WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Nov 30, 2021 – 02:12 pm GMT

PDB ID	:	7B48
Title	:	Structural basis of reactivation of oncogenic p53 mutants by a small molecule: methylene quinuclidinone (MQ). Human p53DBD-R273H mutant bound to MQ: R273H-MQ (II)
Authors	:	Degtjarik, O.; Rozenberg, H.; Shakked, Z.
Deposited on	:	2020-12-02
Resolution	:	2.05 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 (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 (i)

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

The reported resolution of this entry is 2.05 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.
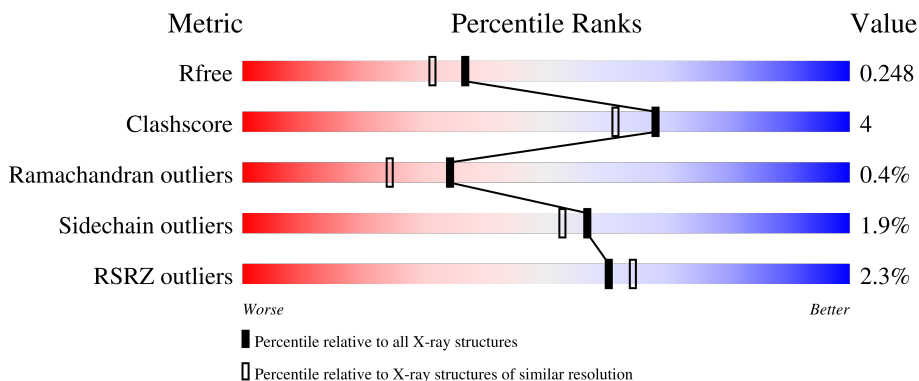
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	A	200	3% / 83% / 13% / •
1	B	200	2% / 86% / 10% / •
1	C	200	2% / 83% / 12% / •
1	D	200	2% / 86% / 8% / 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	D	303	-	-	-	X
6	PEG	C	305	-	-	X	-

2   Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6615 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 Cellular tumor antigen p53.

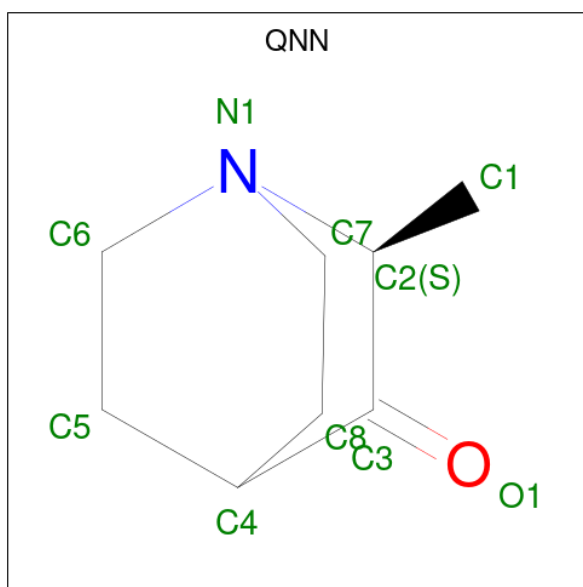| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total	C	N	O	S	0	4	0
			1526	941	277	289	19			
1	B	193	Total	C	N	O	S	0	3	0
			1519	937	277	286	19			
1	C	192	Total	C	N	O	S	0	3	0
			1517	936	277	285	19			
1	D	189	Total	C	N	O	S	0	5	0
			1499	924	276	281	18			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	273	HIS	ARG	engineered mutation	UNP P04637
B	273	HIS	ARG	engineered mutation	UNP P04637
C	273	HIS	ARG	engineered mutation	UNP P04637
D	273	HIS	ARG	engineered mutation	UNP P04637

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		

- Molecule 3 is (2 {S})-2-methyl-1-azabicyclo[2.2.2]octan-3-one (three-letter code: QNN) (formula: $C_8H_{13}NO$) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total 10	C 8	N 1	O 1	0	1
3	C	1	Total 10	C 8	N 1	O 1	0	1

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

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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
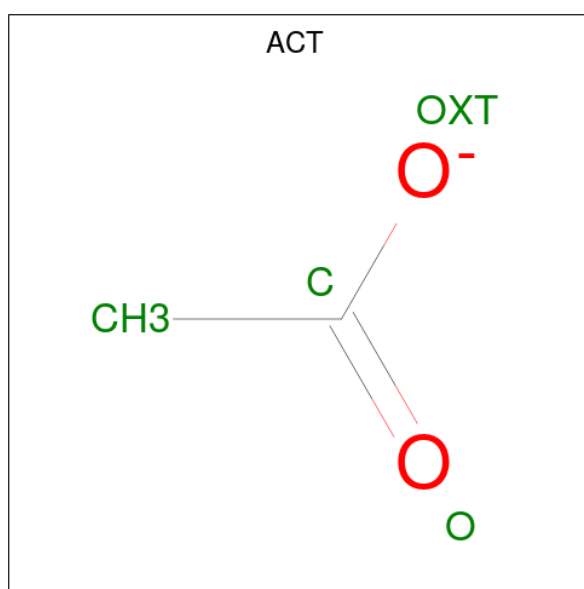
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 4	C 2	O 2	0	0
5	D	1	Total 4	C 2	O 2	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 7	C 4	O 3	0	0
6	C	1	Total 7	C 4	O 3	0	0
6	D	1	Total 7	C 4	O 3	0	0

• Molecule 7 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	142	Total 144	O 144	0	2
7	B	83	Total 83	O 83	0	0
7	C	110	Total 112	O 112	0	2
7	D	134	Total 138	O 138	0	4

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.
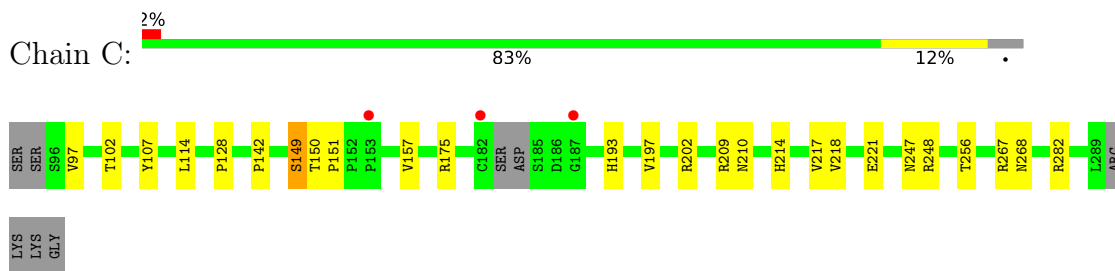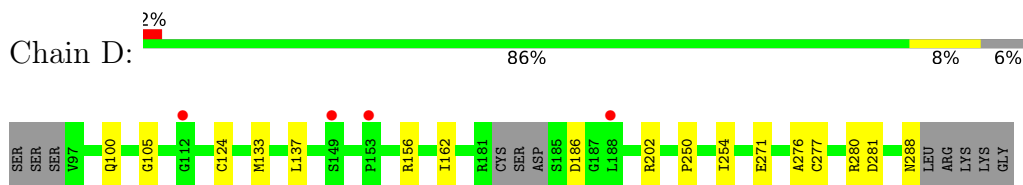
- Molecule 1: Cellular tumor antigen p53
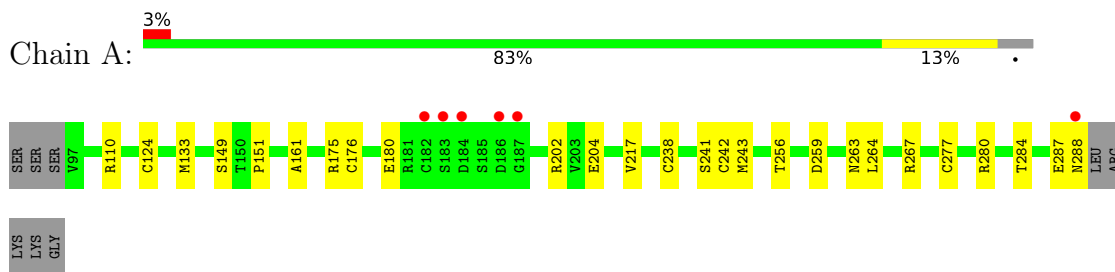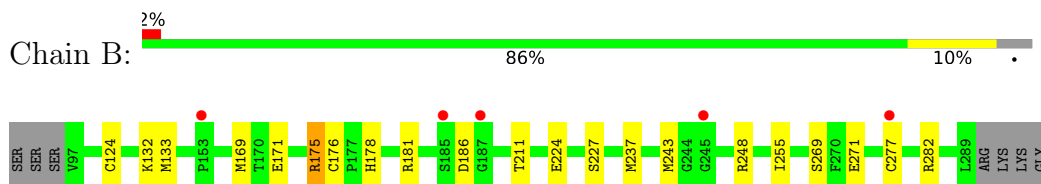
Chain A:

- Molecule 1: Cellular tumor antigen p53

Chain B:

- Molecule 1: Cellular tumor antigen p53

Chain C:

- Molecule 1: Cellular tumor antigen p53

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	69.25Å 72.00Å 85.23Å 90.00° 90.07° 90.00°	Depositor
Resolution (Å)	36.67 – 2.05 36.67 – 2.05	Depositor EDS
% Data completeness (in resolution range)	93.9 (36.67-2.05) 98.0 (36.67-2.05)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.27 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.8.0266	Depositor
R, R_{free}	0.176 , 0.228 0.194 , 0.248	Depositor DCC
R_{free} test set	2503 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.609	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.53, < L^2 > = 0.37$	Xtriage
Estimated twinning fraction	0.000 for -k,-h,-l 0.000 for k,h,-l 0.049 for h,-k,-l	Xtriage
Reported twinning fraction	0.756 for H, K, L 0.244 for -h,-k,l	Depositor
Outliers	0 of 51884 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6615	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.19% 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: ACT, PEG, ZN, EDO, QNN

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	0.93	1/1562 (0.1%)	1.07	7/2122 (0.3%)
1	B	0.91	2/1555 (0.1%)	1.09	4/2115 (0.2%)
1	C	0.87	0/1552	1.06	2/2107 (0.1%)
1	D	0.87	0/1533	1.04	2/2081 (0.1%)
All	All	0.89	3/6202 (0.0%)	1.06	15/8425 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	161	ALA	C-O	6.41	1.35	1.23
1	B	171	GLU	C-O	6.13	1.35	1.23
1	B	255	ILE	C-O	5.21	1.33	1.23

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	175	ARG	NE-CZ-NH1	-8.96	115.82	120.30
1	C	175	ARG	NE-CZ-NH1	-8.88	115.86	120.30
1	A	175	ARG	NE-CZ-NH1	-7.83	116.39	120.30
1	A	202	ARG	NE-CZ-NH1	7.23	123.92	120.30
1	B	175	ARG	NE-CZ-NH2	6.79	123.70	120.30

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	A	1526	0	1455	11	0
1	B	1519	0	1451	11	0
1	C	1517	0	1449	17	0
1	D	1499	0	1422	9	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	B	10	0	0	0	0
3	C	10	0	0	0	0
4	B	8	0	12	2	0
4	C	8	0	12	0	0
4	D	8	0	12	2	0
5	B	4	0	3	0	0
5	D	4	0	3	0	0
6	B	7	0	10	0	0
6	C	7	0	10	5	0
6	D	7	0	10	1	0
7	A	144	0	0	1	0
7	B	83	0	0	1	0
7	C	112	0	0	1	0
7	D	138	0	0	0	0
All	All	6615	0	5849	45	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 4.

The worst 5 of 45 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:D:280:ARG:HG2	1:D:280:ARG:HH11	1.45	0.81
1:D:124:CYS:SG	1:D:133[A]:MET:SD	2.93	0.67
1:D:271:GLU:OE2	6:D:305:PEG:O4	2.14	0.65
1:B:211:THR:OG1	4:B:303:EDO:H22	1.98	0.64
1:C:202:ARG:NH2	1:C:221:GLU:OE2	2.30	0.62

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	194/200 (97%)	190 (98%)	4 (2%)	0	100	100
1	B	194/200 (97%)	188 (97%)	5 (3%)	1 (0%)	29	18
1	C	191/200 (96%)	184 (96%)	6 (3%)	1 (0%)	29	18
1	D	190/200 (95%)	183 (96%)	6 (3%)	1 (0%)	29	18
All	All	769/800 (96%)	745 (97%)	21 (3%)	3 (0%)	34	24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	97	VAL
1	B	186	ASP
1	D	105	GLY

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	173/179 (97%)	170 (98%)	3 (2%)	60	57
1	B	172/179 (96%)	168 (98%)	4 (2%)	50	44
1	C	172/179 (96%)	170 (99%)	2 (1%)	71	70
1	D	167/179 (93%)	162 (97%)	5 (3%)	41	34
All	All	684/716 (96%)	670 (98%)	14 (2%)	57	50

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	149	SER
1	C	150	THR
1	D	288	ASN
1	D	277[B]	CYS
1	D	281	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	273	HIS
1	C	165	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 ⓘ

Of 17 ligands modelled in this entry, 4 are monoatomic - leaving 13 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.
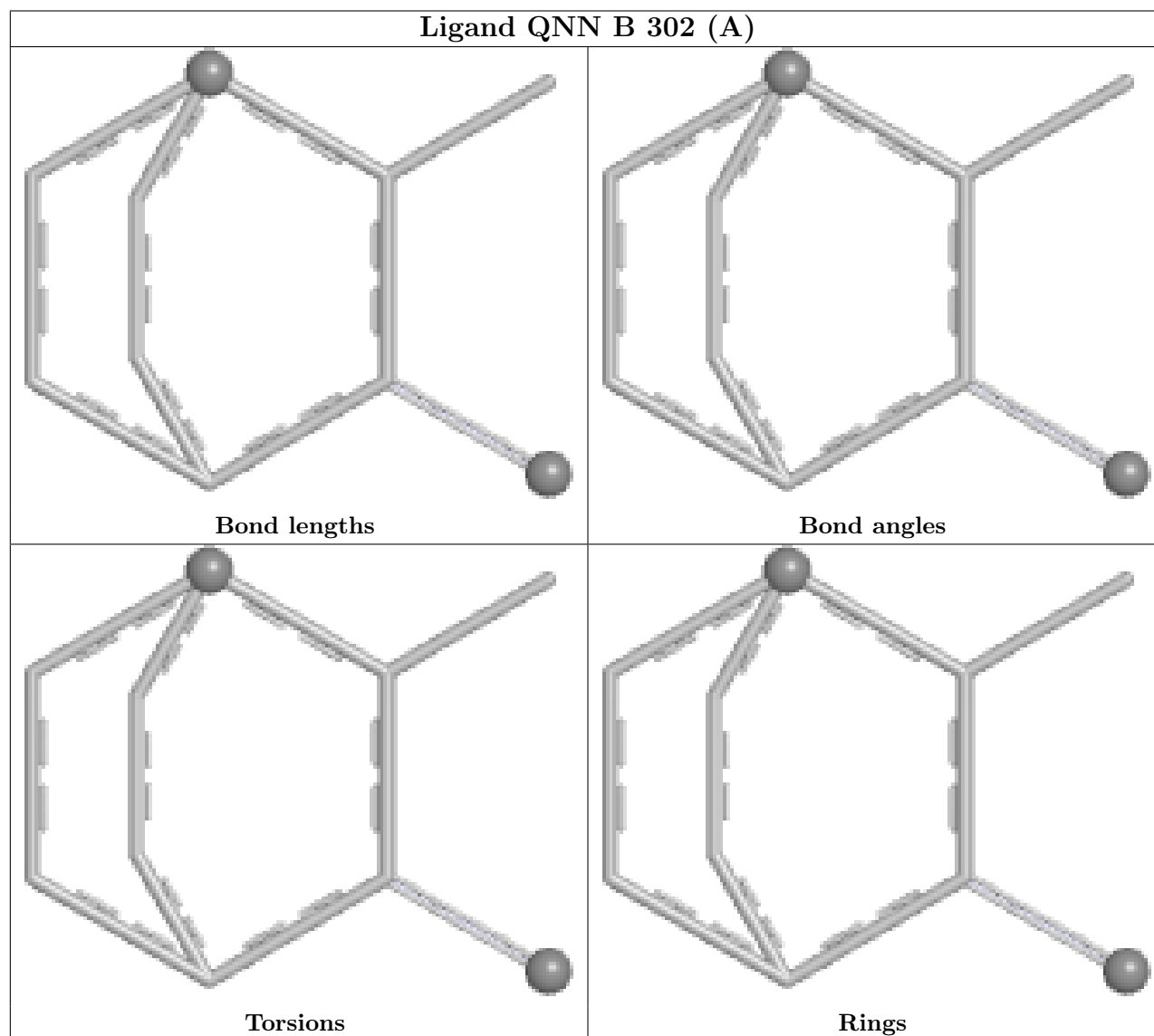
There are no chirality outliers.
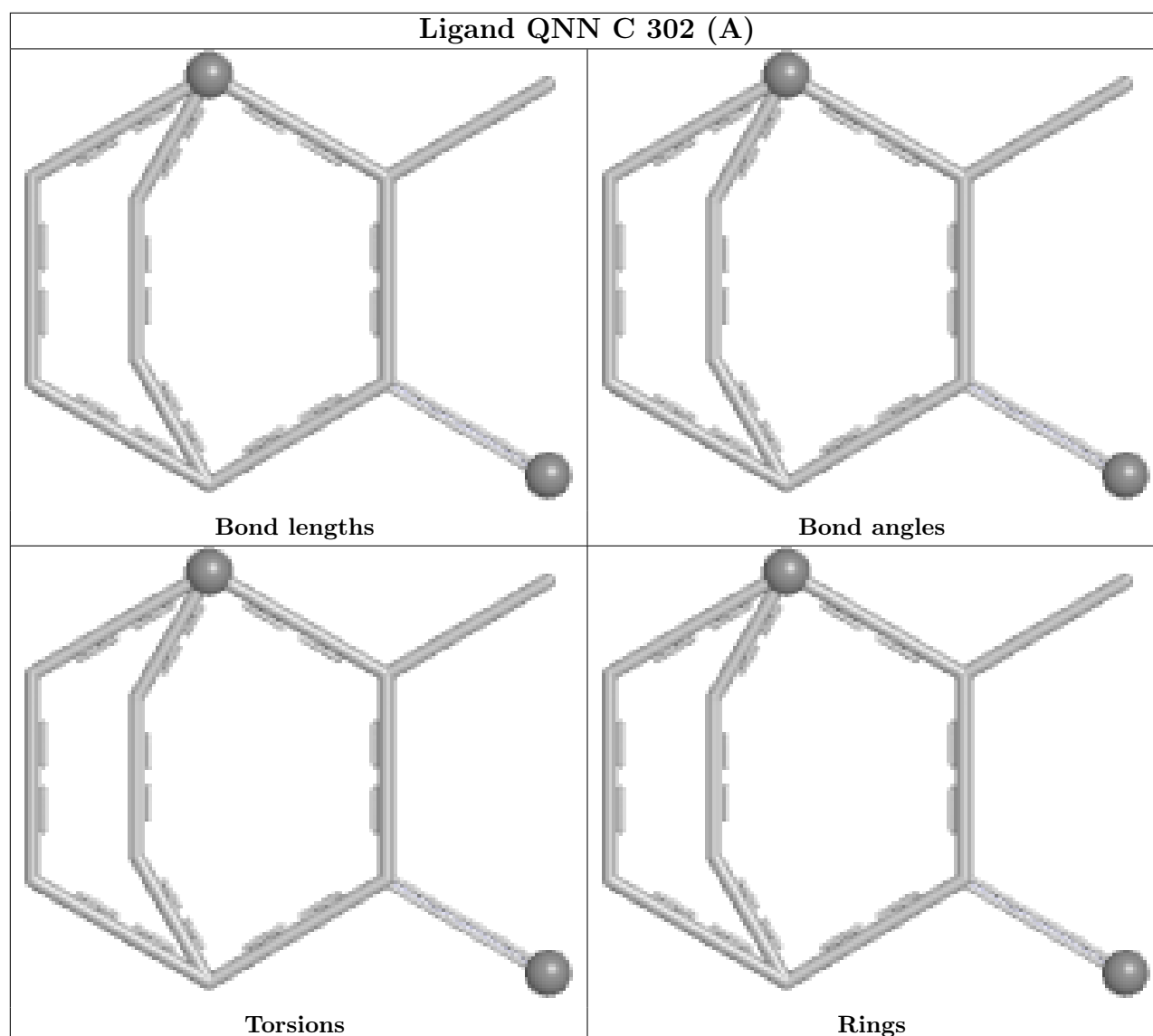
There are no torsion outliers.

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 QNN B 302 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand QNN C 302 (A)

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(Å²)	Q<0.9
1	A	192/200 (96%)	0.04	6 (3%)	49	53	19, 28, 54, 69	0
1	B	193/200 (96%)	0.12	5 (2%)	56	60	17, 34, 53, 64	0
1	C	192/200 (96%)	0.07	3 (1%)	72	74	20, 33, 54, 62	0
1	D	189/200 (94%)	0.06	4 (2%)	63	67	20, 33, 51, 65	0
All	All	766/800 (95%)	0.07	18 (2%)	60	64	17, 32, 54, 69	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	182	CYS	4.3
1	C	187	GLY	3.4
1	B	187	GLY	2.8
1	A	186	ASP	2.8
1	A	182	CYS	2.7

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.
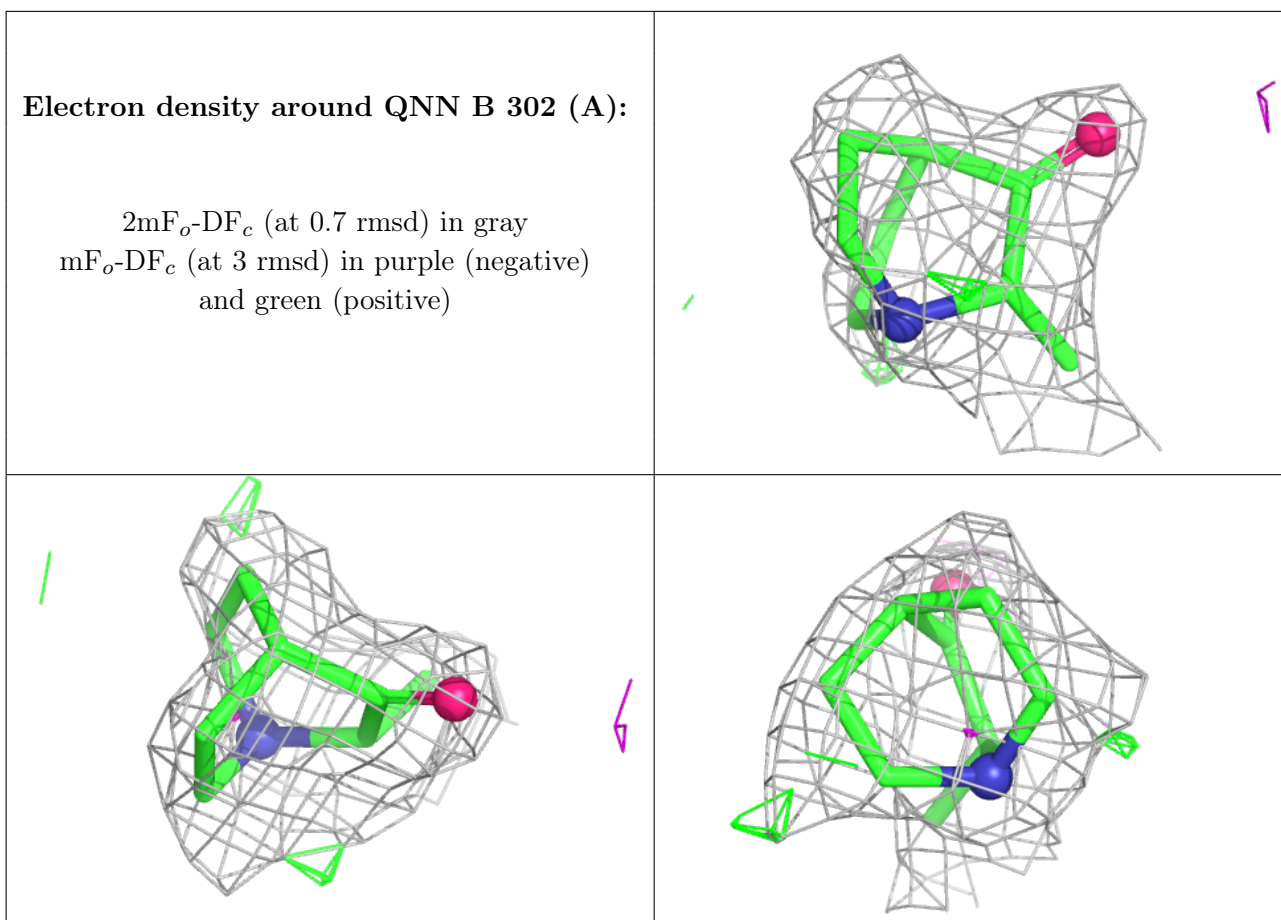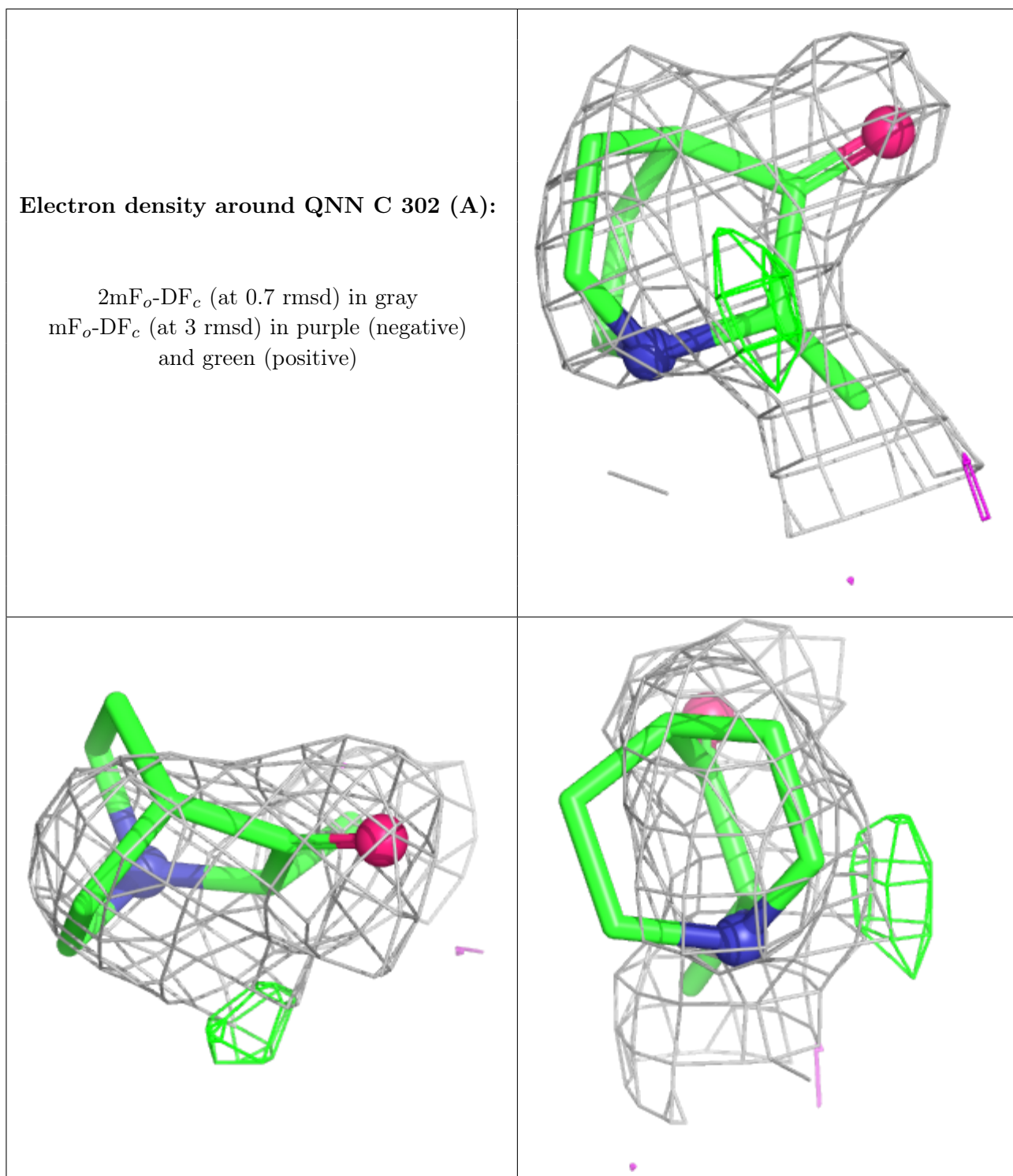
6.4 Ligands ⓘ

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
4	EDO	D	303	4/4	0.49	0.43	45,46,47,47	4
4	EDO	D	302	4/4	0.75	0.29	36,36,37,39	4
4	EDO	B	303	4/4	0.79	0.26	31,35,35,37	4
5	ACT	D	304	4/4	0.79	0.20	43,46,46,47	4
3	QNN	B	302[A]	10/10	0.80	0.21	31,37,38,39	10
3	QNN	C	302[A]	10/10	0.81	0.25	37,53,55,57	10
4	EDO	B	304	4/4	0.84	0.21	42,43,43,45	4
5	ACT	B	305	4/4	0.88	0.10	36,36,40,41	0
6	PEG	C	305	7/7	0.90	0.18	17,19,22,22	7
6	PEG	D	305	7/7	0.90	0.26	35,37,39,39	7
4	EDO	C	304	4/4	0.91	0.28	21,22,22,22	4
6	PEG	B	306	7/7	0.91	0.25	30,33,35,36	7
4	EDO	C	303	4/4	0.94	0.20	23,25,26,28	4
2	ZN	C	301	1/1	0.98	0.07	27,27,27,27	0
2	ZN	B	301	1/1	0.98	0.09	28,28,28,28	0
2	ZN	D	301	1/1	0.99	0.08	34,34,34,34	0
2	ZN	A	301	1/1	0.99	0.08	30,30,30,30	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 QNN B 302 (A):

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

Electron density around QNN C 302 (A):

$2mF_o$-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.