WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Dec 1, 2021 – 02:12 pm GMT

PDB ID	:	7OSX
Title	:	PARP15 catalytic domain in complex with OUL205
Authors	:	Maksimainen, M.M.; Lehtio, L.
Deposited on	:	2021-06-09
Resolution	:	1.60 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 ⓘ symbol.

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

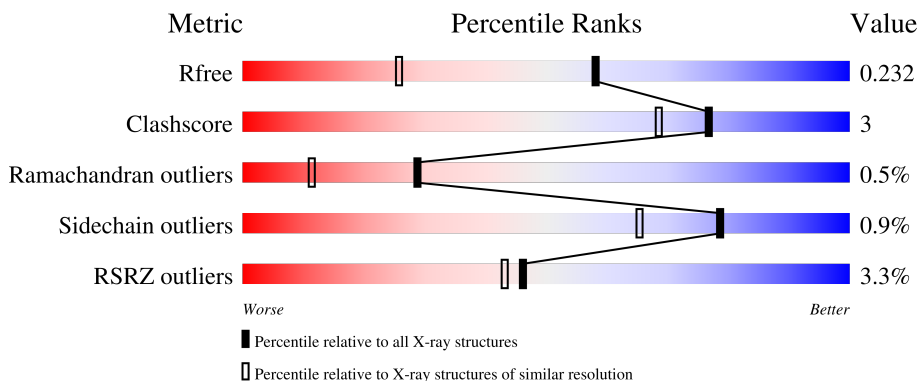|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.60 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		0.232
Clashscore		3
Ramachandran outliers		0.5%
Sidechain outliers		0.9%
RSRZ outliers		3.3%

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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	221	4% / 77% / 9% / 13%
1	B	221	2% / 86% / · / 10%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3468 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 Protein mono-ADP-ribosyltransferase PARP15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total 1611	C 1024	N 280	O 299	S 8	0	8	0
1	B	198	Total 1623	C 1029	N 285	O 301	S 8	0	4	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	458	MET	-	initiating methionine	UNP Q460N3
A	459	HIS	-	expression tag	UNP Q460N3
A	460	HIS	-	expression tag	UNP Q460N3
A	461	HIS	-	expression tag	UNP Q460N3
A	462	HIS	-	expression tag	UNP Q460N3
A	463	HIS	-	expression tag	UNP Q460N3
A	464	HIS	-	expression tag	UNP Q460N3
A	465	SER	-	expression tag	UNP Q460N3
A	466	SER	-	expression tag	UNP Q460N3
A	467	GLY	-	expression tag	UNP Q460N3
A	468	VAL	-	expression tag	UNP Q460N3
A	469	ASP	-	expression tag	UNP Q460N3
A	470	LEU	-	expression tag	UNP Q460N3
A	471	GLY	-	expression tag	UNP Q460N3
A	472	THR	-	expression tag	UNP Q460N3
A	473	GLU	-	expression tag	UNP Q460N3
A	474	ASN	-	expression tag	UNP Q460N3
A	475	LEU	-	expression tag	UNP Q460N3
A	476	TYR	-	expression tag	UNP Q460N3
A	477	PHE	-	expression tag	UNP Q460N3
A	478	GLN	-	expression tag	UNP Q460N3
A	479	SER	-	expression tag	UNP Q460N3
A	480	MET	-	expression tag	UNP Q460N3
B	458	MET	-	initiating methionine	UNP Q460N3
B	459	HIS	-	expression tag	UNP Q460N3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	460	HIS	-	expression tag	UNP Q460N3
B	461	HIS	-	expression tag	UNP Q460N3
B	462	HIS	-	expression tag	UNP Q460N3
B	463	HIS	-	expression tag	UNP Q460N3
B	464	HIS	-	expression tag	UNP Q460N3
B	465	SER	-	expression tag	UNP Q460N3
B	466	SER	-	expression tag	UNP Q460N3
B	467	GLY	-	expression tag	UNP Q460N3
B	468	VAL	-	expression tag	UNP Q460N3
B	469	ASP	-	expression tag	UNP Q460N3
B	470	LEU	-	expression tag	UNP Q460N3
B	471	GLY	-	expression tag	UNP Q460N3
B	472	THR	-	expression tag	UNP Q460N3
B	473	GLU	-	expression tag	UNP Q460N3
B	474	ASN	-	expression tag	UNP Q460N3
B	475	LEU	-	expression tag	UNP Q460N3
B	476	TYR	-	expression tag	UNP Q460N3
B	477	PHE	-	expression tag	UNP Q460N3
B	478	GLN	-	expression tag	UNP Q460N3
B	479	SER	-	expression tag	UNP Q460N3
B	480	MET	-	expression tag	UNP Q460N3

- Molecule 2 is 8-oxidanyl-4 {H}-thieno[2,3-c]isoquinolin-5-one (three-letter code: 0UI) (formula: $C_{11}H_7NO_2S$) (labeled as "Ligand of Interest" by depositor).
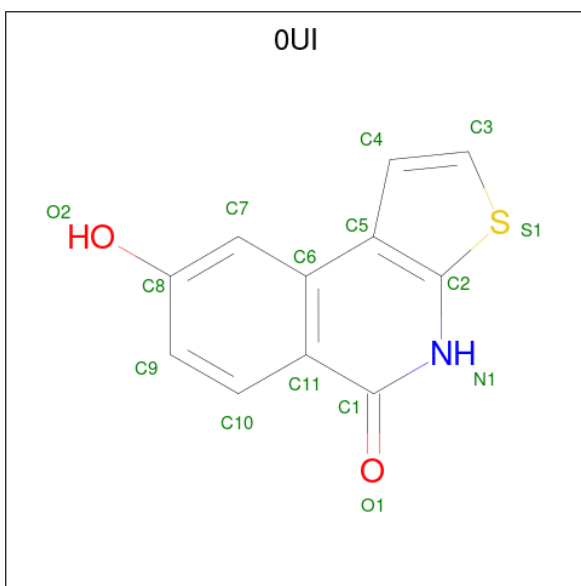
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			15	11	1	2	1		

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	96	Total	O	0	0
			96	96		
4	B	118	Total	O	0	1
			119	119		

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: Protein mono-ADP-ribosyltransferase PARP15

Chain A:
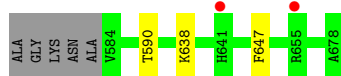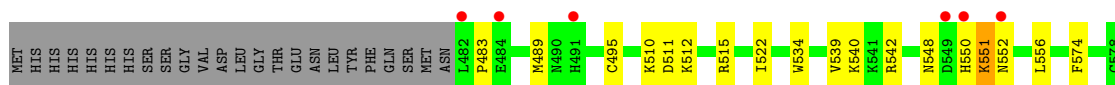
• Molecule 1: Protein mono-ADP-ribosyltransferase PARP15

Chain B:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	45.37Å 68.76Å 161.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.71 – 1.60 43.67 – 1.60	Depositor EDS
% Data completeness (in resolution range)	98.3 (43.71-1.60) 98.3 (43.67-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.198 , 0.223 0.207 , 0.232	Depositor DCC
R_{free} test set	3317 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	20.8	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3468	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% 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: 0UI, DMS

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	0.64	0/1678	0.75	0/2270
1	B	0.63	0/1679	0.74	0/2273
All	All	0.63	0/3357	0.75	0/4543

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	A	1611	0	1571	15	0
1	B	1623	0	1574	6	0
2	A	15	0	0	0	0
3	B	4	0	6	0	0
4	A	96	0	0	4	0
4	B	119	0	0	0	0
All	All	3468	0	3151	19	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.

All (19) 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:590[A]:THR:HG22	4:A:857:HOH:O	1.93	0.67
1:A:510:LYS:HE3	4:A:802:HOH:O	2.01	0.60
1:A:510:LYS:CE	4:A:802:HOH:O	2.52	0.58
1:A:574:PHE:HB3	1:A:590[B]:THR:HG21	1.87	0.56
1:A:510:LYS:NZ	4:A:802:HOH:O	2.32	0.54
1:B:653:ASN:HD21	1:B:655:ARG:NE	2.09	0.49
1:B:655:ARG:HG2	1:B:656:SER:N	2.26	0.49
1:A:539:VAL:HG21	1:B:539:VAL:HG21	1.94	0.49
1:A:542[B]:ARG:NH2	1:B:535:GLN:OE1	2.48	0.46
1:B:510:LYS:O	1:B:514:THR:HG23	2.17	0.45
1:A:574:PHE:HB3	1:A:590[A]:THR:HG21	2.00	0.44
1:A:512:LYS:HD3	1:A:556:LEU:HD23	2.00	0.44
1:B:653:ASN:HD21	1:B:655:ARG:CZ	2.30	0.44
1:A:548:ASN:HB3	1:A:551:LYS:HD2	1.99	0.44
1:A:483:PRO:HG3	1:A:534:TRP:CZ2	2.53	0.43
1:A:638:LYS:HA	1:A:647:PHE:CE2	2.54	0.43
1:A:511[A]:ASP:OD1	1:A:515:ARG:NH2	2.47	0.42
1:A:489:MET:HG2	1:A:495[A]:CYS:SG	2.59	0.41
1:A:540[A]:LYS:HA	1:A:540[A]:LYS:HD2	1.88	0.41

There are no symmetry-related clashes.

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	196/221 (89%)	192 (98%)	2 (1%)	2 (1%)	15	3
1	B	200/221 (90%)	197 (98%)	3 (2%)	0	100	100
All	All	396/442 (90%)	389 (98%)	5 (1%)	2 (0%)	29	11

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	550	HIS
1	A	551	LYS

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	182/198 (92%)	180 (99%)	2 (1%)	73	57
1	B	180/198 (91%)	179 (99%)	1 (1%)	86	77
All	All	362/396 (91%)	359 (99%)	3 (1%)	78	70

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	522	ILE
1	A	552	ASN
1	B	581	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	571	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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 0UI A 701
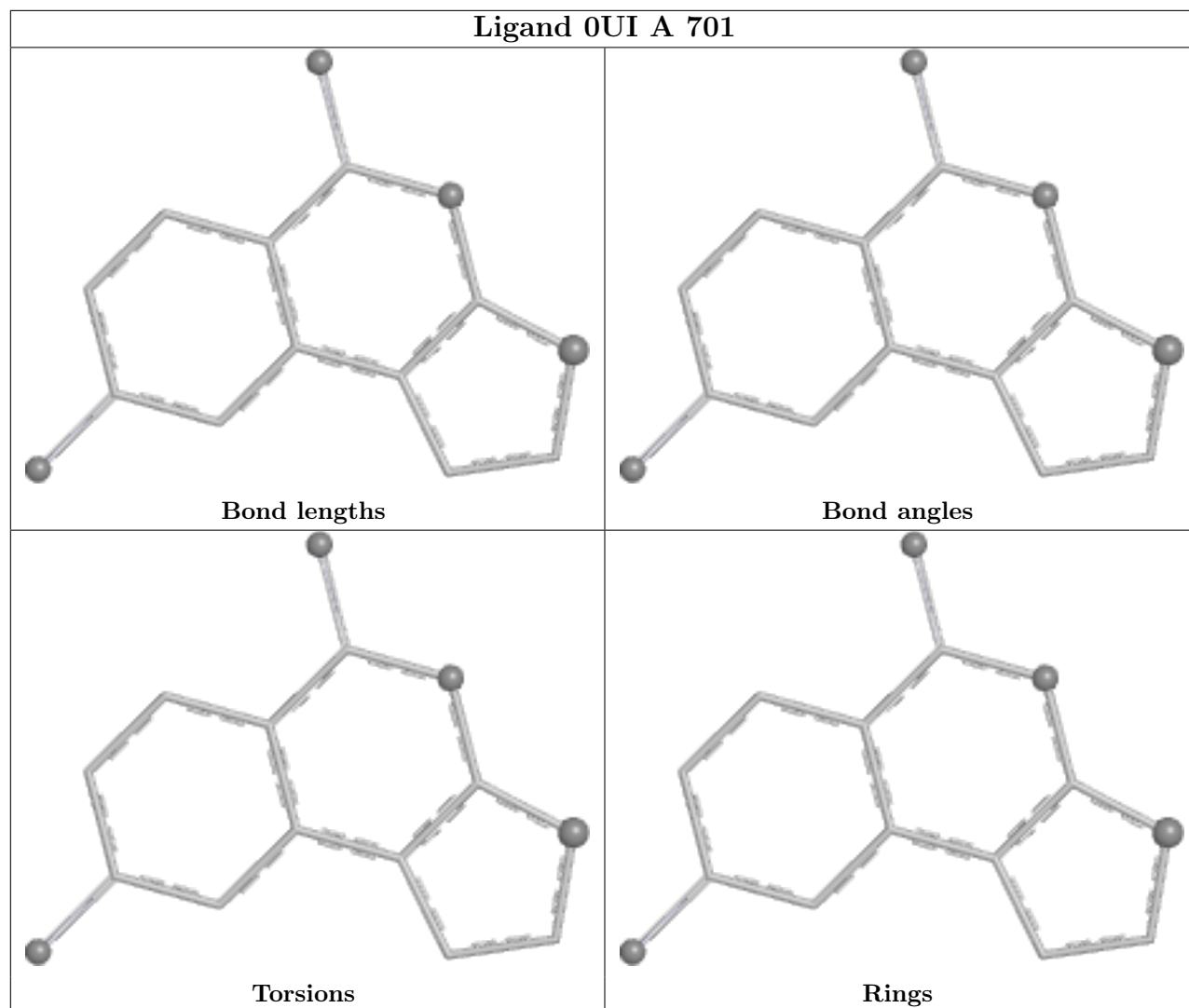
5.7 Other polymers ⓘ

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 (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	192/221 (86%)	0.11	8 (4%)	36	33	16, 23, 49, 72	0
1	B	198/221 (89%)	-0.08	5 (2%)	57	55	14, 21, 38, 67	0
All	All	390/442 (88%)	0.01	13 (3%)	46	43	14, 22, 46, 72	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	550	HIS	7.9
1	B	491	HIS	5.7
1	B	490	ASN	4.3
1	A	552	ASN	3.4
1	A	482	LEU	3.0
1	A	484	GLU	2.8
1	A	491	HIS	2.8
1	B	549	ASP	2.6
1	B	655	ARG	2.5
1	A	641	HIS	2.5
1	A	549	ASP	2.3
1	B	492	GLN	2.3
1	A	655	ARG	2.1

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

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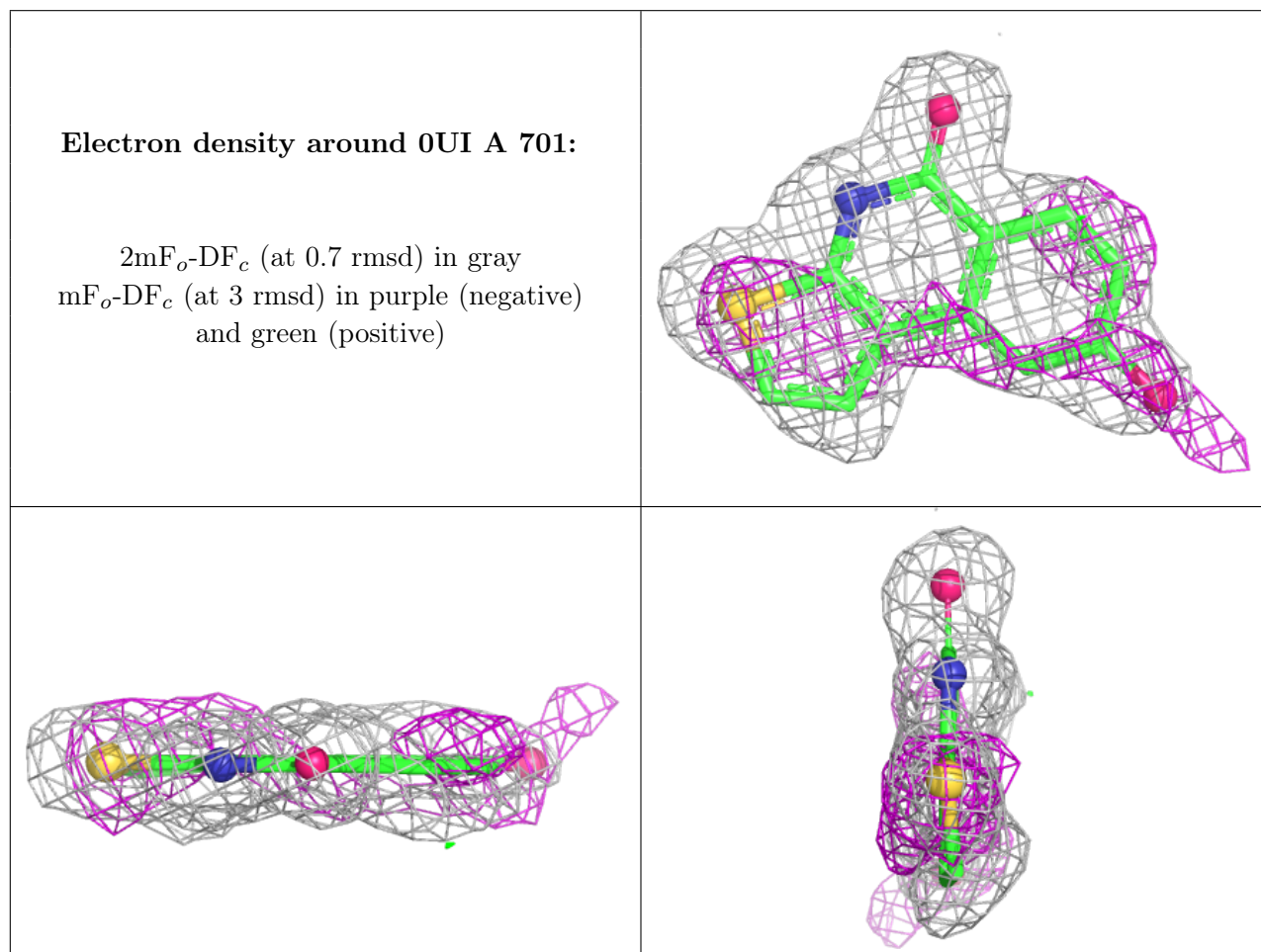
6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4　Ligands ⓘ

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
3	DMS	B	701	4/4	0.84	0.29	23,41,44,46	0
2	0UI	A	701	15/15	0.86	0.24	20,30,36,39	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.

Electron density around 0UI A 701:

$2mF_o$-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 ⓘ

There are no such residues in this entry.