



# Full wwPDB X-ray Structure Validation Report ⓘ

May 9, 2021 – 02:43 PM JST

PDB ID : 7DIY  
Title : Crystal structure of SARS-CoV-2 nsp10 bound to nsp14-exoribonuclease domain  
Authors : Lin, S.; Chen, H.; Chen, Z.M.; Yang, F.L.; Ye, F.; Zheng, Y.; Yang, J.; Lin, X.; Sun, H.L.; Wang, L.L.; Wen, A.; Cao, Y.; Lu, G.W.  
Deposited on : 2020-11-19  
Resolution : 2.69 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.18  
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) : 2.18

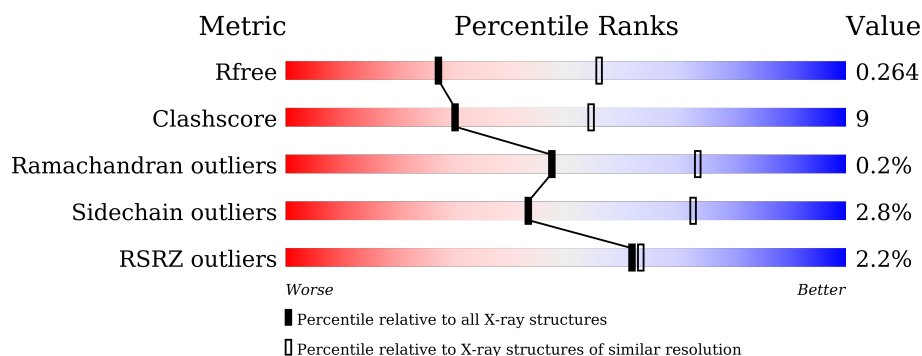
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.69 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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  $\geq 3$ , 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	144	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 21%, green 69%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> <span>%</span> <span>69%</span> <span>21%</span> <span>• 9%</span> </div> </div>
2	B	294	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 2%, yellow 21%, green 75%, grey 21%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> <span>2%</span> <span>75%</span> <span>21%</span> <span>• •</span> </div> </div>

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
3	ZN	A	202	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3249 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 nsp10 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	131	Total	C	N	O	S	0	0	0
			955	593	160	186	16			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P0DTD1
A	-3	PRO	-	expression tag	UNP P0DTD1
A	-2	LEU	-	expression tag	UNP P0DTD1
A	-1	GLY	-	expression tag	UNP P0DTD1
A	0	SER	-	expression tag	UNP P0DTD1

- Molecule 2 is a protein called nsp14-ExoN protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	286	Total	C	N	O	S	0	0	0
			2260	1444	391	403	22			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	GLY	-	expression tag	UNP P0DTD1
B	-3	PRO	-	expression tag	UNP P0DTD1
B	-2	ALA	-	expression tag	UNP P0DTD1
B	-1	GLY	-	expression tag	UNP P0DTD1
B	0	SER	-	expression tag	UNP P0DTD1

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total 2	Zn 2	0	0
3	B	2	Total 2	Zn 2	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 1	Mg 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	7	Total 7	O 7	0	0
5	B	22	Total 22	O 22	0	0



- Molecule 1: nsp10 protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.93Å 68.28Å 109.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.77 – 2.69 42.77 – 2.69	Depositor EDS
% Data completeness (in resolution range)	98.1 (42.77-2.69) 98.1 (42.77-2.69)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.38 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.230 , 0.264 0.230 , 0.264	Depositor DCC
$R_{free}$ test set	563 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.8	Xtriage
Anisotropy	0.737	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 24.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3249	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: ZN, MG

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	0.48	1/976 (0.1%)	0.55	0/1327
2	B	0.41	0/2322	0.57	1/3156 (0.0%)
All	All	0.43	1/3298 (0.0%)	0.56	1/4483 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	84	PRO	N-CD	-7.61	1.37	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	97	THR	N-CA-C	6.04	127.30	111.00

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	A	955	0	911	23	0
2	B	2260	0	2214	43	1
3	A	2	0	0	2	0
3	B	2	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	A	7	0	0	0	0
5	B	22	0	0	4	0
All	All	3249	0	3125	59	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 9.

All (59) 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:A:128:CYS:HG	3:A:202:ZN:ZN	0.80	0.95
2:B:89:PHE:O	5:B:401:HOH:O	1.99	0.80
1:A:6:GLU:HG3	1:A:7:VAL:H	1.48	0.79
2:B:15:ILE:HG23	2:B:16:THR:HG23	1.64	0.78
1:A:128:CYS:SG	3:A:202:ZN:ZN	1.75	0.75
1:A:6:GLU:OE2	1:A:10:ASN:ND2	2.16	0.73
2:B:184:VAL:HG12	2:B:277:THR:HG22	1.72	0.71
2:B:60:PHE:HB3	2:B:62:MET:HE2	1.80	0.64
2:B:32:LYS:O	5:B:404:HOH:O	2.15	0.64
2:B:107:LEU:HD22	2:B:153:MET:HE2	1.78	0.63
1:A:39:THR:O	1:A:78:ARG:NH2	2.33	0.61
1:A:115:THR:HG23	1:A:124:LYS:HD3	1.83	0.61
2:B:76:ARG:O	2:B:80:ILE:HG12	2.01	0.60
1:A:6:GLU:HG3	1:A:7:VAL:N	2.16	0.60
1:A:5:THR:O	2:B:25:THR:HG21	2.02	0.60
2:B:181:VAL:HG22	2:B:227:TRP:HZ2	1.68	0.58
2:B:181:VAL:HG22	2:B:227:TRP:CZ2	2.39	0.58
2:B:208:CYS:O	2:B:209:LEU:HB2	2.05	0.57
1:A:128:CYS:SG	1:A:130:CYS:SG	3.02	0.57
1:A:21:VAL:HG22	2:B:200:LYS:HB3	1.86	0.56
2:B:92:GLU:OE2	5:B:405:HOH:O	2.18	0.55
1:A:80:HIS:CE1	2:B:55:ILE:HG21	2.42	0.54
1:A:57:VAL:HG13	1:A:58:THR:HG23	1.90	0.53
2:B:267:ALA:O	2:B:269:VAL:N	2.43	0.52
2:B:210:CYS:HB3	2:B:212:ARG:H	1.75	0.52
1:A:126:TYR:N	1:A:127:GLY:HA2	2.26	0.50
2:B:3:ASN:OD1	5:B:406:HOH:O	2.20	0.50
2:B:120:VAL:H	2:B:157:LEU:H	1.59	0.50
2:B:184:VAL:HG13	2:B:242:ILE:HD11	1.94	0.49
2:B:207:CYS:SG	2:B:210:CYS:HB2	2.53	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:40:ASN:OD1	2:B:25:THR:OG1	2.30	0.49
2:B:113:THR:O	2:B:115:VAL:N	2.46	0.49
2:B:267:ALA:O	2:B:269:VAL:HG23	2.13	0.48
2:B:120:VAL:N	2:B:157:LEU:H	2.11	0.48
2:B:17:GLY:HA3	2:B:53:ARG:NH2	2.28	0.48
2:B:80:ILE:O	2:B:83:VAL:HG23	2.13	0.48
2:B:207:CYS:SG	2:B:208:CYS:O	2.72	0.47
1:A:54:ALA:HB1	1:A:123:TRP:CE2	2.50	0.47
1:A:80:HIS:ND1	2:B:55:ILE:HG21	2.30	0.47
2:B:221:SER:O	2:B:223:THR:N	2.49	0.46
2:B:250:THR:HG23	2:B:251:GLY:N	2.30	0.46
2:B:207:CYS:HB3	2:B:210:CYS:HB2	1.96	0.46
2:B:185:LEU:CD2	2:B:241:MET:HB2	2.46	0.46
1:A:53:GLN:O	1:A:98:GLN:HG2	2.17	0.44
2:B:111:PHE:CD2	2:B:169:MET:HE1	2.53	0.44
1:A:41:CYS:O	2:B:26:HIS:HB2	2.17	0.44
1:A:12:THR:OG1	2:B:63:ASN:OD1	2.16	0.43
2:B:214:ALA:HB1	2:B:224:TYR:HB3	2.00	0.43
2:B:260:TYR:HB3	2:B:283:HIS:CD2	2.52	0.43
1:A:74:CYS:HB3	1:A:77:CYS:HB2	2.02	0.42
2:B:14:VAL:HG12	2:B:16:THR:H	1.85	0.42
1:A:117:CYS:SG	1:A:119:VAL:HB	2.59	0.41
2:B:215:THR:HG21	2:B:227:TRP:CD2	2.56	0.41
1:A:85:ASN:OD1	1:A:85:ASN:N	2.54	0.41
2:B:69:TYR:HA	2:B:70:PRO:HD3	1.91	0.41
2:B:82:HIS:CD2	2:B:180:ARG:HD2	2.55	0.41
1:A:65:GLN:HE21	1:A:65:GLN:HB3	1.72	0.41
2:B:88:GLY:C	2:B:274:ALA:HB1	2.42	0.41
2:B:141:PRO:HB2	2:B:146:PHE:CD2	2.57	0.40

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)
2:B:155:LYS:O	2:B:221:SER:OG[4_455]	2.15	0.05

## 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	A	129/144 (90%)	123 (95%)	6 (5%)	0	100	100
2	B	284/294 (97%)	264 (93%)	19 (7%)	1 (0%)	34	60
All	All	413/438 (94%)	387 (94%)	25 (6%)	1 (0%)	47	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	114	GLY

### 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	A	105/116 (90%)	103 (98%)	2 (2%)	57	82
2	B	251/255 (98%)	243 (97%)	8 (3%)	39	68
All	All	356/371 (96%)	346 (97%)	10 (3%)	43	73

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	62	ASN
1	A	85	ASN
2	B	57	MET
2	B	118	VAL
2	B	122	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	174	LEU
2	B	181	VAL
2	B	215	THR
2	B	223	THR
2	B	261	CYS

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	80	HIS
1	A	98	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 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 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.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	131/144 (90%)	0.11	2 (1%) 73 76	33, 40, 50, 62	0
2	B	286/294 (97%)	0.10	7 (2%) 59 60	33, 38, 45, 49	0
All	All	417/438 (95%)	0.10	9 (2%) 62 63	33, 38, 47, 62	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	268	HIS	4.2
1	A	86	PRO	3.3
1	A	89	PHE	3.2
2	B	209	LEU	3.1
2	B	267	ALA	2.9
2	B	250	THR	2.5
2	B	155	LYS	2.2
2	B	146	PHE	2.1
2	B	266	ASN	2.1

### 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.

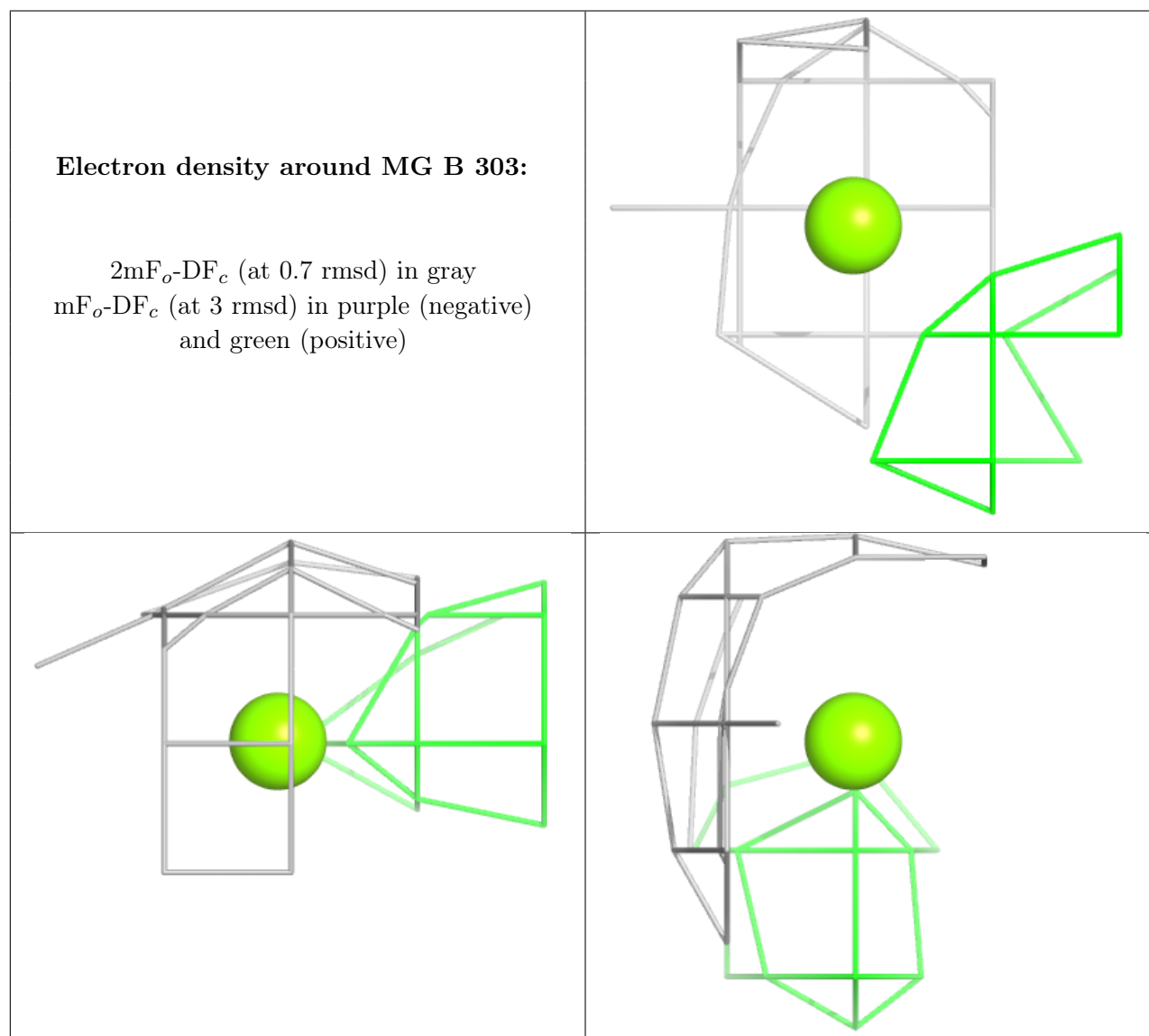
### 6.4 Ligands [i](#)

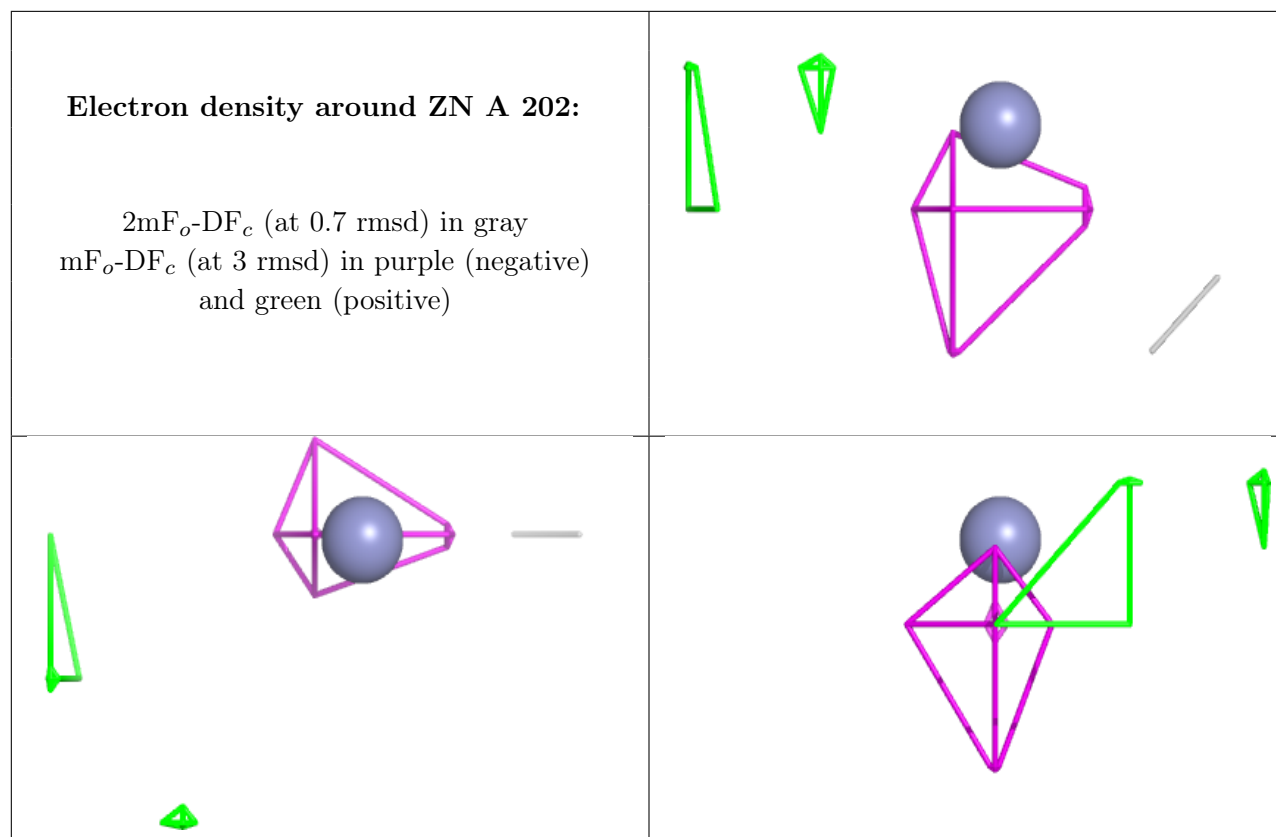
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	MG	B	303	1/1	0.83	0.10	36,36,36,36	0
3	ZN	A	202	1/1	0.94	0.05	52,52,52,52	0
3	ZN	B	302	1/1	0.98	0.16	40,40,40,40	0
3	ZN	A	201	1/1	0.99	0.08	36,36,36,36	0
3	ZN	B	301	1/1	0.99	0.07	38,38,38,38	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.

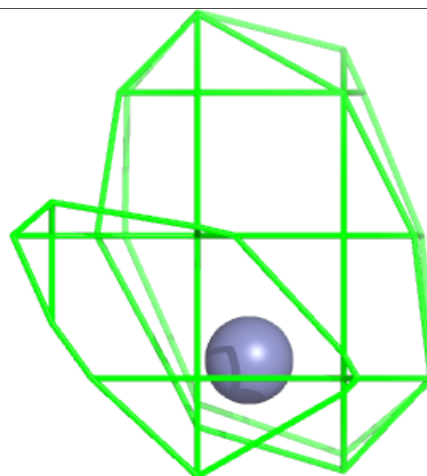
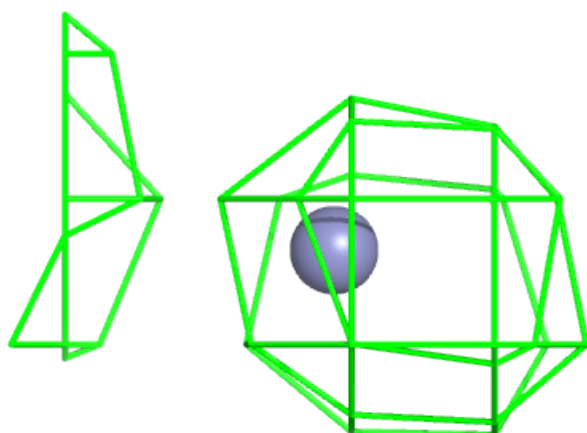
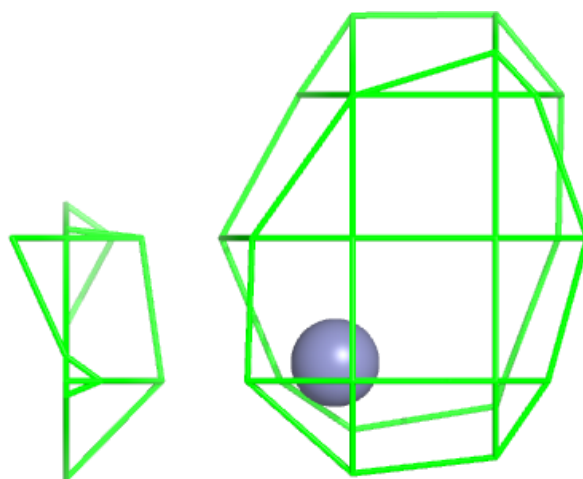






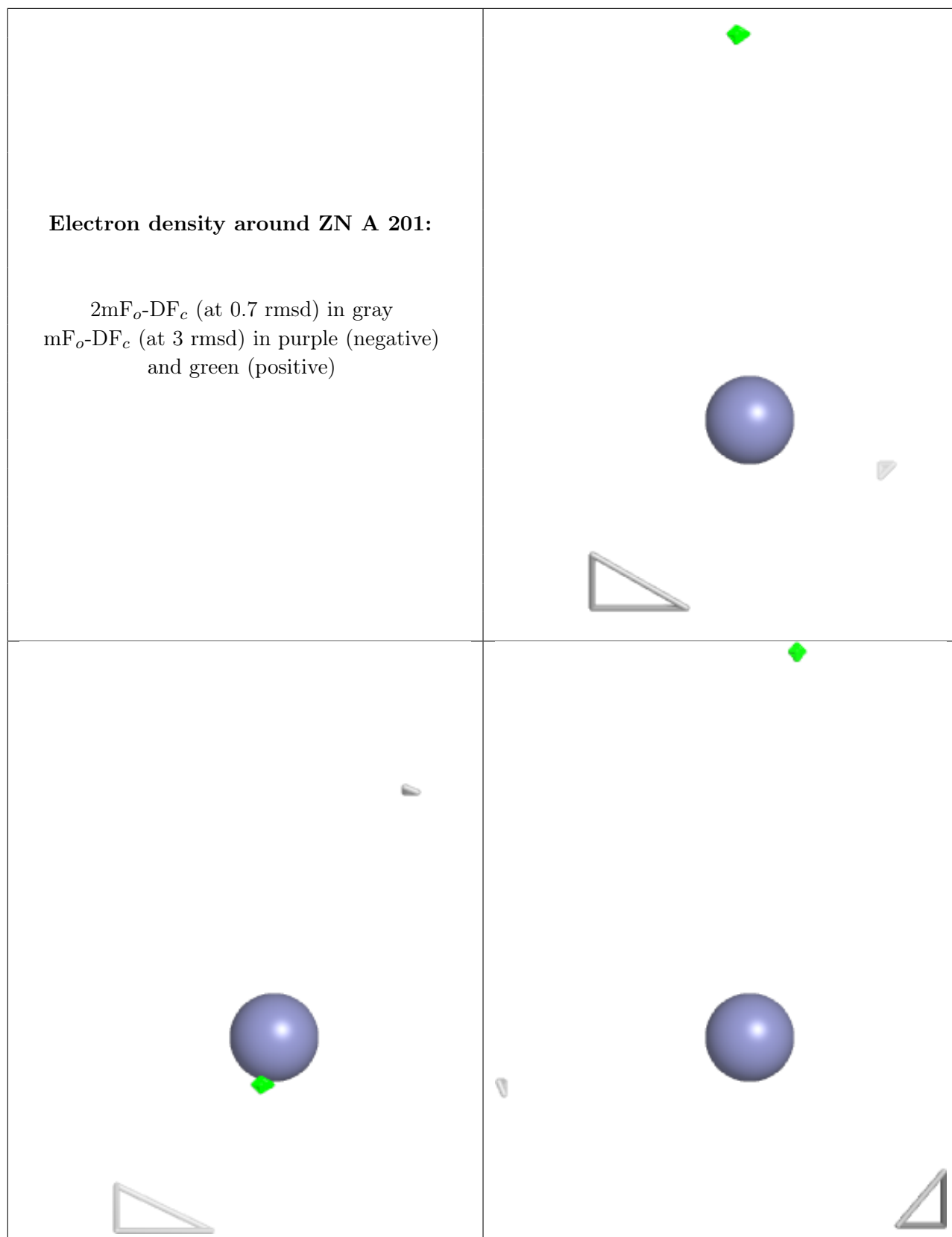
**Electron density around ZN B 302:**

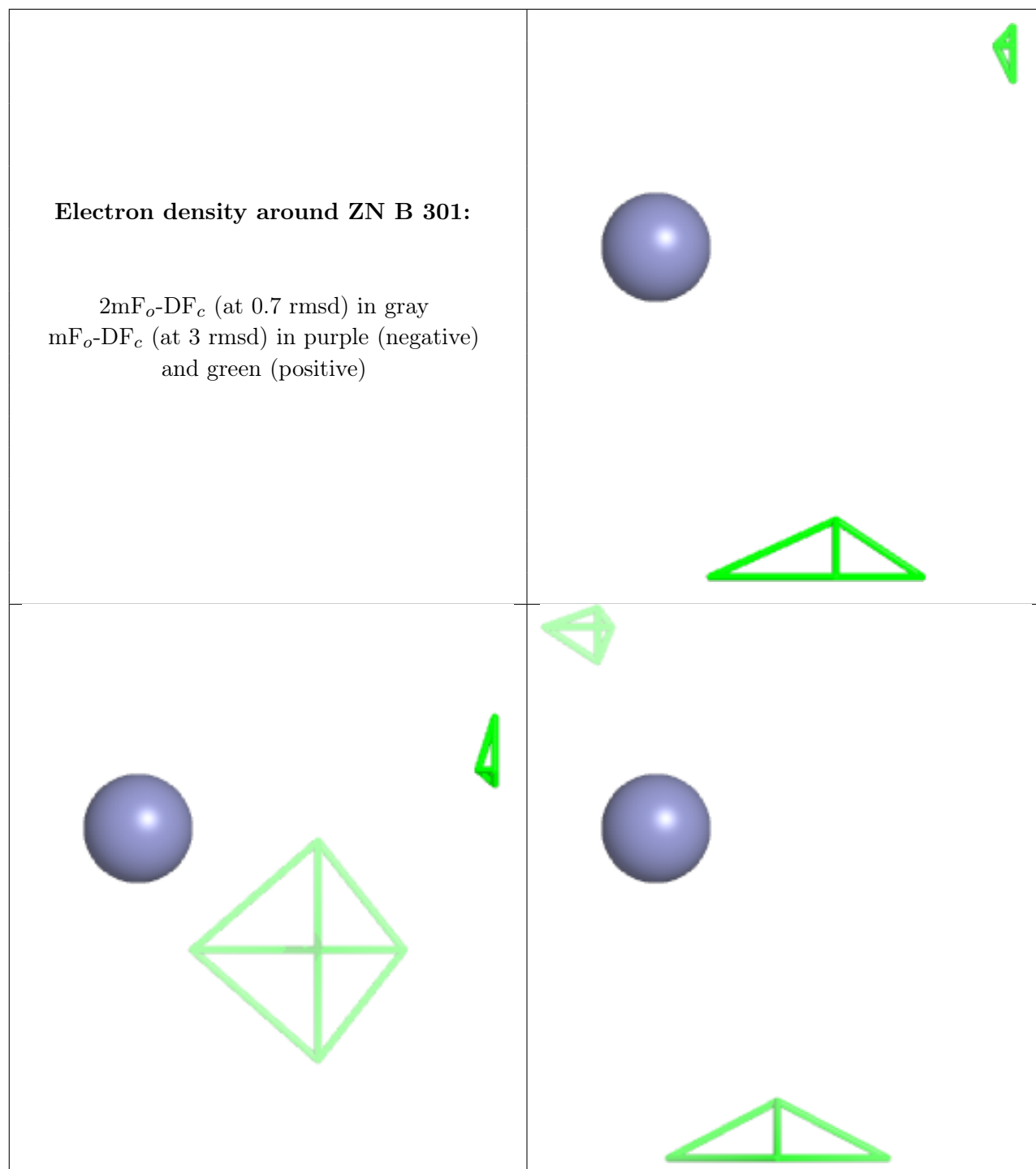
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around ZN 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)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.