wwPDB X-ray Structure Validation Summary Report (i)

Apr 21, 2024 – 04:22 pm BST

PDB ID	:	6TQ2
Title	:	N-TERMINAL BROMODOMAIN OF HUMAN BRD2 WITH 5-(4-(4-fluoro phenyl)-1H-imidazol-5-yl)-1-methylpyridin-2(1H)-one
Authors	:	Chung, C.
Deposited on	:	2019-12-15
Resolution	:	2.26 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 (i) 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 (i)) were used in the production of this report:

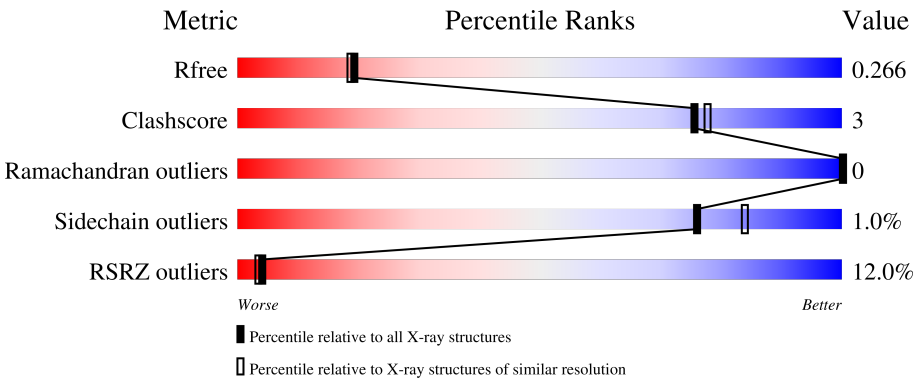
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.26 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	155	8% 66% 6% 28%
1	BBB	155	14% 68% 5% 27%
1	CCC	155	5% 63% 5% • 31%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3244 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 Bromodomain-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	112	954	618	158	168	10	1	2	0
1	BBB	113	961	620	163	168	10	0	2	0
1	CCC	107	913	592	154	158	9	0	2	0

There are 63 discrepancies between the modelled and reference sequences:

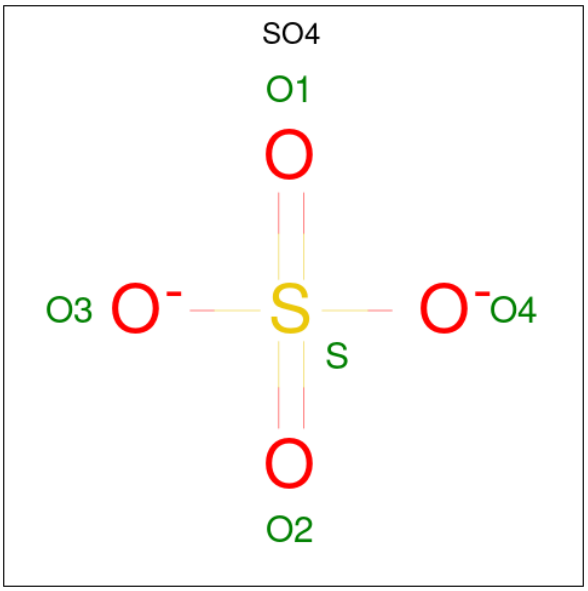
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	46	MET	-	initiating methionine	UNP P25440
AAA	47	GLY	-	expression tag	UNP P25440
AAA	48	SER	-	expression tag	UNP P25440
AAA	49	SER	-	expression tag	UNP P25440
AAA	50	HIS	-	expression tag	UNP P25440
AAA	51	HIS	-	expression tag	UNP P25440
AAA	52	HIS	-	expression tag	UNP P25440
AAA	53	HIS	-	expression tag	UNP P25440
AAA	54	HIS	-	expression tag	UNP P25440
AAA	55	HIS	-	expression tag	UNP P25440
AAA	56	SER	-	expression tag	UNP P25440
AAA	57	SER	-	expression tag	UNP P25440
AAA	58	GLY	-	expression tag	UNP P25440
AAA	59	LEU	-	expression tag	UNP P25440
AAA	60	VAL	-	expression tag	UNP P25440
AAA	61	PRO	-	expression tag	UNP P25440
AAA	62	ARG	-	expression tag	UNP P25440
AAA	63	GLY	-	expression tag	UNP P25440
AAA	64	SER	-	expression tag	UNP P25440
AAA	65	HIS	-	expression tag	UNP P25440
AAA	66	MET	-	expression tag	UNP P25440
BBB	46	MET	-	initiating methionine	UNP P25440
BBB	47	GLY	-	expression tag	UNP P25440

Continued on next page...

Continued from previous page...

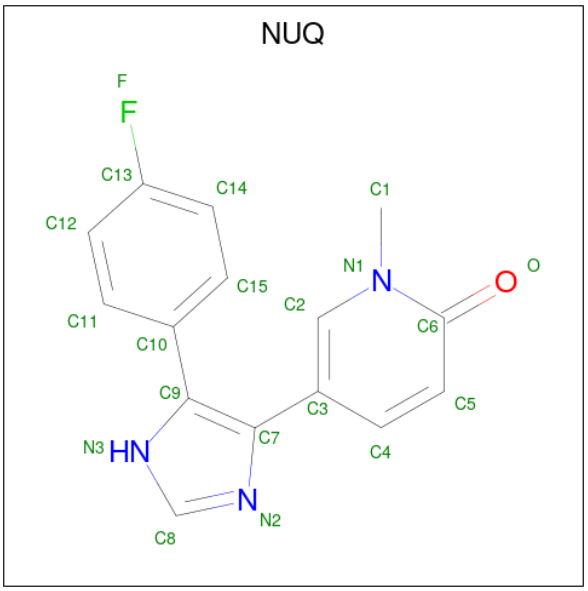
Chain	Residue	Modelled	Actual	Comment	Reference
BBB	48	SER	-	expression tag	UNP P25440
BBB	49	SER	-	expression tag	UNP P25440
BBB	50	HIS	-	expression tag	UNP P25440
BBB	51	HIS	-	expression tag	UNP P25440
BBB	52	HIS	-	expression tag	UNP P25440
BBB	53	HIS	-	expression tag	UNP P25440
BBB	54	HIS	-	expression tag	UNP P25440
BBB	55	HIS	-	expression tag	UNP P25440
BBB	56	SER	-	expression tag	UNP P25440
BBB	57	SER	-	expression tag	UNP P25440
BBB	58	GLY	-	expression tag	UNP P25440
BBB	59	LEU	-	expression tag	UNP P25440
BBB	60	VAL	-	expression tag	UNP P25440
BBB	61	PRO	-	expression tag	UNP P25440
BBB	62	ARG	-	expression tag	UNP P25440
BBB	63	GLY	-	expression tag	UNP P25440
BBB	64	SER	-	expression tag	UNP P25440
BBB	65	HIS	-	expression tag	UNP P25440
BBB	66	MET	-	expression tag	UNP P25440
CCC	46	MET	-	initiating methionine	UNP P25440
CCC	47	GLY	-	expression tag	UNP P25440
CCC	48	SER	-	expression tag	UNP P25440
CCC	49	SER	-	expression tag	UNP P25440
CCC	50	HIS	-	expression tag	UNP P25440
CCC	51	HIS	-	expression tag	UNP P25440
CCC	52	HIS	-	expression tag	UNP P25440
CCC	53	HIS	-	expression tag	UNP P25440
CCC	54	HIS	-	expression tag	UNP P25440
CCC	55	HIS	-	expression tag	UNP P25440
CCC	56	SER	-	expression tag	UNP P25440
CCC	57	SER	-	expression tag	UNP P25440
CCC	58	GLY	-	expression tag	UNP P25440
CCC	59	LEU	-	expression tag	UNP P25440
CCC	60	VAL	-	expression tag	UNP P25440
CCC	61	PRO	-	expression tag	UNP P25440
CCC	62	ARG	-	expression tag	UNP P25440
CCC	63	GLY	-	expression tag	UNP P25440
CCC	64	SER	-	expression tag	UNP P25440
CCC	65	HIS	-	expression tag	UNP P25440
CCC	66	MET	-	expression tag	UNP P25440

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



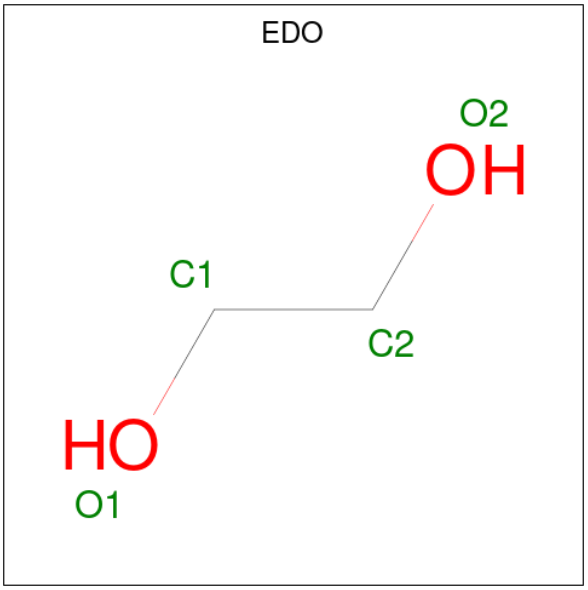
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total O S 5 4 1	0	0
2	AAA	1	Total O S 5 4 1	0	0
2	CCC	1	Total O S 5 4 1	0	0
2	CCC	1	Total O S 5 4 1	0	0

- Molecule 3 is 5-[5-(4-fluorophenyl)-1 {H}-imidazol-4-yl]-1-methyl-pyridin-2-one (three-letter code: NUQ) (formula: $C_{15}H_{12}FN_3O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
3	AAA	1	20	15	1	3	1	0	0
3	BBB	1	20	15	1	3	1	0	0
3	CCC	1	20	15	1	3	1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	AAA	1	4	2	2	0	0

- Molecule 5 is water.

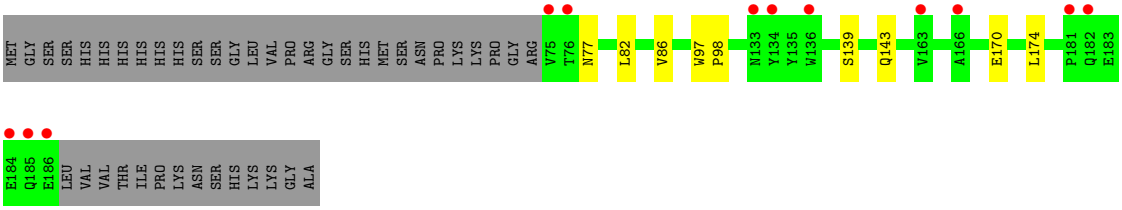
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	AAA	138	138	138	0	0
5	BBB	109	109	109	0	0
5	CCC	85	85	85	0	0

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.

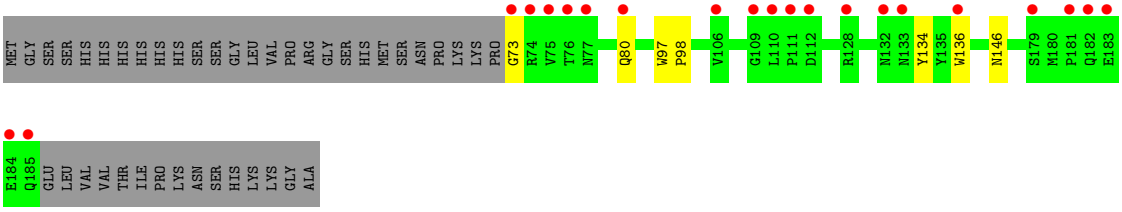
• Molecule 1: Bromodomain-containing protein 2

Chain AAA:



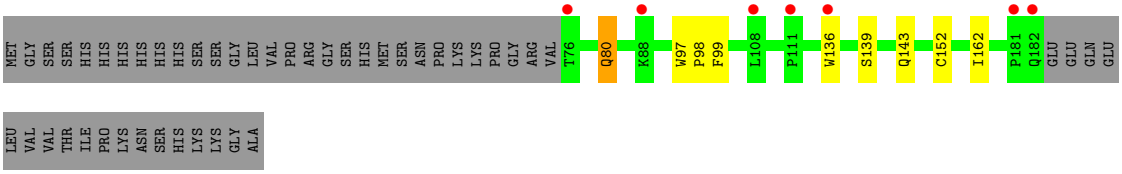
• Molecule 1: Bromodomain-containing protein 2

Chain BBB:



• Molecule 1: Bromodomain-containing protein 2

Chain CCC:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	115.79Å 55.85Å 67.44Å 90.00° 94.03° 90.00°	Depositor
Resolution (Å)	40.89 – 2.26 40.88 – 2.26	Depositor EDS
% Data completeness (in resolution range)	98.8 (40.89-2.26) 98.8 (40.88-2.26)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.94 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.202 , 0.269 0.209 , 0.266	Depositor DCC
R_{free} test set	1022 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	30.9	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 54.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3244	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.86% 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: SO4, NUQ, EDO

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	0.65	0/979	0.69	0/1324
1	BBB	0.63	0/986	0.67	0/1332
1	CCC	0.64	0/938	0.67	0/1269
All	All	0.64	0/2903	0.68	0/3925

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	954	0	948	7	0
1	BBB	961	0	955	6	0
1	CCC	913	0	909	8	0
2	AAA	10	0	0	0	0
2	CCC	10	0	0	0	0
3	AAA	20	0	0	0	0
3	BBB	20	0	0	0	0
3	CCC	20	0	0	0	0
4	AAA	4	0	6	0	0
5	AAA	138	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	BBB	109	0	0	1	0
5	CCC	85	0	0	1	0
All	All	3244	0	2818	18	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 18 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:CCC:80[A]:GLN:HE21	1:CCC:80[A]:GLN:HA	1.49	0.78
1:BBB:134:TYR:OH	5:BBB:401:HOH:O	2.09	0.69
1:CCC:80[A]:GLN:HA	1:CCC:80[A]:GLN:NE2	2.17	0.60
1:AAA:97:TRP:CG	1:AAA:98:PRO:HD3	2.43	0.54
1:AAA:174:LEU:HD11	1:BBB:146:ASN:HB2	1.92	0.51

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	112/155 (72%)	111 (99%)	1 (1%)	0	100	100
1	BBB	113/155 (73%)	110 (97%)	3 (3%)	0	100	100
1	CCC	107/155 (69%)	104 (97%)	3 (3%)	0	100	100
All	All	332/465 (71%)	325 (98%)	7 (2%)	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	105/140 (75%)	105 (100%)	0	100	100
1	BBB	105/140 (75%)	104 (99%)	1 (1%)	76	84
1	CCC	100/140 (71%)	97 (97%)	3 (3%)	41	50
All	All	310/420 (74%)	306 (99%)	4 (1%)	76	79

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	80	GLN
1	CCC	80[A]	GLN
1	CCC	80[B]	GLN
1	CCC	136	TRP

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)

8 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	SO4	AAA	302	-	4,4,4	0.37	0	6,6,6	0.05	0
2	SO4	CCC	301	-	4,4,4	0.42	0	6,6,6	0.09	0
3	NUQ	AAA	303	-	18,22,22	0.88	1 (5%)	23,31,31	1.03	2 (8%)
3	NUQ	BBB	301	-	18,22,22	0.88	1 (5%)	23,31,31	0.95	2 (8%)
4	EDO	AAA	304	-	3,3,3	0.07	0	2,2,2	0.15	0
2	SO4	CCC	302	-	4,4,4	0.38	0	6,6,6	0.07	0
2	SO4	AAA	301	-	4,4,4	0.38	0	6,6,6	0.07	0
3	NUQ	CCC	303	-	18,22,22	0.89	1 (5%)	23,31,31	1.00	2 (8%)

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	EDO	AAA	304	-	-	1/1/1/1	-
3	NUQ	BBB	301	-	-	2/4/8/8	0/3/3/3
3	NUQ	AAA	303	-	-	1/4/8/8	0/3/3/3
3	NUQ	CCC	303	-	-	2/4/8/8	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	BBB	301	NUQ	C7-C9	-2.64	1.39	1.43
3	CCC	303	NUQ	C7-C9	-2.54	1.39	1.43
3	AAA	303	NUQ	C7-C9	-2.53	1.39	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	303	NUQ	C7-C9-N3	-3.14	105.29	113.76
3	CCC	303	NUQ	C7-C9-N3	-3.06	105.50	113.76
3	BBB	301	NUQ	C7-C9-N3	-3.01	105.65	113.76

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	303	NUQ	C8-N3-C9	2.46	107.67	102.99
3	CCC	303	NUQ	C8-N3-C9	2.31	107.38	102.99

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

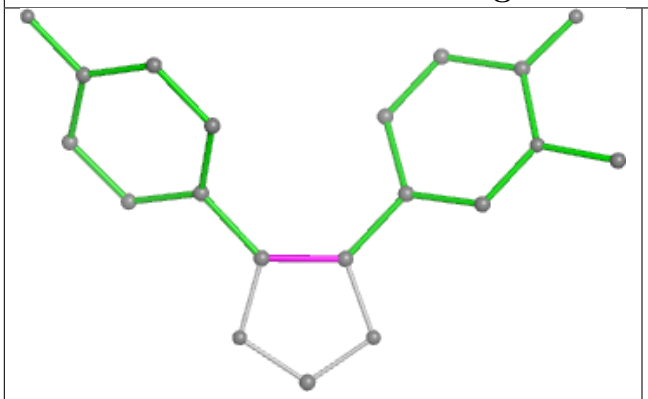
Mol	Chain	Res	Type	Atoms
3	CCC	303	NUQ	C15-C10-C9-N3
3	CCC	303	NUQ	C11-C10-C9-N3
3	BBB	301	NUQ	C11-C10-C9-N3
4	AAA	304	EDO	O1-C1-C2-O2
3	AAA	303	NUQ	C11-C10-C9-N3

There are no ring outliers.

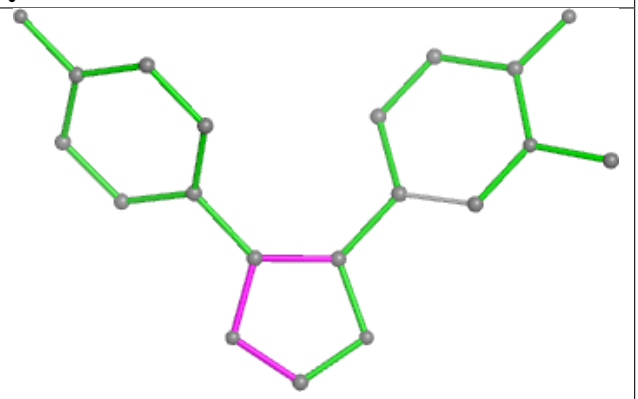
No monomer is involved in short contacts.

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.

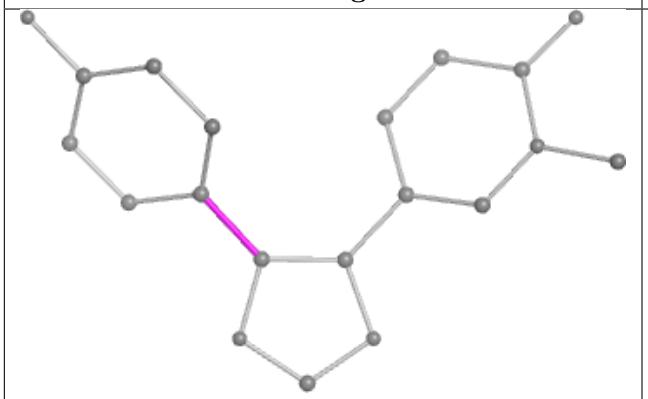
Ligand NUQ AAA 303



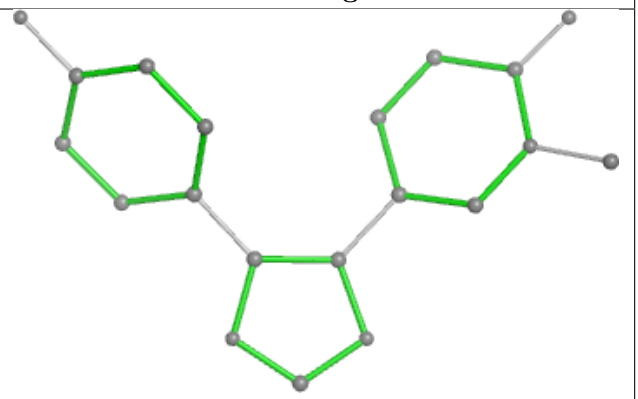
Bond lengths



Bond angles

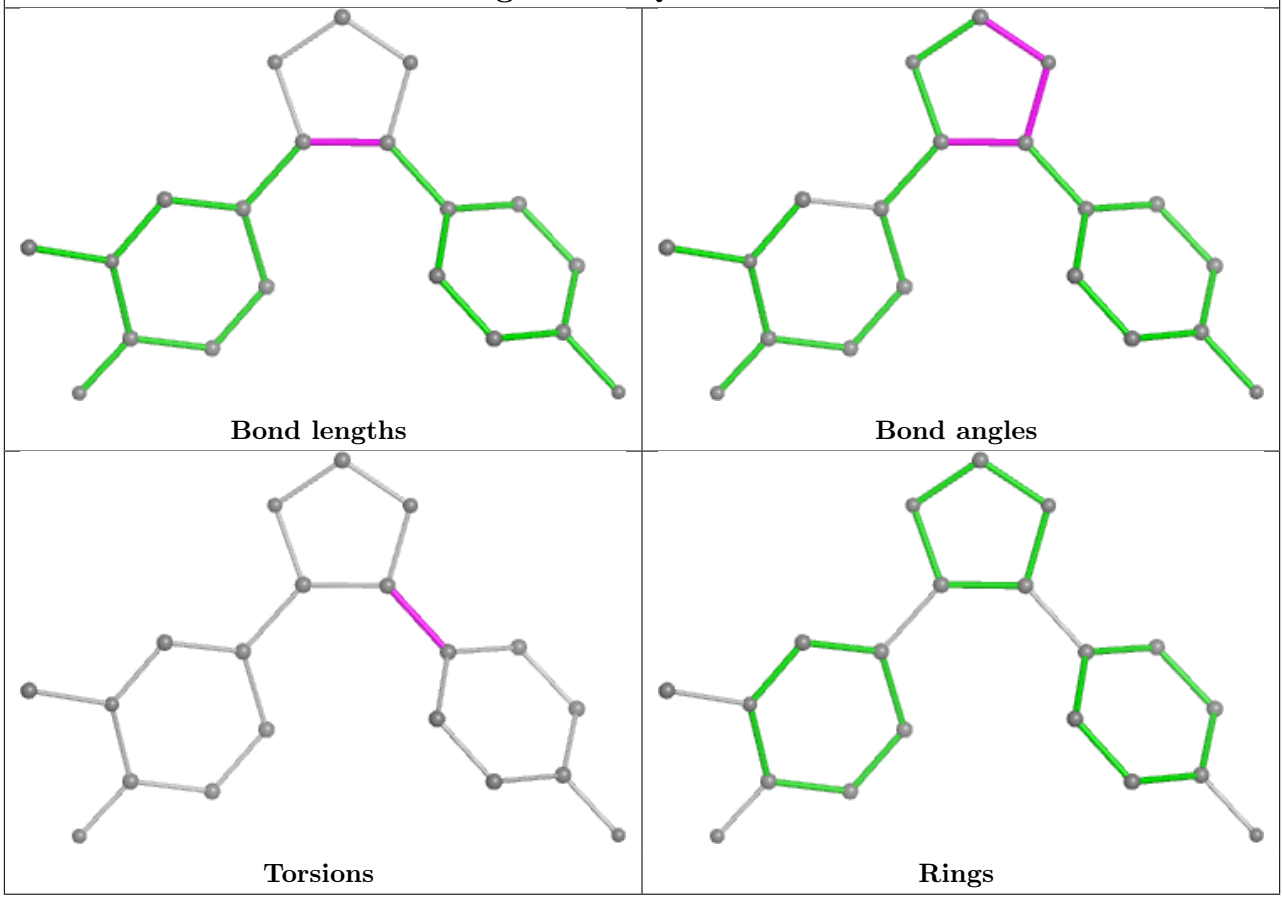


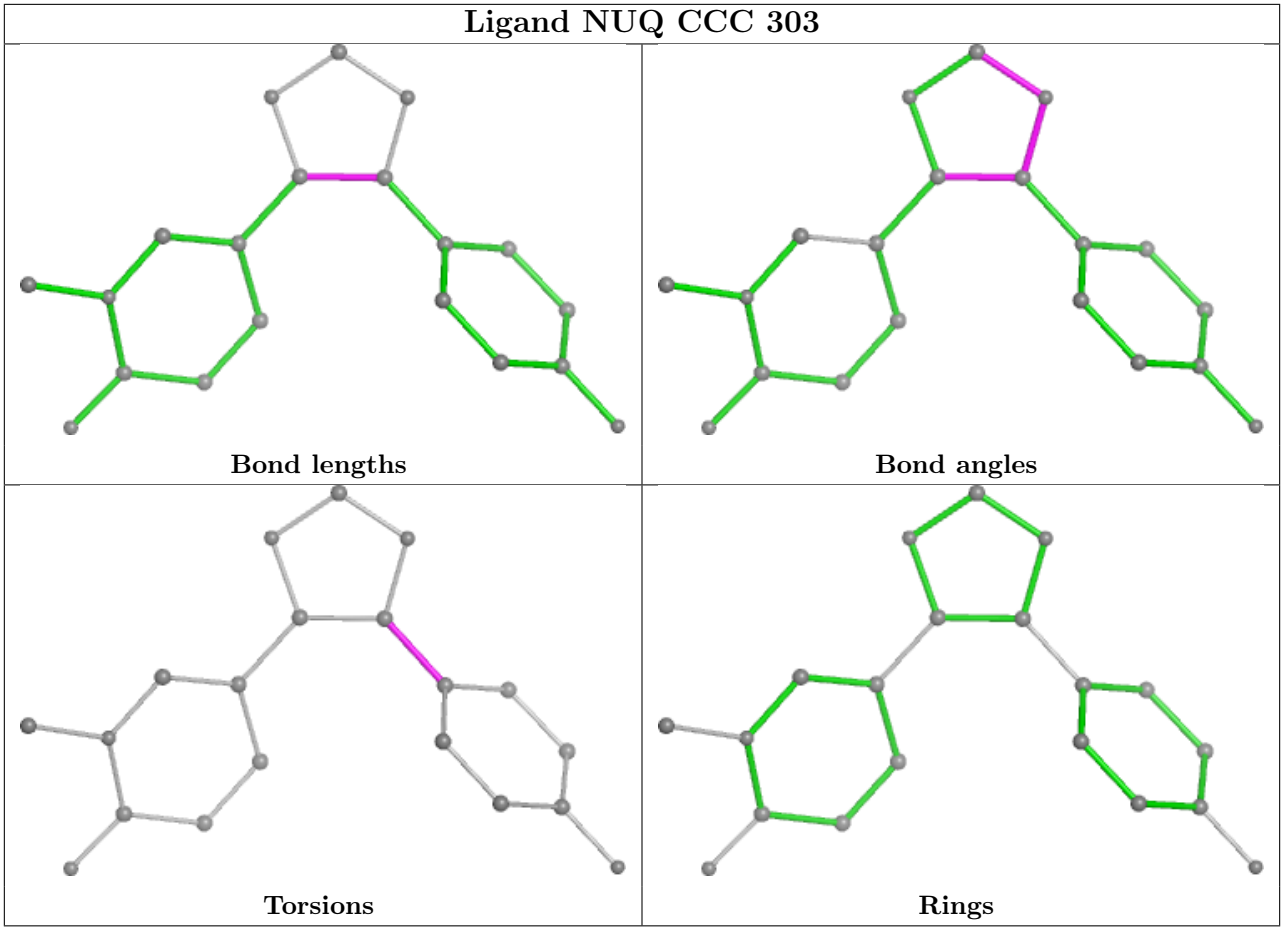
Torsions



Rings

Ligand NUQ BBB 301





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	AAA	112/155 (72%)	0.82	12 (10%)	6	5	35, 52, 89, 114	0
1	BBB	113/155 (72%)	1.02	21 (18%)	1	1	35, 54, 94, 120	0
1	CCC	107/155 (69%)	0.60	7 (6%)	18	20	38, 55, 81, 105	0
All	All	332/465 (71%)	0.82	40 (12%)	4	3	35, 53, 89, 120	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	75	VAL	9.4
1	AAA	136	TRP	8.1
1	BBB	77	ASN	5.8
1	CCC	76	THR	5.4
1	BBB	111	PRO	4.7

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)

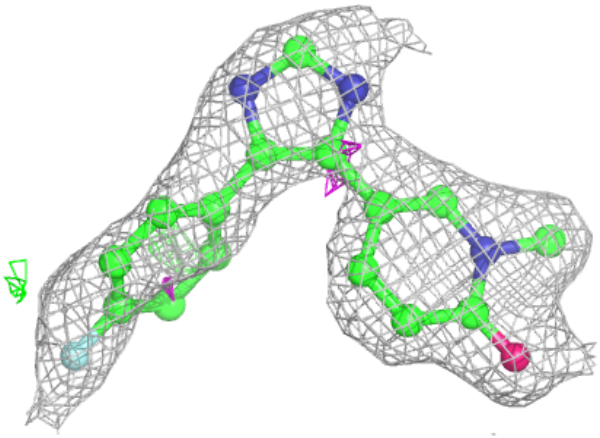
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

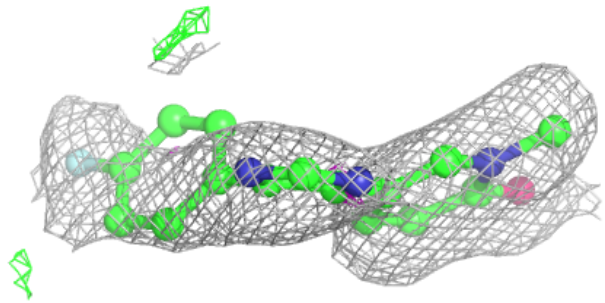
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	AAA	304	4/4	0.69	0.27	92,95,96,97	0
3	NUQ	CCC	303	20/20	0.80	0.22	48,68,89,92	0
3	NUQ	AAA	303	20/20	0.87	0.18	42,52,83,86	0
2	SO4	CCC	302	5/5	0.88	0.23	104,105,114,118	0
3	NUQ	BBB	301	20/20	0.88	0.19	46,60,88,89	0
2	SO4	AAA	302	5/5	0.90	0.34	107,115,123,128	0
2	SO4	AAA	301	5/5	0.92	0.16	97,101,108,115	0
2	SO4	CCC	301	5/5	0.98	0.16	59,61,65,70	0

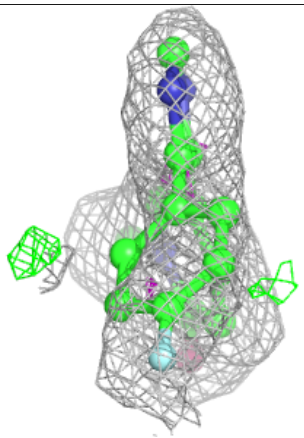
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around NUQ CCC 303:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

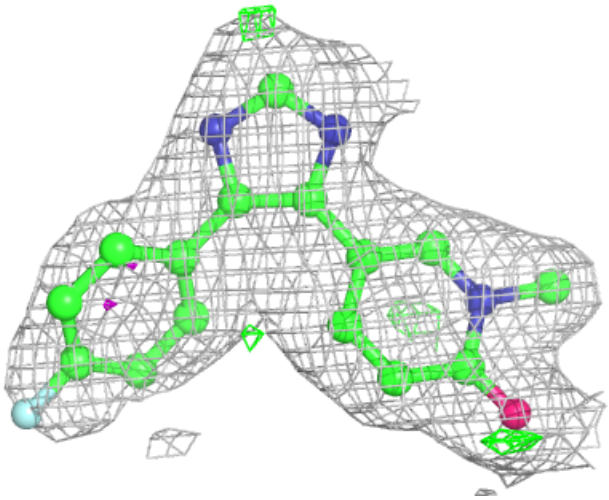


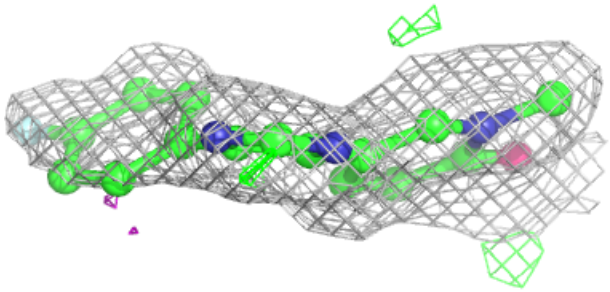


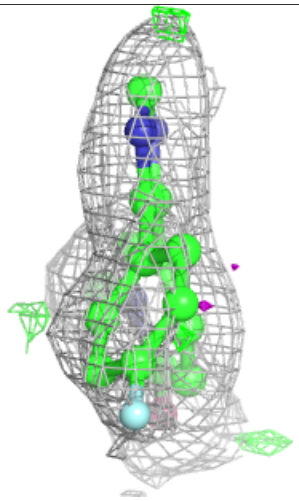


Electron density around NUQ AAA 303:

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

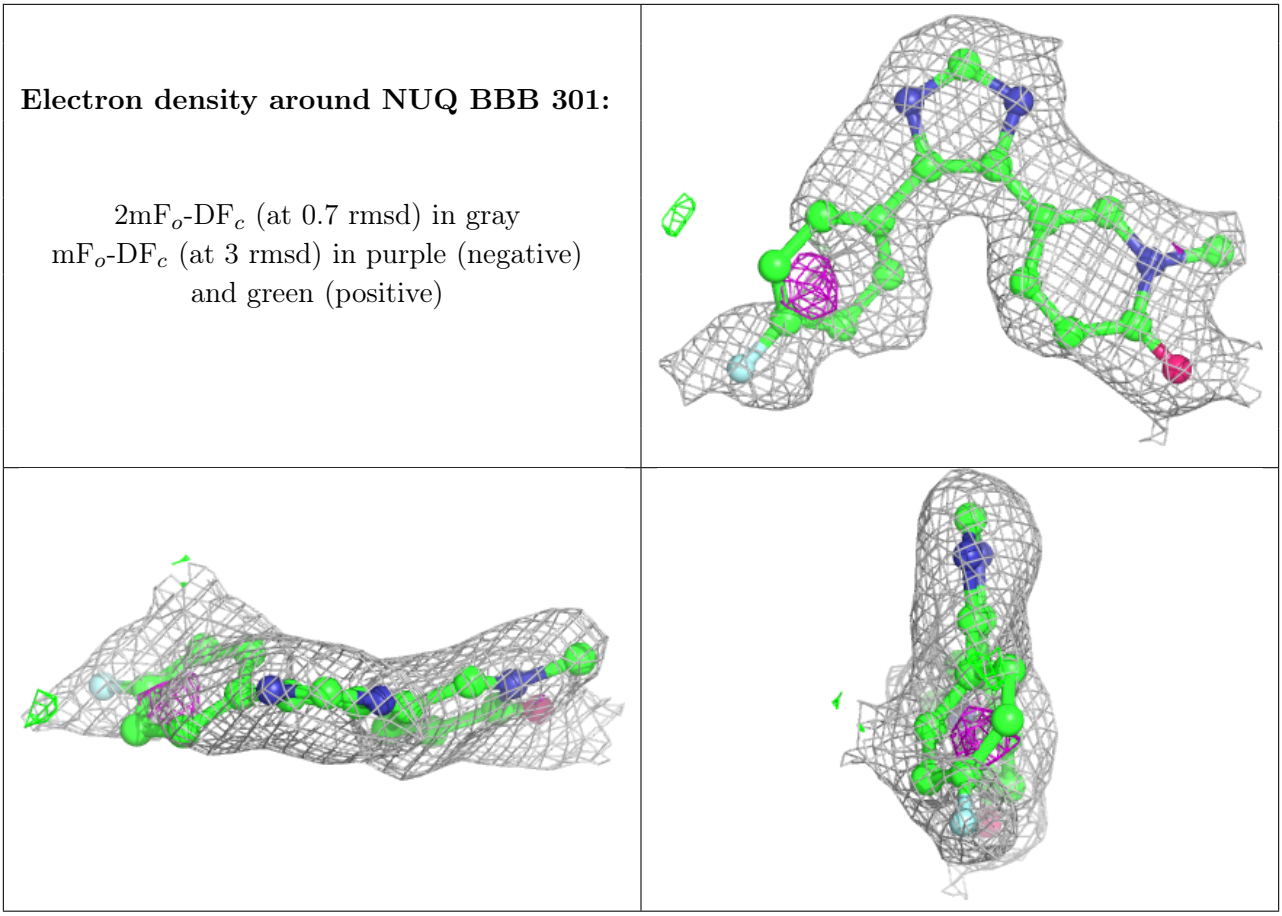






Electron density around NUQ BBB 301:

$2mF_o$-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.