG	NE-CZ-NH1	-8.80	115.90	120.30
1	B	454	LEU	CB-CG-CD2	7.95	124.52	111.00
1	B	225	MET	CG-SD-CE	-6.49	89.81	100.20
1	B	343	ARG	NE-CZ-NH2	6.22	123.41	120.30
1	B	454	LEU	CB-CG-CD1	-5.82	101.11	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	344	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	B	343	ARG	Sidechain
1	B	549	ARG	Sidechain

## 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	2716	2668	2656	16	0
1	B	2699	2644	2630	19	0
2	A	41	38	0	0	0
2	B	41	38	0	0	0
3	A	28	0	0	0	0
3	B	39	0	0	1	0
All	All	5564	5388	5286	32	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 3.

The worst 5 of 32 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:B:455:VAL:HG13	1:B:511:THR:HB	1.58	0.84
1:A:227:SER:HB3	1:A:454:LEU:HD13	1.80	0.64
1:A:320:ILE:HB	1:A:321:PRO:HD3	1.79	0.63
1:A:553:GLN:O	1:A:554:LYS:HB3	2.02	0.59
1:A:256:VAL:HB	1:A:257:PRO:HD3	1.86	0.57

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	331/360 (92%)	314 (95%)	17 (5%)	0	100	100
1	B	327/360 (91%)	315 (96%)	12 (4%)	0	100	100
All	All	658/720 (91%)	629 (96%)	29 (4%)	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	302/327 (92%)	265 (88%)	37 (12%)	4	11
1	B	300/327 (92%)	256 (85%)	44 (15%)	2	7
All	All	602/654 (92%)	521 (86%)	81 (14%)	3	9

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	370	VAL
1	B	459	ASP
1	B	383	GLU
1	B	419	ILE
1	B	487	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	218	ASN
1	A	384	HIS
1	A	512	ASN
1	B	237	GLN
1	B	418	ASN

### 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	A1ICW	B	601	-	46,46,46	0.53	0	58,65,65	1.04	5 (8%)
2	A1ICW	A	601	-	46,46,46	0.48	0	58,65,65	0.80	2 (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.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	A1ICW	B	601	-	-	4/22/47/47	0/6/6/6
2	A1ICW	A	601	-	-	4/22/47/47	0/6/6/6

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	A1ICW	O1-C1-N2	-3.95	117.04	121.67
2	B	601	A1ICW	C14-C13-N1	3.21	115.65	110.63
2	A	601	A1ICW	C14-C13-N1	-2.85	106.17	110.63
2	B	601	A1ICW	C5-N1-C13	-2.75	105.20	111.66
2	A	601	A1ICW	C31-C32-N3	2.68	115.36	113.80

There are no chirality outliers.

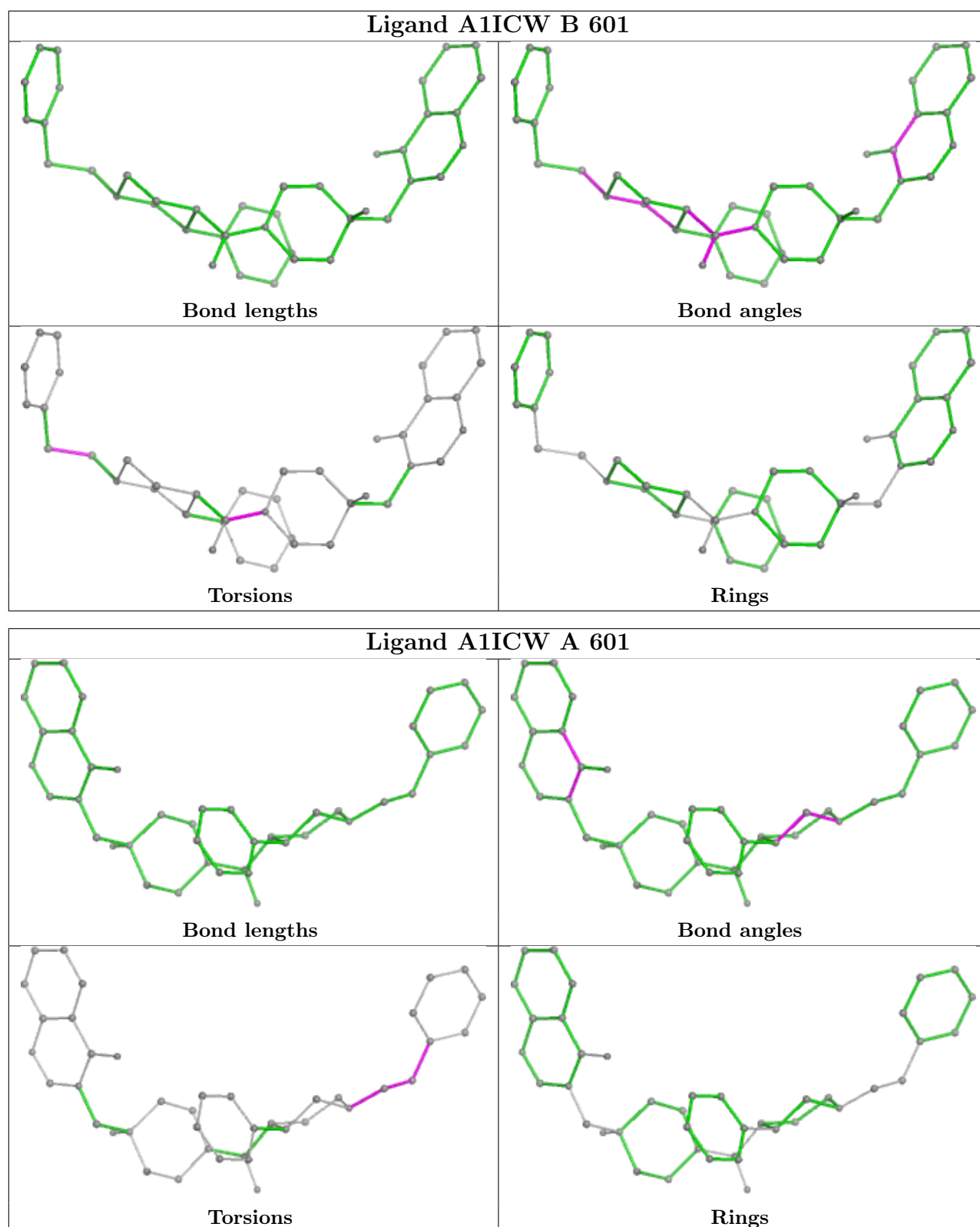
5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	601	A1ICW	O1-C1-N2-C34
2	A	601	A1ICW	C6-C5-N1-C4
2	A	601	A1ICW	N1-C5-C6-C7
2	B	601	A1ICW	N1-C5-C6-C7
2	B	601	A1ICW	O1-C1-N2-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 [\(i\)](#)

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	335/360 (93%)	-0.69	1 (0%) 90 88	39, 64, 110, 189	0
1	B	333/360 (92%)	-0.71	1 (0%) 90 88	35, 63, 108, 141	0
All	All	668/720 (92%)	-0.70	2 (0%) 90 88	35, 64, 109, 189	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	500	GLY	2.9
1	B	500	GLY	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](#)

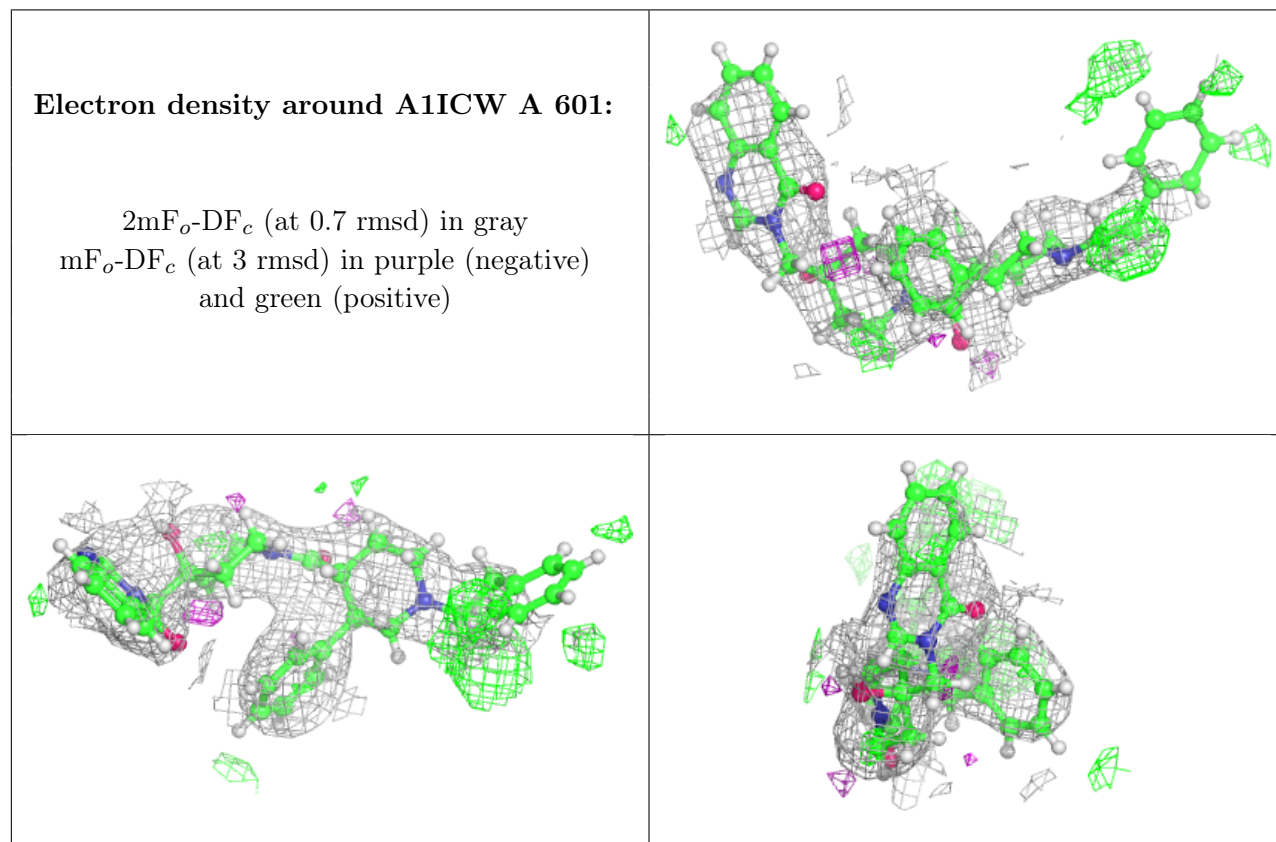
There are no monosaccharides in this entry.

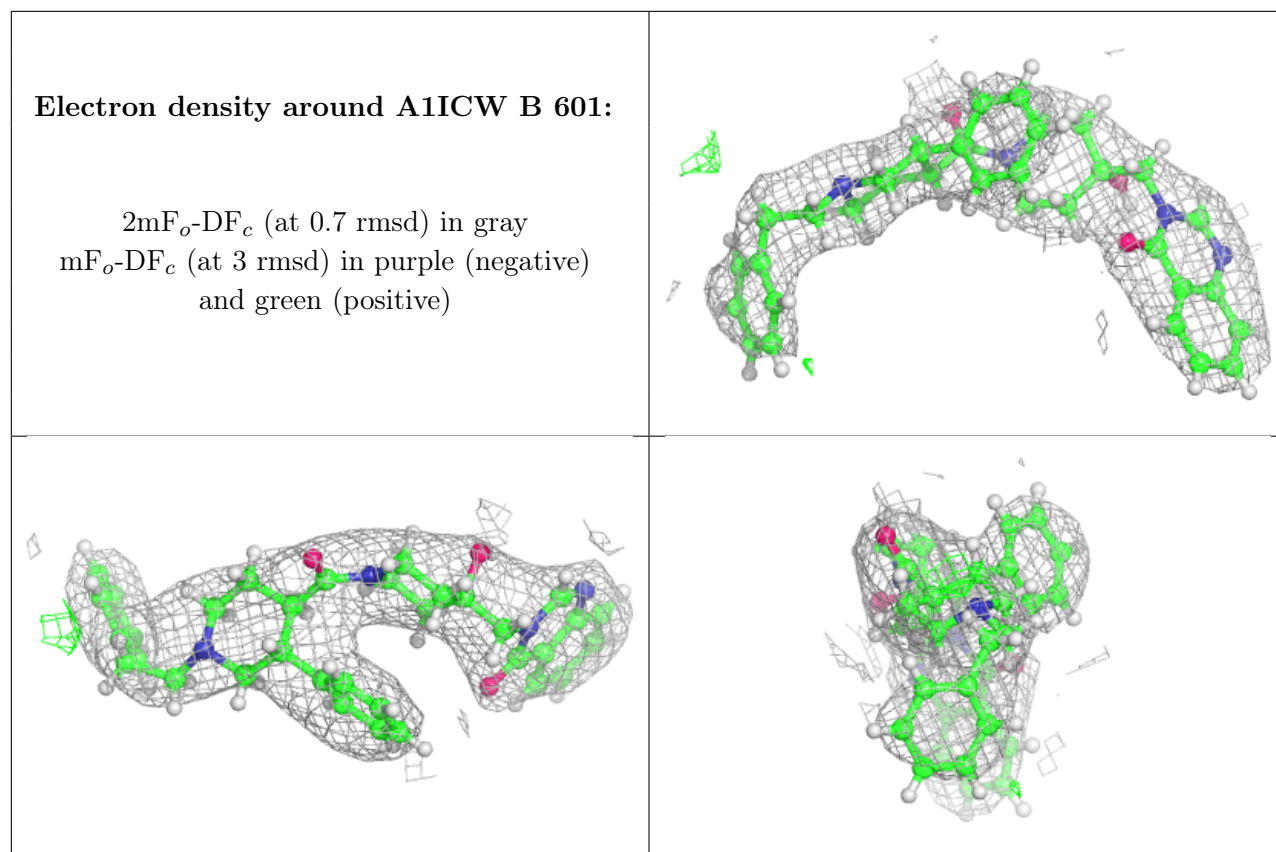
### 6.4 Ligands [i](#)

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<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(Å <sup>2</sup> )	Q<0.9
2	A1ICW	A	601	41/41	0.86	0.16	60,113,184,191	1
2	A1ICW	B	601	41/41	0.94	0.09	30,61,80,83	1

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.