W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 05:03 pm GMT

PDB ID	:	6YDV
Title	:	Crystal Structure of the Jmjc Domain of Human JMJD1B in complex with FM001511a from the DSPL fragment library
Authors	:	Snee, M.; Nowak, R.; Johansson, C.; Burgess-Brown, N.A.; Arrowsmith, C.H.; Bountra, C.; Edwards, A.M.; Oppermann, U.
Deposited on	:	2020-03-21
Resolution	:	2.07 Å(reported)

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

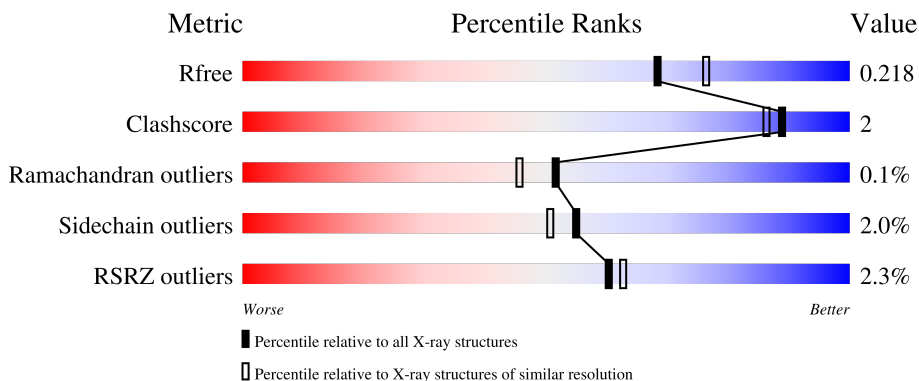|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.07 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		0.218
Clashscore		2
Ramachandran outliers		0.1%
Sidechain outliers		2.0%
RSRZ outliers		2.3%

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	2684 (2.08-2.04)
Clashscore	141614	2801 (2.08-2.04)
Ramachandran outliers	138981	2768 (2.08-2.04)
Sidechain outliers	138945	2768 (2.08-2.04)
RSRZ outliers	127900	2646 (2.08-2.04)

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	AAA	372	% 85% 6% · 9%
1	BBB	372	3% 87% 5% 8%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 11507 atoms, of which 5350 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 JMJD1B protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	340	Total 5367	C 1729	H 2639	N 474	O 510	S 15	124	0	0
1	BBB	344	Total 5477	C 1763	H 2699	N 487	O 513	S 15	132	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1357	MET	-	initiating methionine	UNP A4FUT8
AAA	1358	HIS	-	expression tag	UNP A4FUT8
AAA	1359	HIS	-	expression tag	UNP A4FUT8
AAA	1360	HIS	-	expression tag	UNP A4FUT8
AAA	1361	HIS	-	expression tag	UNP A4FUT8
AAA	1362	HIS	-	expression tag	UNP A4FUT8
AAA	1363	HIS	-	expression tag	UNP A4FUT8
AAA	1364	SER	-	expression tag	UNP A4FUT8
AAA	1365	SER	-	expression tag	UNP A4FUT8
AAA	1366	GLY	-	expression tag	UNP A4FUT8
AAA	1367	VAL	-	expression tag	UNP A4FUT8
AAA	1368	ASP	-	expression tag	UNP A4FUT8
AAA	1369	LEU	-	expression tag	UNP A4FUT8
AAA	1370	GLY	-	expression tag	UNP A4FUT8
AAA	1371	THR	-	expression tag	UNP A4FUT8
AAA	1372	GLU	-	expression tag	UNP A4FUT8
AAA	1373	ASN	-	expression tag	UNP A4FUT8
AAA	1374	LEU	-	expression tag	UNP A4FUT8
AAA	1375	TYR	-	expression tag	UNP A4FUT8
AAA	1376	PHE	-	expression tag	UNP A4FUT8
AAA	1377	GLN	-	expression tag	UNP A4FUT8
AAA	1378	SER	-	expression tag	UNP A4FUT8
AAA	1379	MET	-	expression tag	UNP A4FUT8
AAA	1601	GLU	GLN	conflict	UNP A4FUT8
AAA	1606	HIS	GLY	conflict	UNP A4FUT8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	1357	MET	-	initiating methionine	UNP A4FUT8
BBB	1358	HIS	-	expression tag	UNP A4FUT8
BBB	1359	HIS	-	expression tag	UNP A4FUT8
BBB	1360	HIS	-	expression tag	UNP A4FUT8
BBB	1361	HIS	-	expression tag	UNP A4FUT8
BBB	1362	HIS	-	expression tag	UNP A4FUT8
BBB	1363	HIS	-	expression tag	UNP A4FUT8
BBB	1364	SER	-	expression tag	UNP A4FUT8
BBB	1365	SER	-	expression tag	UNP A4FUT8
BBB	1366	GLY	-	expression tag	UNP A4FUT8
BBB	1367	VAL	-	expression tag	UNP A4FUT8
BBB	1368	ASP	-	expression tag	UNP A4FUT8
BBB	1369	LEU	-	expression tag	UNP A4FUT8
BBB	1370	GLY	-	expression tag	UNP A4FUT8
BBB	1371	THR	-	expression tag	UNP A4FUT8
BBB	1372	GLU	-	expression tag	UNP A4FUT8
BBB	1373	ASN	-	expression tag	UNP A4FUT8
BBB	1374	LEU	-	expression tag	UNP A4FUT8
BBB	1375	TYR	-	expression tag	UNP A4FUT8
BBB	1376	PHE	-	expression tag	UNP A4FUT8
BBB	1377	GLN	-	expression tag	UNP A4FUT8
BBB	1378	SER	-	expression tag	UNP A4FUT8
BBB	1379	MET	-	expression tag	UNP A4FUT8
BBB	1601	GLU	GLN	conflict	UNP A4FUT8
BBB	1606	HIS	GLY	conflict	UNP A4FUT8

- Molecule 2 is {N}-(2-azanyl-2-oxidanylidene-ethyl)-4-methoxy-benzamide (three-letter code: ONQ) (formula: $C_{10}H_{12}N_2O_3$) (labeled as "Ligand of Interest" by depositor).

ONQ
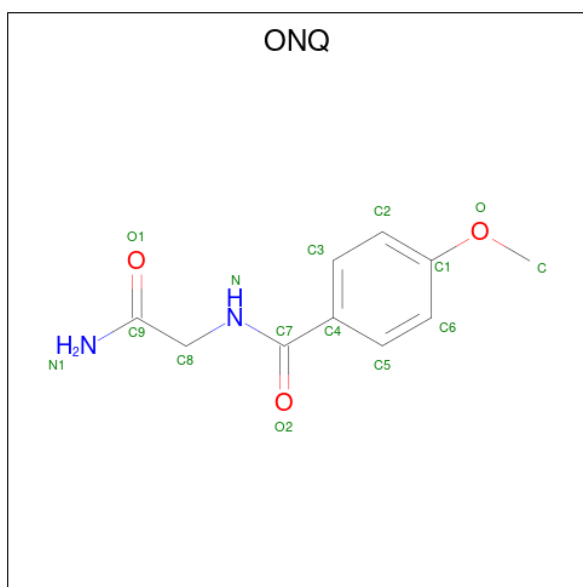
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	2	0
			27	10	12	2	3		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	4	Total	Cl	0	0
			4	4		
3	BBB	1	Total	Cl	0	0
			1	1		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	1	Total	Mn	0	0
			1	1		
4	BBB	1	Total	Mn	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	336	Total	O	0	0
			336	336		
5	BBB	293	Total	O	0	0
			293	293		

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: JMJD1B protein

Chain AAA:

- Molecule 1: JMJD1B protein

Chain BBB:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.60Å 93.79Å 92.61Å 90.00° 107.50° 90.00°	Depositor
Resolution (Å)	54.93 – 2.07 54.93 – 2.07	Depositor EDS
% Data completeness (in resolution range)	99.1 (54.93-2.07) 99.1 (54.93-2.07)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.70 (at 2.07Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.170 , 0.214 0.179 , 0.218	Depositor DCC
R_{free} test set	2774 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	31.3	Xtriage
Anisotropy	0.619	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 51.1	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.036 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	11507	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ONQ, CL, MN

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	AAA	0.67	0/2801	0.75	0/3803
1	BBB	0.68	0/2852	0.72	0/3869
All	All	0.67	0/5653	0.74	0/7672

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	AAA	2728	2639	2610	12	0
1	BBB	2778	2699	2668	9	0
2	AAA	15	12	0	1	0
3	AAA	4	0	0	0	0
3	BBB	1	0	0	0	0
4	AAA	1	0	0	0	0
4	BBB	1	0	0	0	0
5	AAA	336	0	0	2	1
5	BBB	293	0	0	1	1
All	All	6157	5350	5278	21	1

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 2.

The worst 5 of 21 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:AAA:1582:HIS:CD2	1:AAA:1585:GLU:HG3	2.29	0.67
1:BBB:1644:HIS:NE2	1:BBB:1649:ASP:OD2	2.33	0.62
1:AAA:1389:ARG:HG3	2:AAA:1801:ONQ:O2	2.02	0.58
1:BBB:1443:VAL:HG21	1:BBB:1485:LEU:HD11	1.87	0.55
1:BBB:1505:LEU:HD22	1:BBB:1700:VAL:HG11	1.87	0.55

All (1) 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)
5:AAA:2184:HOH:O	5:BBB:2176:HOH:O[2_657]	0.88	1.32

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	AAA	338/372 (91%)	328 (97%)	9 (3%)	1 (0%)	41	32
1	BBB	340/372 (91%)	323 (95%)	17 (5%)	0	100	100
All	All	678/744 (91%)	651 (96%)	26 (4%)	1 (0%)	51	45

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	1379	MET

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	AAA	294/331 (89%)	286 (97%)	8 (3%)	44	39
1	BBB	300/331 (91%)	296 (99%)	4 (1%)	69	67
All	All	594/662 (90%)	582 (98%)	12 (2%)	55	51

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	1717	THR
1	BBB	1582	HIS
1	BBB	1607	LYS
1	BBB	1588	LYS
1	AAA	1410	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 7 are 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	$\#|Z| > 2$
2	ONQ	AAA	1801	-	15,15,15	1.55	3 (20%)	19,19,19	1.18	1 (5%)

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	ONQ	AAA	1801	-	-	4/11/11/11	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	1801	ONQ	C2-C1	2.59	1.43	1.38
2	AAA	1801	ONQ	C7-N	2.50	1.39	1.33
2	AAA	1801	ONQ	C9-N1	2.17	1.39	1.32

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	AAA	1801	ONQ	C-O-C1	4.24	126.71	117.51

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	1801	ONQ	C5-C4-C7-N
2	AAA	1801	ONQ	C3-C4-C7-N
2	AAA	1801	ONQ	C5-C4-C7-O2
2	AAA	1801	ONQ	C3-C4-C7-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	1801	ONQ	1	0

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 ONQ AAA 1801

Bond lengths   Bond angles

Torsions   Rings
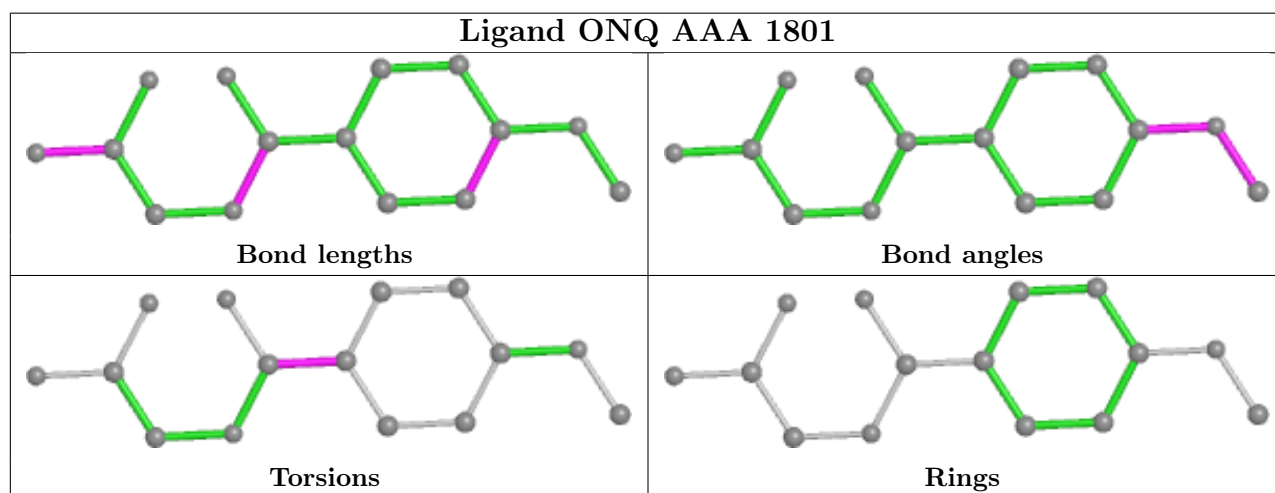
5.7 Other polymers ⓘ

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(Å^2)	Q<0.9
1	AAA	340/372 (91%)	-0.24	5 (1%)	73	75	21, 35, 74, 115	0
1	BBB	344/372 (92%)	-0.10	11 (3%)	47	50	24, 41, 83, 146	0
All	All	684/744 (91%)	-0.17	16 (2%)	60	63	21, 38, 79, 146	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	1577	ILE	5.6
1	AAA	1582	HIS	4.9
1	BBB	1581	ALA	4.6
1	AAA	1581	ALA	3.9
1	BBB	1643	ASP	3.4

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

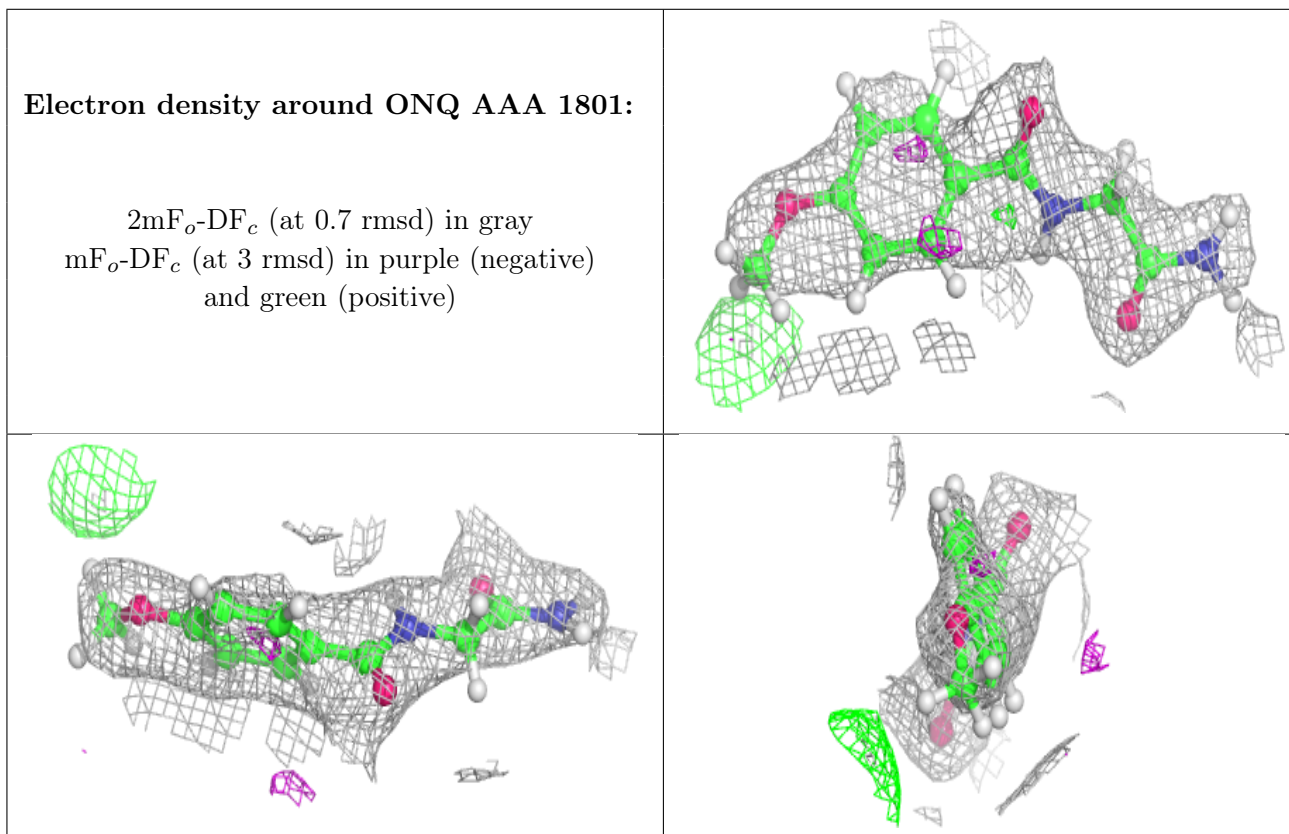There are no non-standard protein/DNA/RNA residues in this entry.

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(Å²)	Q<0.9
2	ONQ	AAA	1801	15/15	0.89	0.17	41,51,61,62	2
4	MN	AAA	1806	1/1	0.96	0.22	63,63,63,63	0
4	MN	BBB	1802	1/1	0.97	0.22	48,48,48,48	0
3	CL	AAA	1805	1/1	0.99	0.11	54,54,54,54	0
3	CL	BBB	1801	1/1	0.99	0.04	44,44,44,44	0
3	CL	AAA	1803	1/1	1.00	0.07	43,43,43,43	0
3	CL	AAA	1804	1/1	1.00	0.07	37,37,37,37	0
3	CL	AAA	1802	1/1	1.00	0.12	28,28,28,28	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 ONQ AAA 1801:

$2mF_o\text{-}DF_c$ (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.