



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 13, 2023 – 03:40 pm GMT

PDB ID : 7Z3Y
Title : Structure of the mouse 8-oxoguanine DNA Glycosylase mOGG1 in complex with ligand TH013545
Authors : Scaletti, E.R.; Stenmark, P.
Deposited on : 2022-03-02
Resolution : 2.35 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 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 : 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

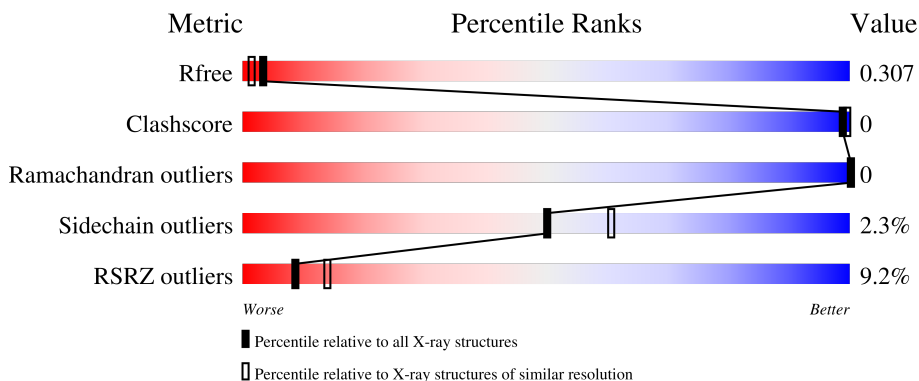
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

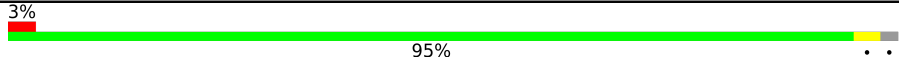
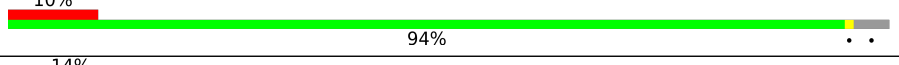
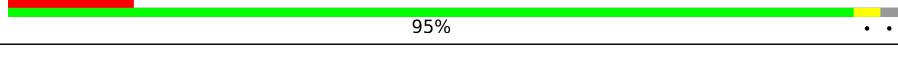
The reported resolution of this entry is 2.35 Å.

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}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	AAA	318	 3% 95%
1	BBB	318	 10% 94%
1	CCC	318	 14% 95%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7295 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.

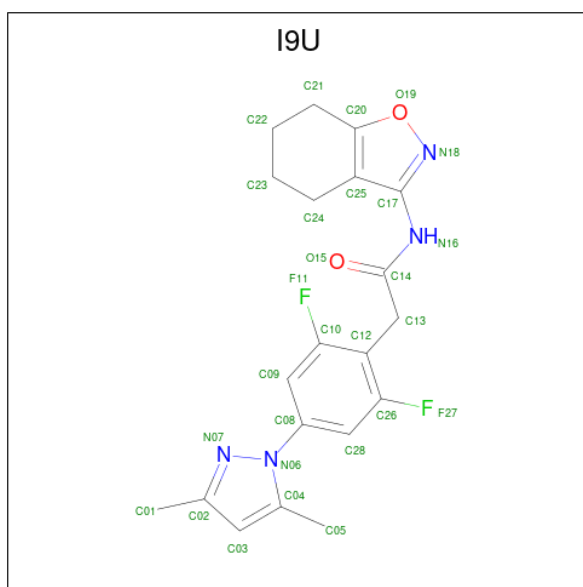
- Molecule 1 is a protein called N-glycosylase/DNA lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	313	2471	1574	447	439	11	0	0	0
1	BBB	304	2369	1513	429	415	12	0	1	0
1	CCC	311	2287	1465	410	402	10	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	8	GLY	-	expression tag	UNP O08760
AAA	10	HIS	SER	conflict	UNP O08760
BBB	8	GLY	-	expression tag	UNP O08760
BBB	10	HIS	SER	conflict	UNP O08760
CCC	8	GLY	-	expression tag	UNP O08760
CCC	10	HIS	SER	conflict	UNP O08760

- Molecule 2 is 2-[4-(3,5-dimethylpyrazol-1-yl)-2,6-bis(fluoranyl)phenyl]- {N}-(4,5,6,7-tetrahydro-1,2-benzoxazol-3-yl)ethanamide (three-letter code: I9U) (formula: C₂₀H₂₀F₂N₄O₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	AAA	1	28	20	2	4	2	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	AAA	1	6	3	3	0	0
3	BBB	1	6	3	3	0	0

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total Ni 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	73	Total O 73 73	0	0
5	BBB	30	Total O 30 30	0	0
5	CCC	24	Total O 24 24	0	0

3 Residue-property plots [i](#)

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.

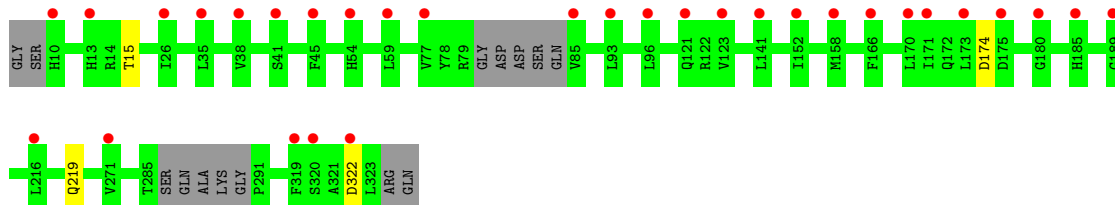
- Molecule 1: N-glycosylase/DNA lyase

Chain AAA: 3% 95%



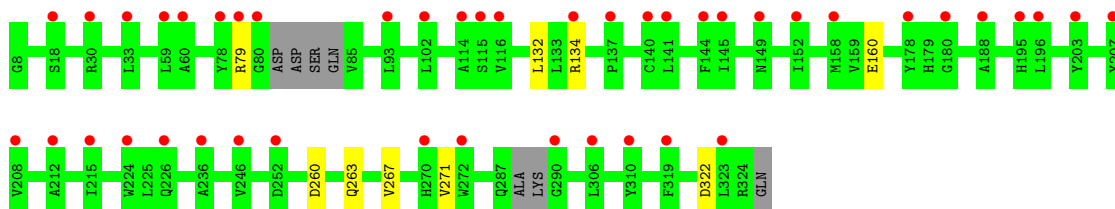
- Molecule 1: N-glycosylase/DNA lyase

Chain BBB: 10% 94%



- Molecule 1: N-glycosylase/DNA lyase

Chain CCC: 14% 95%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.00Å 81.83Å 168.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.63 – 2.35 57.57 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.9 (57.63-2.35) 99.9 (57.57-2.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.30 (at 2.34Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.260 , 0.310 0.261 , 0.307	Depositor DCC
R_{free} test set	2300 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å ²)	67.7	Xtrriage
Anisotropy	0.050	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 49.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7295	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NI, GOL, I9U

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	0.64	0/2541	0.70	0/3462
1	BBB	0.65	0/2437	0.70	0/3325
1	CCC	0.67	0/2352	0.70	0/3224
All	All	0.66	0/7330	0.70	0/10011

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	AAA	2471	0	2396	4	0
1	BBB	2369	0	2268	1	0
1	CCC	2287	0	2073	1	0
2	AAA	28	0	0	0	0
3	AAA	6	0	8	0	0
3	BBB	6	0	8	0	0
4	AAA	1	0	0	0	0
5	AAA	73	0	0	0	0
5	BBB	30	0	0	1	0
5	CCC	24	0	0	0	0
All	All	7295	0	6753	6	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 0.

All (6) 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:BBB:219:GLN:N	5:BBB:501:HOH:O	2.34	0.60
1:CCC:267:VAL:HA	1:CCC:271:VAL:HG11	1.98	0.45
1:AAA:100:PHE:O	1:AAA:131:ARG:HD3	2.18	0.44
1:AAA:275:ALA:O	1:AAA:279:TYR:HB2	2.20	0.41
1:AAA:192:ALA:HB3	1:AAA:209:ARG:HG3	2.03	0.41
1:AAA:216:LEU:HD12	1:AAA:216:LEU:HA	1.87	0.41

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	AAA	309/318 (97%)	300 (97%)	9 (3%)	0	100	100
1	BBB	299/318 (94%)	284 (95%)	15 (5%)	0	100	100
1	CCC	305/318 (96%)	283 (93%)	22 (7%)	0	100	100
All	All	913/954 (96%)	867 (95%)	46 (5%)	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	AAA	256/266 (96%)	250 (98%)	6 (2%)	50	61
1	BBB	240/266 (90%)	237 (99%)	3 (1%)	69	80
1	CCC	207/266 (78%)	200 (97%)	7 (3%)	37	46
All	All	703/798 (88%)	687 (98%)	16 (2%)	50	61

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	11	MET
1	AAA	30	ARG
1	AAA	58	VAL
1	AAA	132	LEU
1	AAA	216	LEU
1	AAA	272	TRP
1	BBB	15	THR
1	BBB	174	ASP
1	BBB	322	ASP
1	CCC	79	ARG
1	CCC	132	LEU
1	CCC	134	ARG
1	CCC	160	GLU
1	CCC	260	ASP
1	CCC	263	GLN
1	CCC	322	ASP

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 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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
3	GOL	AAA	402	-	5,5,5	0.10	0	5,5,5	0.28	0
2	I9U	AAA	401	-	25,31,31	1.75	6 (24%)	27,45,45	2.61	10 (37%)
3	GOL	BBB	401	-	5,5,5	0.10	0	5,5,5	0.32	0

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
3	GOL	AAA	402	-	-	2/4/4/4	-
2	I9U	AAA	401	-	-	0/10/19/19	0/4/4/4
3	GOL	BBB	401	-	-	2/4/4/4	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	401	I9U	C14-N16	4.71	1.46	1.35
2	AAA	401	I9U	C08-N06	-3.27	1.34	1.44
2	AAA	401	I9U	C17-N16	3.15	1.45	1.39
2	AAA	401	I9U	C21-C20	3.11	1.53	1.50
2	AAA	401	I9U	C24-C25	2.55	1.55	1.51
2	AAA	401	I9U	O15-C14	-2.01	1.19	1.23

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	401	I9U	C28-C26-C12	-6.66	119.91	124.59
2	AAA	401	I9U	C09-C10-C12	-5.54	120.70	124.59
2	AAA	401	I9U	C04-N06-N07	-4.65	107.68	111.81

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	401	I9U	C02-N07-N06	3.84	110.11	105.66
2	AAA	401	I9U	C26-C12-C10	3.54	119.44	114.51
2	AAA	401	I9U	C09-C08-N06	3.33	123.31	119.07
2	AAA	401	I9U	C01-C02-N07	3.13	126.31	120.07
2	AAA	401	I9U	C05-C04-C03	-3.07	123.09	128.72
2	AAA	401	I9U	C09-C08-C28	-2.40	118.51	121.61
2	AAA	401	I9U	F27-C26-C12	2.20	120.26	117.63

There are no chirality outliers.

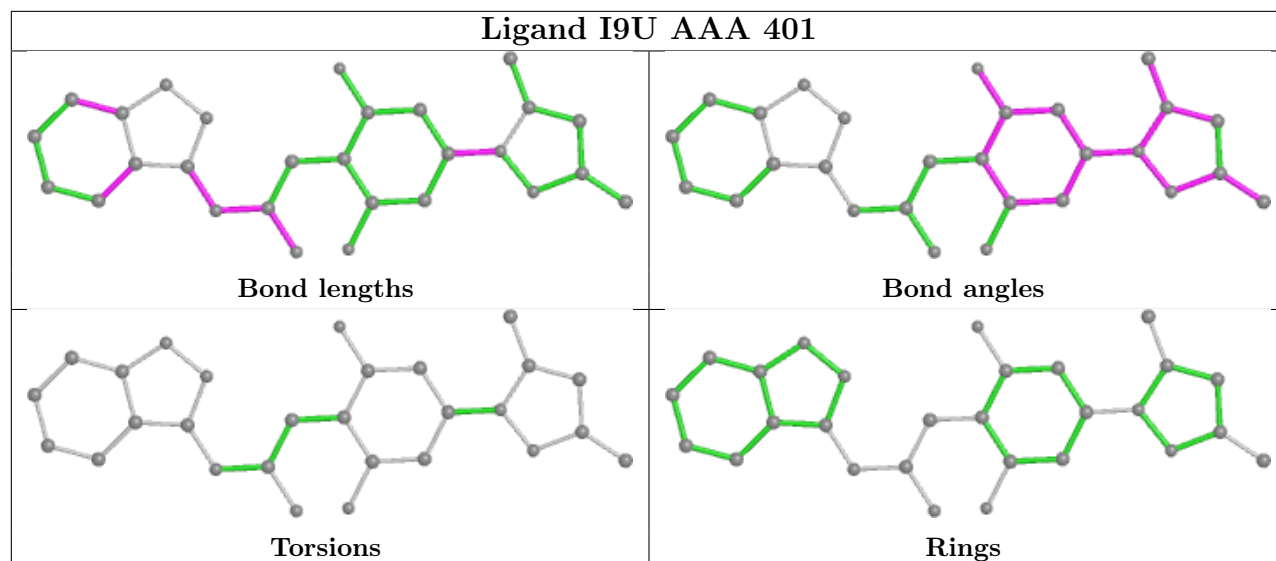
All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	402	GOL	O1-C1-C2-O2
3	AAA	402	GOL	O1-C1-C2-C3
3	BBB	401	GOL	C1-C2-C3-O3
3	BBB	401	GOL	O2-C2-C3-O3

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

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	313/318 (98%)	0.69	10 (3%) 47 59	43, 62, 96, 125	0
1	BBB	304/318 (95%)	0.89	31 (10%) 6 11	54, 87, 119, 133	0
1	CCC	311/318 (97%)	0.98	44 (14%) 2 4	56, 96, 128, 133	0
All	All	928/954 (97%)	0.85	85 (9%) 9 14	43, 83, 120, 133	0

All (85) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	246	VAL	5.3
1	CCC	196	LEU	4.7
1	CCC	149	ASN	4.7
1	CCC	203	TYR	4.4
1	CCC	290	GLY	4.4
1	CCC	310	TYR	4.4
1	BBB	96	LEU	4.3
1	CCC	207	TYR	4.2
1	BBB	10	HIS	4.1
1	BBB	319	PHE	4.0
1	CCC	140	CYS	4.0
1	CCC	60	ALA	3.9
1	CCC	114	ALA	3.9
1	CCC	188	ALA	3.7
1	CCC	59	LEU	3.6
1	CCC	115	SER	3.6
1	BBB	171	ILE	3.5
1	BBB	175	ASP	3.5
1	BBB	173	LEU	3.4
1	CCC	178	TYR	3.2
1	BBB	123	VAL	3.2
1	AAA	288	ALA	2.9
1	AAA	273	GLN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	BBB	59	LEU	2.9
1	CCC	145	ILE	2.9
1	BBB	158	MET	2.8
1	CCC	144	PHE	2.8
1	CCC	93	LEU	2.8
1	CCC	80	GLY	2.8
1	CCC	270	HIS	2.8
1	CCC	158	MET	2.8
1	BBB	77	VAL	2.8
1	BBB	189	GLY	2.7
1	BBB	166	PHE	2.7
1	CCC	152	ILE	2.7
1	CCC	78	TYR	2.7
1	BBB	85	VAL	2.7
1	BBB	41	SER	2.7
1	BBB	141	LEU	2.7
1	BBB	322	ASP	2.7
1	CCC	33	LEU	2.7
1	CCC	141	LEU	2.7
1	BBB	26	ILE	2.6
1	AAA	80	GLY	2.6
1	CCC	137	PRO	2.6
1	CCC	18	SER	2.6
1	BBB	38	VAL	2.6
1	AAA	269	VAL	2.5
1	CCC	224	TRP	2.5
1	AAA	60	ALA	2.5
1	CCC	134	ARG	2.5
1	CCC	272	TRP	2.4
1	BBB	216	LEU	2.4
1	CCC	306	LEU	2.4
1	AAA	272	TRP	2.4
1	CCC	252	ASP	2.3
1	BBB	13	HIS	2.3
1	CCC	102	LEU	2.3
1	AAA	81	ASP	2.3
1	CCC	236	ALA	2.3
1	CCC	323	LEU	2.3
1	AAA	27	PRO	2.3
1	CCC	212	ALA	2.3
1	BBB	170	LEU	2.2
1	BBB	180	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	BBB	185	HIS	2.2
1	BBB	45	PHE	2.2
1	CCC	226	GLN	2.2
1	BBB	271	VAL	2.2
1	CCC	30	ARG	2.1
1	CCC	116	VAL	2.1
1	BBB	152	ILE	2.1
1	CCC	208	VAL	2.1
1	CCC	319	PHE	2.1
1	CCC	79	ARG	2.1
1	AAA	96	LEU	2.1
1	BBB	35	LEU	2.1
1	BBB	320	SER	2.1
1	BBB	54	HIS	2.1
1	BBB	93	LEU	2.1
1	CCC	180	GLY	2.1
1	CCC	195	HIS	2.0
1	CCC	215	ILE	2.0
1	BBB	121	GLN	2.0
1	AAA	71	ASP	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, 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.

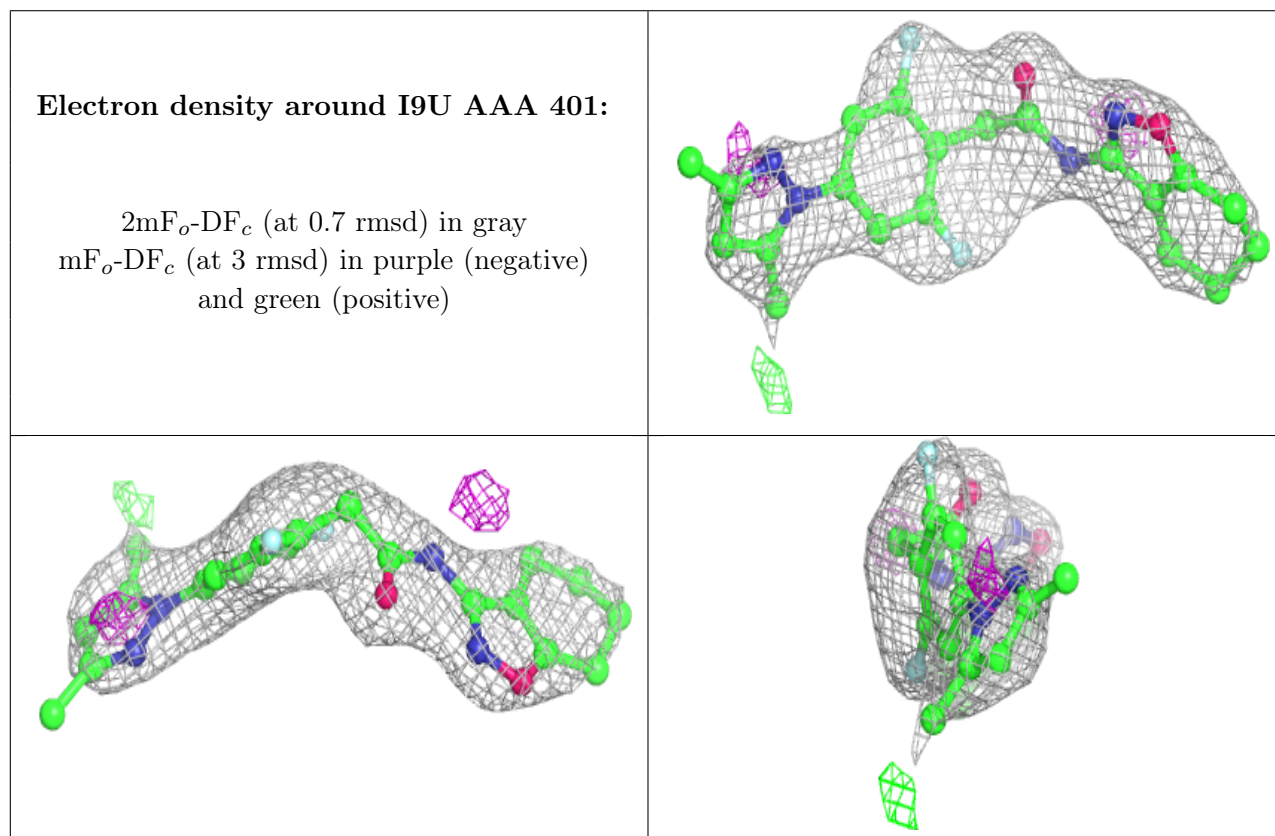
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	GOL	BBB	401	6/6	0.76	0.21	90,93,95,96	0
3	GOL	AAA	402	6/6	0.83	0.24	73,78,78,79	0
2	I9U	AAA	401	28/28	0.88	0.21	78,85,95,95	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	NI	AAA	403	1/1	0.98	0.23	61,61,61,61	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.



6.5 Other polymers [i](#)

There are no such residues in this entry.