



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 18, 2025 – 10:45 AM EDT

PDB ID : 9NWN / pdb_00009nwn
Title : Structure of minimal CRISPR polymerase bound to ApNHpp
Authors : Doherty, E.E.; Bolling, C.B.; Wilcox, X.E.; Doudna, J.A.
Deposited on : 2025-03-24
Resolution : 2.28 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 ⓘ 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 : 4-5-2 with Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.45.1

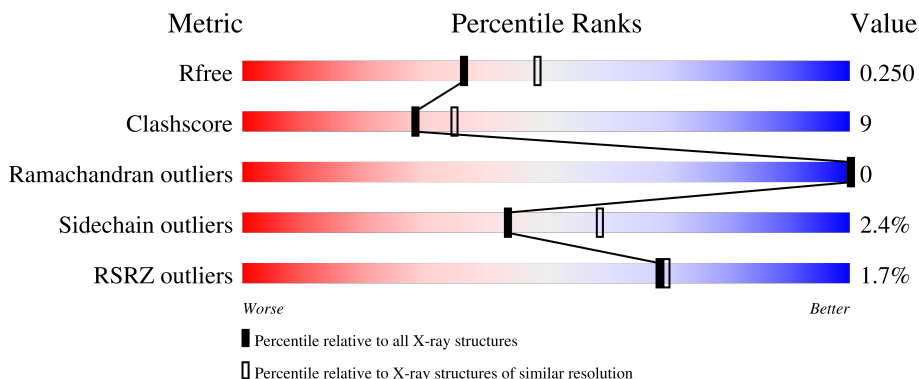
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.28 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}	164625	8487 (2.30-2.26)
Clashscore	180529	9437 (2.30-2.26)
Ramachandran outliers	177936	9341 (2.30-2.26)
Sidechain outliers	177891	9342 (2.30-2.26)
RSRZ outliers	164620	8487 (2.30-2.26)

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 $\leq 5\%$. 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	A	125	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>18%</div> <div>• •</div> </div> </div>
1	B	125	<div> <div>2%</div> <div> <div></div> <div>70%</div> <div>22%</div> <div>• • •</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3957 atoms, of which 1925 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.

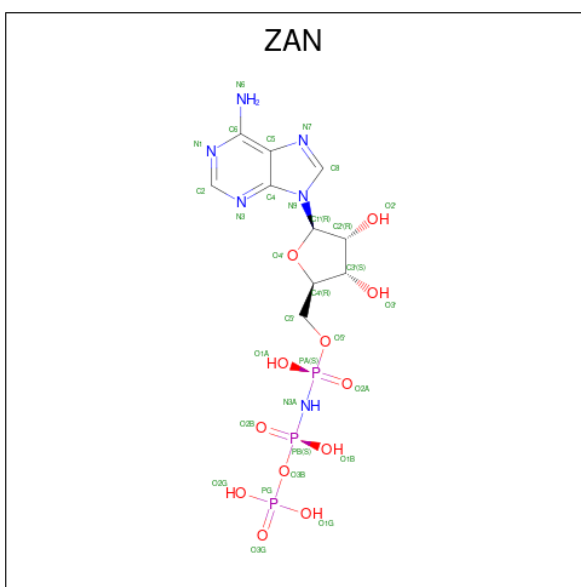
- Molecule 1 is a protein called MCpol domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	120	Total	C	H	N	O	S	0	2	0
			1901	599	952	153	194	3			
1	A	120	Total	C	H	N	O	S	0	1	0
			1896	599	948	153	193	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	120	GLU	-	expression tag	UNP A0A426EXS8
B	121	ASN	-	expression tag	UNP A0A426EXS8
B	122	LEU	-	expression tag	UNP A0A426EXS8
B	123	TYR	-	expression tag	UNP A0A426EXS8
B	124	PHE	-	expression tag	UNP A0A426EXS8
B	125	GLN	-	expression tag	UNP A0A426EXS8
A	120	GLU	-	expression tag	UNP A0A426EXS8
A	121	ASN	-	expression tag	UNP A0A426EXS8
A	122	LEU	-	expression tag	UNP A0A426EXS8
A	123	TYR	-	expression tag	UNP A0A426EXS8
A	124	PHE	-	expression tag	UNP A0A426EXS8
A	125	GLN	-	expression tag	UNP A0A426EXS8

- Molecule 2 is 5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]amino}phosphoryl]adenosine (CCD ID: ZAN) (formula: C₁₀H₁₇N₆O₁₂P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	B	1	Total 44	C 10	H 13	N 6	O 12	P 3	0	0
2	A	1	Total 43	C 10	H 12	N 6	O 12	P 3	0	0

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0

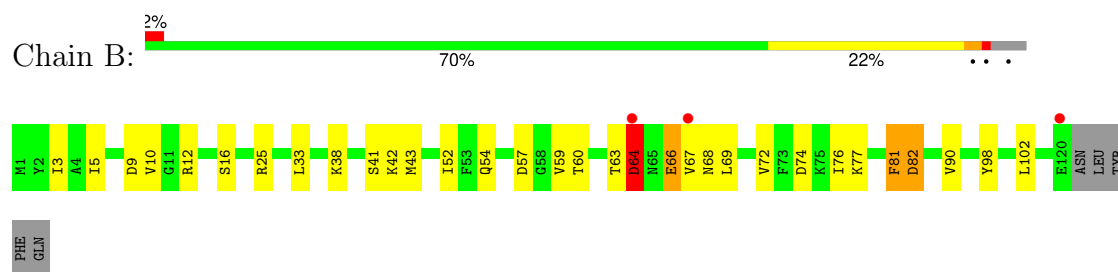
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	34	Total O 34 34	0	0
4	A	37	Total O 37 37	0	0

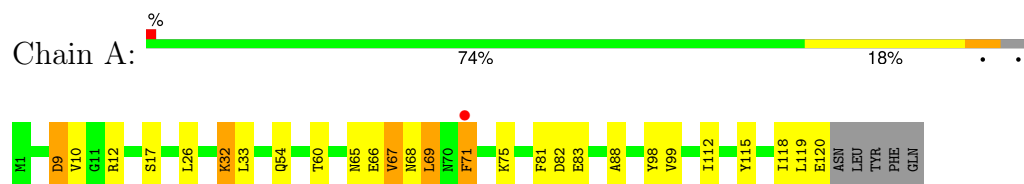
3 Residue-property plots [i](#)

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: MCpol domain-containing protein



- Molecule 1: MCpol domain-containing protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	49.13Å 56.77Å 93.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.55 – 2.28 48.55 – 2.28	Depositor EDS
% Data completeness (in resolution range)	98.2 (48.55-2.28) 98.2 (48.55-2.28)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.99 (at 2.27Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, R_{free}	0.224 , 0.249 0.225 , 0.250	Depositor DCC
R_{free} test set	1232 reflections (9.84%)	wwPDB-VP
Wilson B-factor (Å ²)	35.3	Xtriage
Anisotropy	0.787	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 36.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3957	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.92% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ZAN

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	0.81	4/959 (0.4%)	1.07	4/1289 (0.3%)
1	B	0.82	3/967 (0.3%)	1.10	3/1300 (0.2%)
All	All	0.82	7/1926 (0.4%)	1.09	7/2589 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	B	0	4
All	All	0	8

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	9	ASP	C-N	-9.09	1.23	1.33
1	A	10	VAL	C-N	6.99	1.43	1.33
1	A	32	LYS	C-O	-6.87	1.15	1.24
1	A	33	LEU	C-O	-6.21	1.17	1.24
1	B	66[A]	GLU	N-CA	5.70	1.52	1.45

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	82	ASP	CA-CB-CG	8.64	121.24	112.60
1	A	71	PHE	N-CA-C	-6.82	103.46	111.03
1	A	60	THR	CA-C-O	-5.42	114.89	121.28
1	A	68	ASN	N-CA-C	-5.19	101.31	109.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	64[A]	ASP	N-CA-C	-5.18	106.98	113.15

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	65	ASN	Mainchain
1	A	66[A]	GLU	Mainchain
1	B	63	THR	Mainchain
1	B	64[A]	ASP	Mainchain
1	B	64[B]	ASP	Mainchain

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles [i](#)

5.2.1 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	A	119/125 (95%)	113 (95%)	6 (5%)	0	100	100
1	B	120/125 (96%)	113 (94%)	7 (6%)	0	100	100
All	All	239/250 (96%)	226 (95%)	13 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.2.2 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	A	106/110 (96%)	103 (97%)	3 (3%)	38	52
1	B	107/110 (97%)	105 (98%)	2 (2%)	52	67
All	All	213/220 (97%)	208 (98%)	5 (2%)	44	60

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	67	VAL
1	B	69	LEU
1	A	54	GLN
1	A	67	VAL
1	A	69	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	54	GLN
1	B	65	ASN
1	A	31	ASN
1	A	68	ASN

5.2.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.5 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	ZAN	A	201	3	28,33,33	3.39	10 (35%)	32,52,52	3.03	14 (43%)
2	ZAN	B	201	3	28,33,33	1.31	5 (17%)	32,52,52	2.16	7 (21%)

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	ZAN	A	201	3	-	8/15/38/38	0/3/3/3
2	ZAN	B	201	3	-	3/15/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	201	ZAN	C3'-C4'	-7.42	1.34	1.53
2	A	201	ZAN	O4'-C4'	7.38	1.61	1.45
2	A	201	ZAN	O4'-C1'	-6.58	1.32	1.40
2	A	201	ZAN	PB-O3B	6.09	1.66	1.59
2	A	201	ZAN	PA-N3A	5.15	1.76	1.63

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	201	ZAN	C5-C6-N6	8.81	133.74	120.31
2	B	201	ZAN	O2A-PA-N3A	5.91	120.47	111.77
2	B	201	ZAN	O1B-PB-O3B	5.82	124.08	104.64
2	A	201	ZAN	N3-C2-N1	-5.50	121.20	128.67
2	A	201	ZAN	N6-C6-N1	-5.28	107.05	118.33

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	201	ZAN	PB-N3A-PA-O5'

Continued on next page...

Continued from previous page...

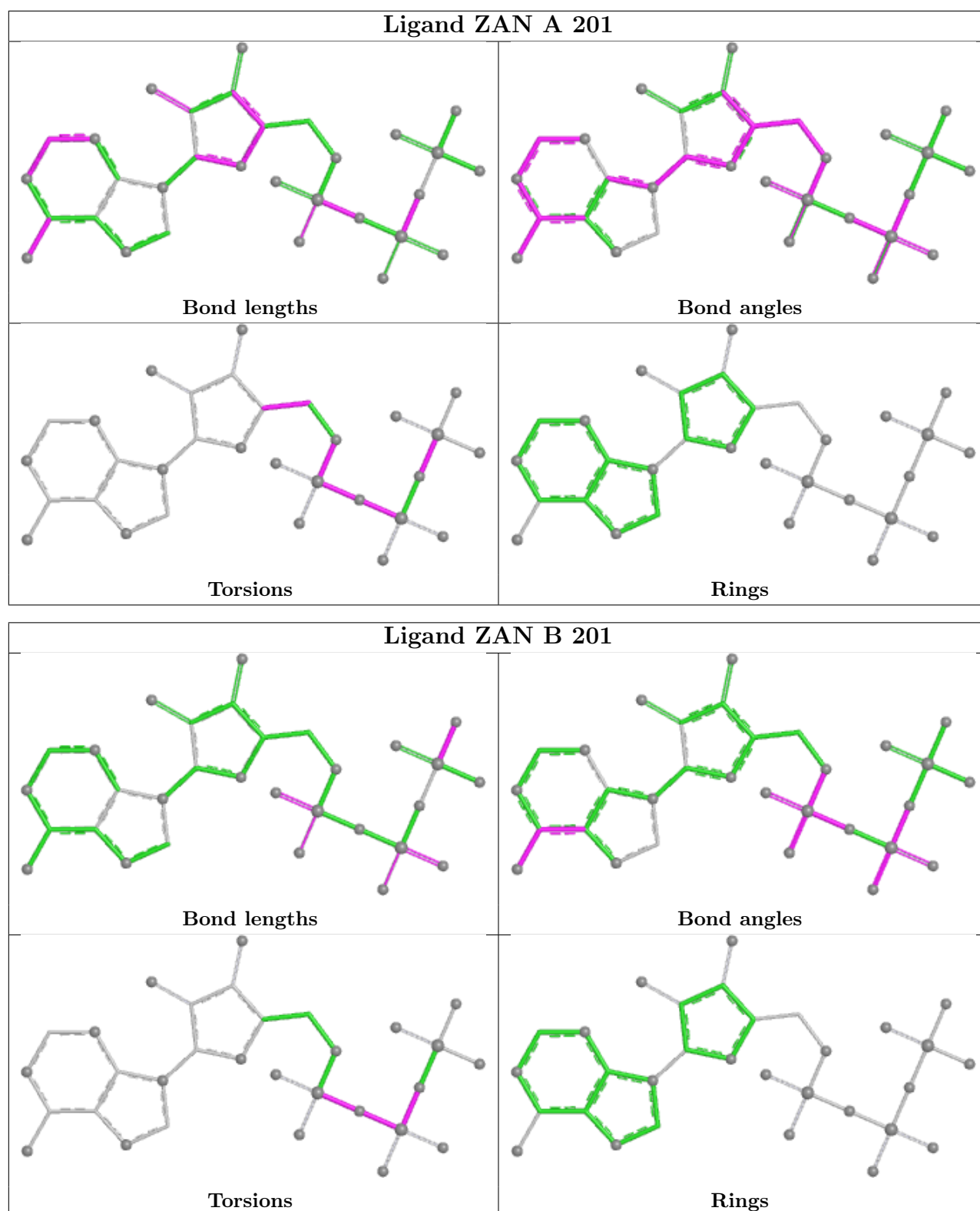
Mol	Chain	Res	Type	Atoms
2	B	201	ZAN	PA-N3A-PB-O3B
2	B	201	ZAN	PG-O3B-PB-O1B
2	A	201	ZAN	PB-N3A-PA-O2A
2	A	201	ZAN	C5'-O5'-PA-O1A

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	201	ZAN	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 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.



5.6 Other polymers ⓘ

There are no such residues in this entry.

5.7 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	120/125 (96%)	0.37	1 (0%) 82 83	37, 49, 75, 97	1 (0%)
1	B	120/125 (96%)	0.42	3 (2%) 58 59	38, 50, 76, 95	2 (1%)
All	All	240/250 (96%)	0.39	4 (1%) 69 70	37, 50, 75, 97	3 (1%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	120	GLU	2.9
1	A	71	PHE	2.6
1	B	64[A]	ASP	2.2
1	B	67	VAL	2.2

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 oligosaccharides 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	MG	B	202	1/1	0.84	0.19	43,43,43,43	0

Continued on next page...

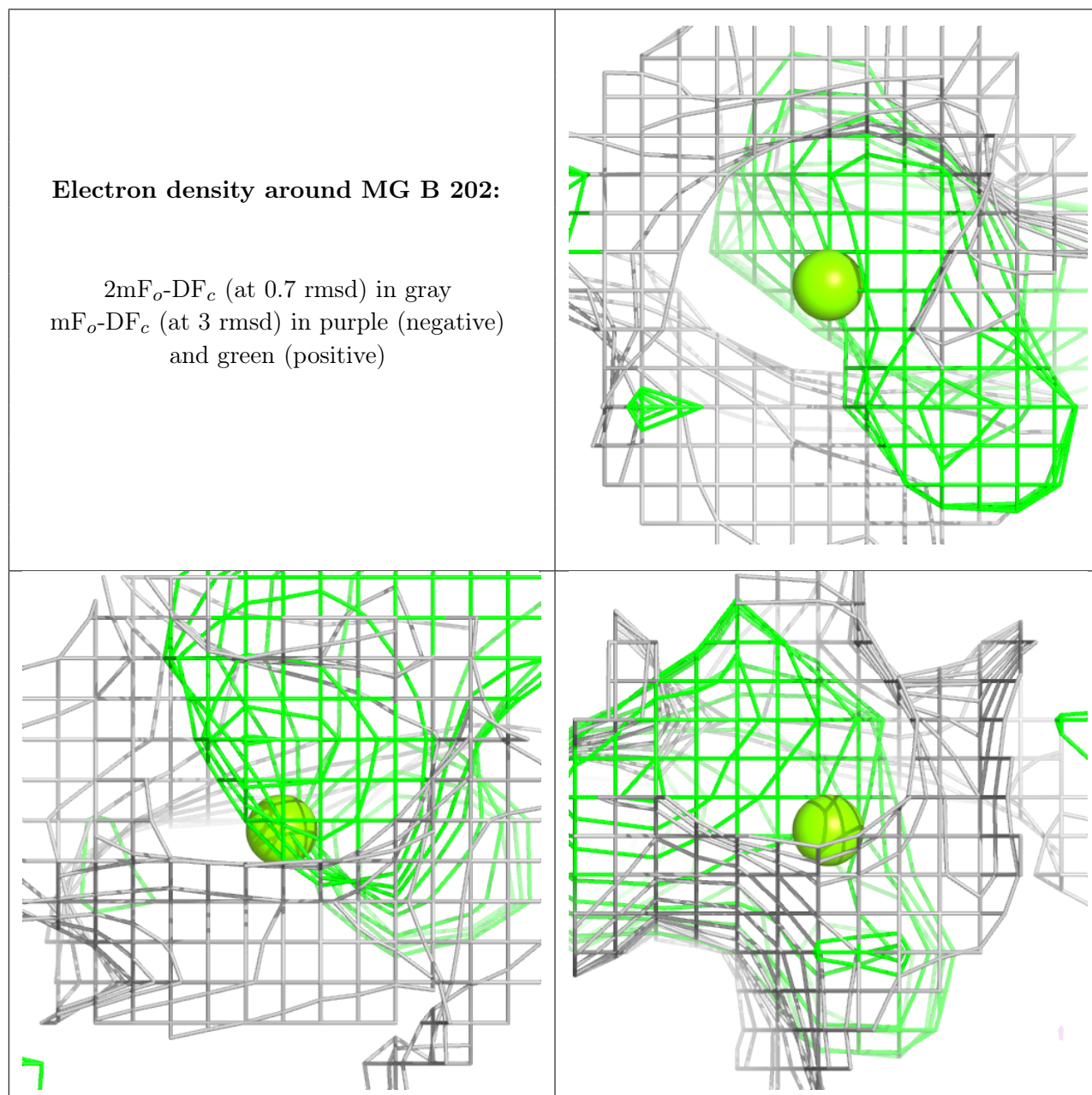
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	ZAN	B	201	31/31	0.86	0.14	34,49,61,69	0
2	ZAN	A	201	31/31	0.95	0.08	27,38,48,52	0
3	MG	A	202	1/1	0.97	0.04	26,26,26,26	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.

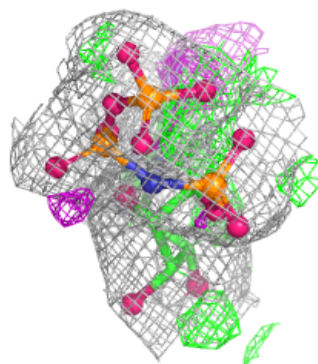
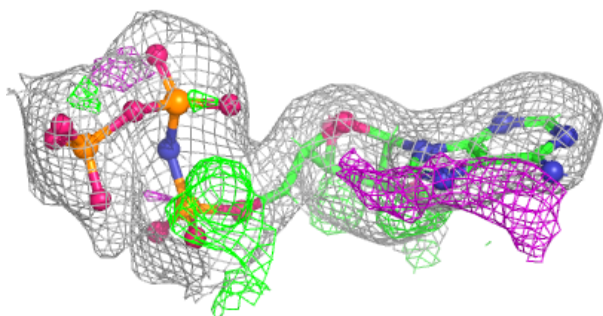
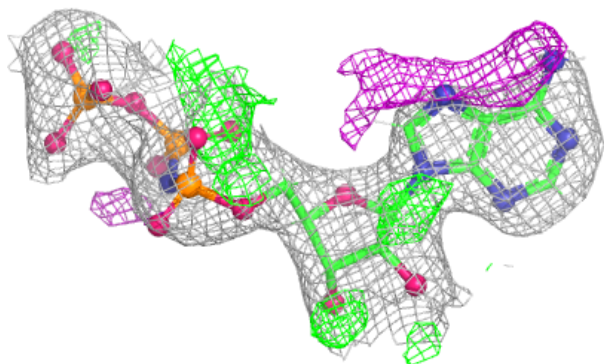
Electron density around MG B 202:

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

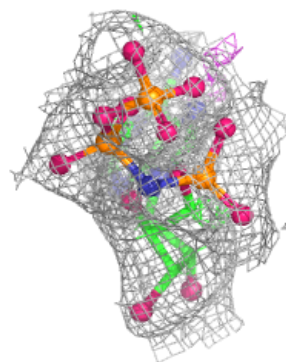
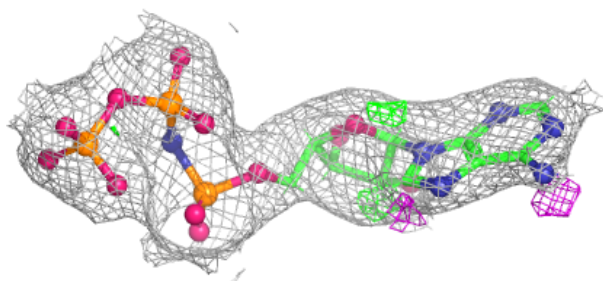
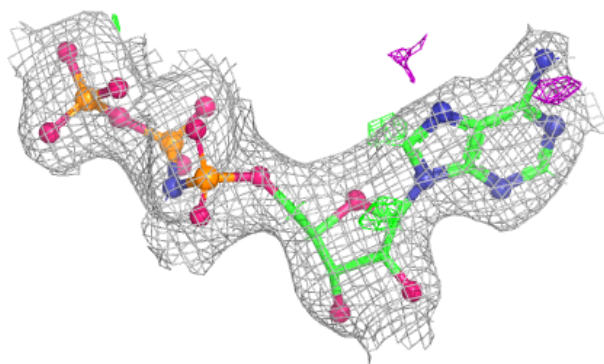


Electron density around ZAN B 201:

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

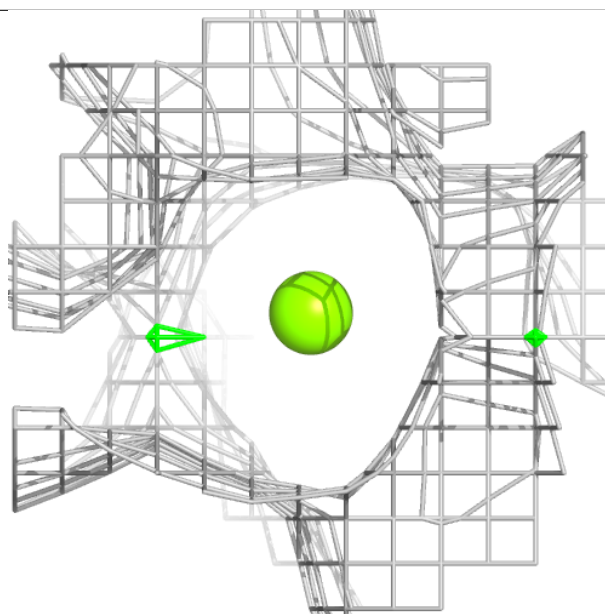
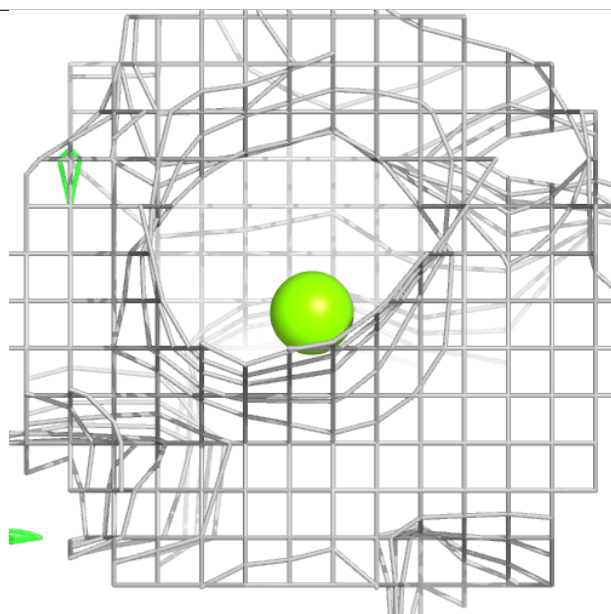
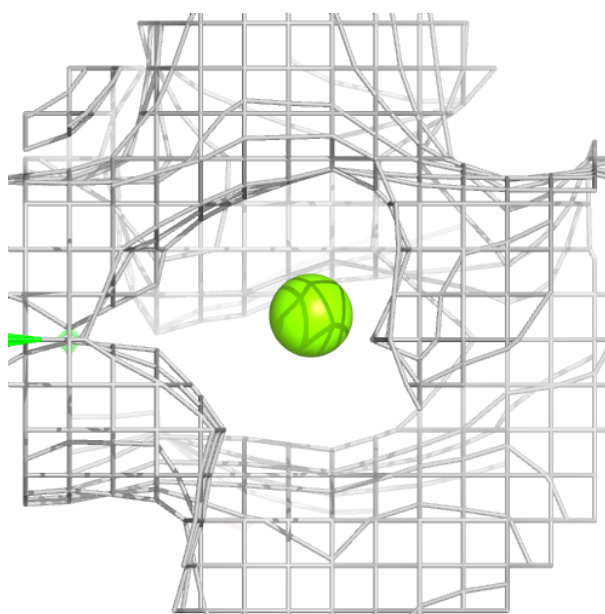
**Electron density around ZAN A 201:**

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



Electron density around MG A 202:

$2mF_o-DF_c$ (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.