Full wwPDB X-ray Structure Validation Report ⓘ

Jan 11, 2024 – 03:53 PM EST

PDB ID	:	8UQ9
Title	:	Crystal structure of RNF168 (RING)-UbcH5c fused to H2A-H2B via a 4-residue linker
Authors	:	Hu, Q.; Botuyan, M.V.; Mer, G.
Deposited on	:	2023-10-23
Resolution	:	2.30 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:

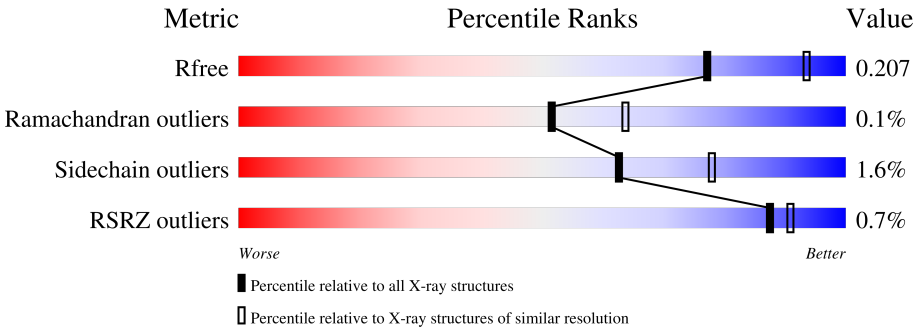
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.30 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	5042 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	439	% 97%
1	a	439	97%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7314 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 E3 ubiquitin-protein ligase RNF168,Ubiquitin-conjugating enzyme E2 D3,Histone H2B type 2-E,Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	432	Total 3370	C 2127	N 606	O 616	S 21	17	1	0
1	a	433	Total 3385	C 2133	N 614	O 617	S 21	5	2	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q8IYW5
A	0	HIS	-	expression tag	UNP Q8IYW5
A	95	SER	-	linker	UNP Q8IYW5
A	96	GLY	-	linker	UNP Q8IYW5
A	97	SER	-	linker	UNP Q8IYW5
A	98	GLY	-	linker	UNP Q8IYW5
A	99	SER	-	linker	UNP Q8IYW5
A	100	GLY	-	linker	UNP Q8IYW5
A	101	SER	-	linker	UNP Q8IYW5
A	1148	GLY	-	linker	UNP P61077
A	1149	SER	-	linker	UNP P61077
A	1150	GLY	-	linker	UNP P61077
A	1151	GLY	-	linker	UNP P61077
A	3011	SER	-	linker	UNP Q16778
a	-1	GLY	-	expression tag	UNP Q8IYW5
a	0	HIS	-	expression tag	UNP Q8IYW5
a	95	SER	-	linker	UNP Q8IYW5
a	96	GLY	-	linker	UNP Q8IYW5
a	97	SER	-	linker	UNP Q8IYW5
a	98	GLY	-	linker	UNP Q8IYW5
a	99	SER	-	linker	UNP Q8IYW5
a	100	GLY	-	linker	UNP Q8IYW5
a	101	SER	-	linker	UNP Q8IYW5
a	1148	GLY	-	linker	UNP P61077

Continued on next page...

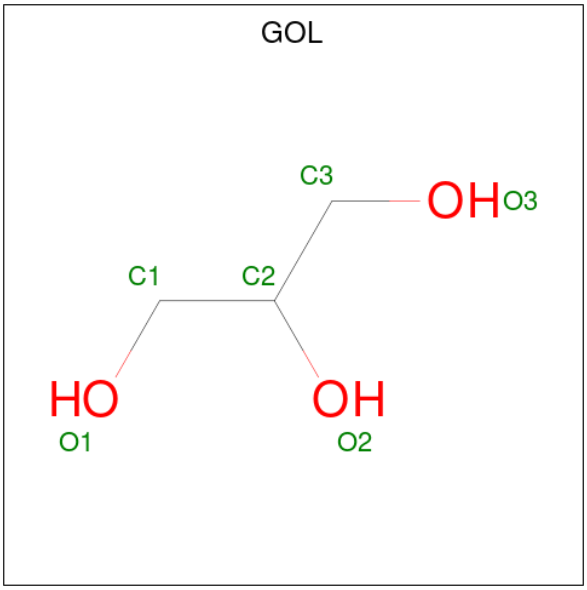
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
a	1149	SER	-	linker	UNP P61077
a	1150	GLY	-	linker	UNP P61077
a	1151	GLY	-	linker	UNP P61077
a	3011	SER	-	linker	UNP Q16778

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	17	Total Cl 17 17	0	0
2	a	19	Total Cl 19 19	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	a	1	Total C O 6 3 3	0	0
3	a	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Zn 2 2	0	0
4	a	2	Total Zn 2 2	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Na 1 1	0	0
5	a	3	Total Na 3 3	0	0

- Molecule 6 is water.

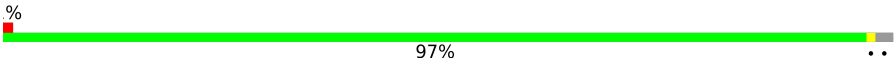
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	235	Total O 235 235	0	0
6	a	244	Total O 244 244	0	0

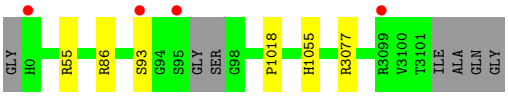
3 Residue-property plots (i)

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: E3 ubiquitin-protein ligase RNF168,Ubiquitin-conjugating enzyme E2 D3,Histone H2B type 2-E,Histone H2A type 1-B/E

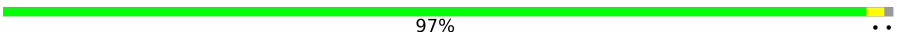
Chain A:





• Molecule 1: E3 ubiquitin-protein ligase RNF168,Ubiquitin-conjugating enzyme E2 D3,Histone H2B type 2-E,Histone H2A type 1-B/E

Chain a:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	108.14Å 108.14Å 114.01Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.18 – 2.30 36.18 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (36.18-2.30) 99.8 (36.18-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.56 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.183 , 0.206 0.188 , 0.207	Depositor DCC
R_{free} test set	2033 reflections (3.08%)	wwPDB-VP
Wilson B-factor (Å^2)	33.2	Xtriage
Anisotropy	0.721	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 31.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.488 for -h,-k,l 0.039 for h,-h-k,-l 0.039 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7314	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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: GOL, ZN, NA, CL

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	0.25	0/3447	0.49	0/4672
1	a	0.25	0/3466	0.50	0/4699
All	All	0.25	0/6913	0.50	0/9371

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

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	429/439 (98%)	419 (98%)	9 (2%)	1 (0%)	47	58
1	a	433/439 (99%)	426 (98%)	7 (2%)	0	100	100
All	All	862/878 (98%)	845 (98%)	16 (2%)	1 (0%)	51	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1018	PRO

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	368/377 (98%)	363 (99%)	5 (1%)	67 81
1	a	371/377 (98%)	364 (98%)	7 (2%)	57 73
All	All	739/754 (98%)	727 (98%)	12 (2%)	62 78

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	55	ARG
1	A	86	ARG
1	A	93	SER
1	A	1055	HIS
1	A	3077	ARG
1	a	93	SER
1	a	1022	SER
1	a	1038	MET
1	a	2120	LYS
1	a	3013	LYS
1	a	3073	ASN
1	a	3084	GLN

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	3082	HIS
1	a	2084	ASN

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 50 ligands modelled in this entry, 44 are monoatomic - leaving 6 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	a	4011	-	5,5,5	0.87	0	5,5,5	1.01	0
3	GOL	A	4007	-	5,5,5	0.89	0	5,5,5	0.98	0
3	GOL	A	4009	-	5,5,5	0.87	0	5,5,5	1.02	0
3	GOL	a	4010	-	5,5,5	0.90	0	5,5,5	0.96	0
3	GOL	a	4024	-	5,5,5	0.88	0	5,5,5	0.95	0
3	GOL	A	4008	-	5,5,5	0.88	0	5,5,5	0.98	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	a	4011	-	-	2/4/4/4	-
3	GOL	A	4007	-	-	2/4/4/4	-
3	GOL	A	4009	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	a	4010	-	-	0/4/4/4	-
3	GOL	a	4024	-	-	2/4/4/4	-
3	GOL	A	4008	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	4009	GOL	C1-C2-C3-O3
3	a	4011	GOL	O1-C1-C2-C3
3	A	4007	GOL	O1-C1-C2-C3
3	a	4024	GOL	O1-C1-C2-C3
3	A	4007	GOL	O1-C1-C2-O2
3	a	4011	GOL	O1-C1-C2-O2
3	A	4009	GOL	O2-C2-C3-O3
3	a	4024	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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	432/439 (98%)	-0.26	4 (0%)	84	88	23, 39, 68, 100	8 (1%)
1	a	433/439 (98%)	-0.27	2 (0%)	91	94	24, 38, 69, 101	1 (0%)
All	All	865/878 (98%)	-0.26	6 (0%)	87	91	23, 39, 69, 101	9 (1%)

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	a	93	SER	3.2
1	a	3100	VAL	2.8
1	A	93	SER	2.7
1	A	0	HIS	2.1
1	A	95	SER	2.1
1	A	3099	ARG	2.0

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, 95^{th} 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
2	CL	a	4004	1/1	0.81	0.09	84,84,84,84	0
2	CL	a	4015	1/1	0.82	0.08	73,73,73,73	0
3	GOL	a	4010	6/6	0.84	0.16	48,57,61,64	0
2	CL	a	4016	1/1	0.86	0.22	61,61,61,61	0
3	GOL	a	4011	6/6	0.88	0.18	39,53,55,63	0
5	NA	a	4019	1/1	0.88	0.14	43,43,43,43	0
5	NA	a	4023	1/1	0.88	0.07	44,44,44,44	0
3	GOL	A	4008	6/6	0.90	0.13	58,62,69,73	0
3	GOL	A	4007	6/6	0.91	0.19	51,56,60,65	0
2	CL	A	4013	1/1	0.91	0.16	63,63,63,63	0
2	CL	A	4014	1/1	0.92	0.11	62,62,62,62	0
3	GOL	A	4009	6/6	0.93	0.16	38,49,55,59	0
3	GOL	a	4024	6/6	0.93	0.15	37,46,49,58	0
2	CL	a	4005	1/1	0.94	0.06	54,54,54,54	0
2	CL	a	4014	1/1	0.94	0.16	62,62,62,62	0
2	CL	a	4003	1/1	0.95	0.09	51,51,51,51	0
2	CL	A	4003	1/1	0.95	0.07	54,54,54,54	0
2	CL	A	4017	1/1	0.95	0.05	68,68,68,68	0
2	CL	a	4020	1/1	0.95	0.06	58,58,58,58	0
2	CL	a	4018	1/1	0.96	0.14	68,68,68,68	0
2	CL	a	4008	1/1	0.96	0.11	75,75,75,75	0
2	CL	A	4021	1/1	0.96	0.18	61,61,61,61	0
5	NA	A	4023	1/1	0.96	0.09	50,50,50,50	0
2	CL	A	4018	1/1	0.96	0.10	65,65,65,65	0
2	CL	a	4006	1/1	0.96	0.15	57,57,57,57	0
2	CL	A	4012	1/1	0.97	0.14	59,59,59,59	0
2	CL	A	4004	1/1	0.98	0.09	66,66,66,66	0
2	CL	a	4021	1/1	0.98	0.19	61,61,61,61	0
2	CL	a	4025	1/1	0.98	0.06	30,30,30,30	0
2	CL	A	4019	1/1	0.98	0.07	54,54,54,54	0
5	NA	a	4009	1/1	0.98	0.13	35,35,35,35	0
2	CL	a	4017	1/1	0.98	0.09	69,69,69,69	0
2	CL	A	4005	1/1	0.98	0.09	54,54,54,54	0
2	CL	a	4027	1/1	0.99	0.09	56,56,56,56	0
2	CL	A	4022	1/1	0.99	0.09	49,49,49,49	0
2	CL	a	4002	1/1	0.99	0.09	41,41,41,41	0
2	CL	A	4006	1/1	0.99	0.07	49,49,49,49	0
2	CL	A	4002	1/1	0.99	0.07	40,40,40,40	0
2	CL	A	4015	1/1	0.99	0.06	41,41,41,41	0
2	CL	A	4020	1/1	0.99	0.09	51,51,51,51	0
4	ZN	A	4010	1/1	0.99	0.14	32,32,32,32	0
4	ZN	A	4011	1/1	0.99	0.16	30,30,30,30	0
4	ZN	a	4012	1/1	0.99	0.14	32,32,32,32	0

Continued on next page...

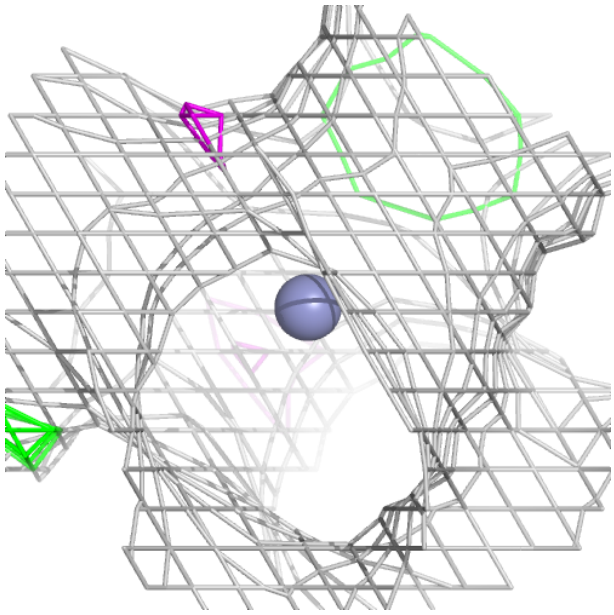
Continued from previous page...

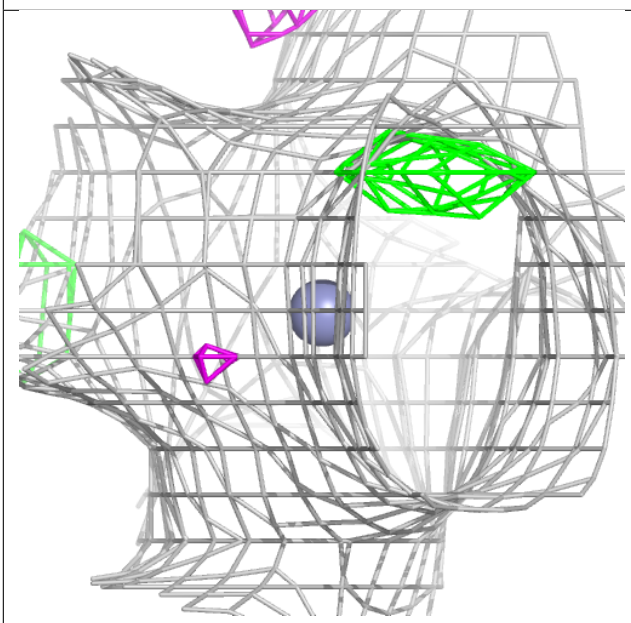
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CL	a	4007	1/1	0.99	0.15	47,47,47,47	0
2	CL	a	4022	1/1	0.99	0.11	65,65,65,65	0
2	CL	A	4016	1/1	0.99	0.06	30,30,30,30	0
2	CL	a	4026	1/1	0.99	0.08	50,50,50,50	0
2	CL	A	4001	1/1	1.00	0.07	32,32,32,32	0
4	ZN	a	4013	1/1	1.00	0.16	30,30,30,30	0
2	CL	a	4001	1/1	1.00	0.07	32,32,32,32	0

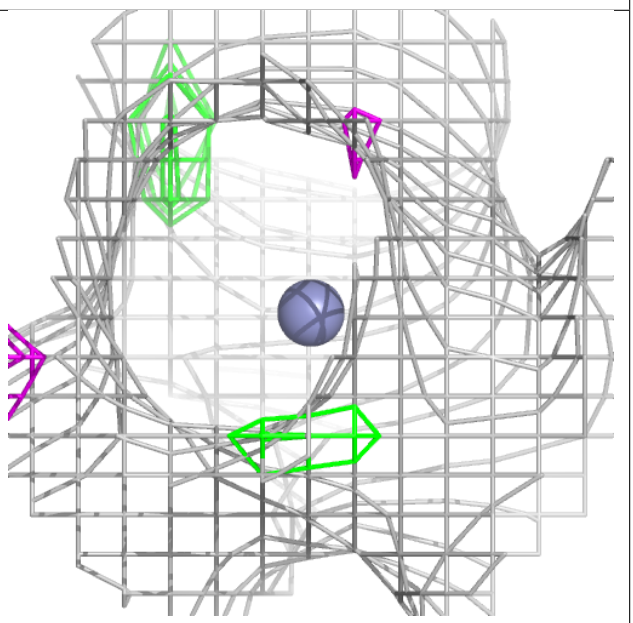
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ZN A 4010:

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

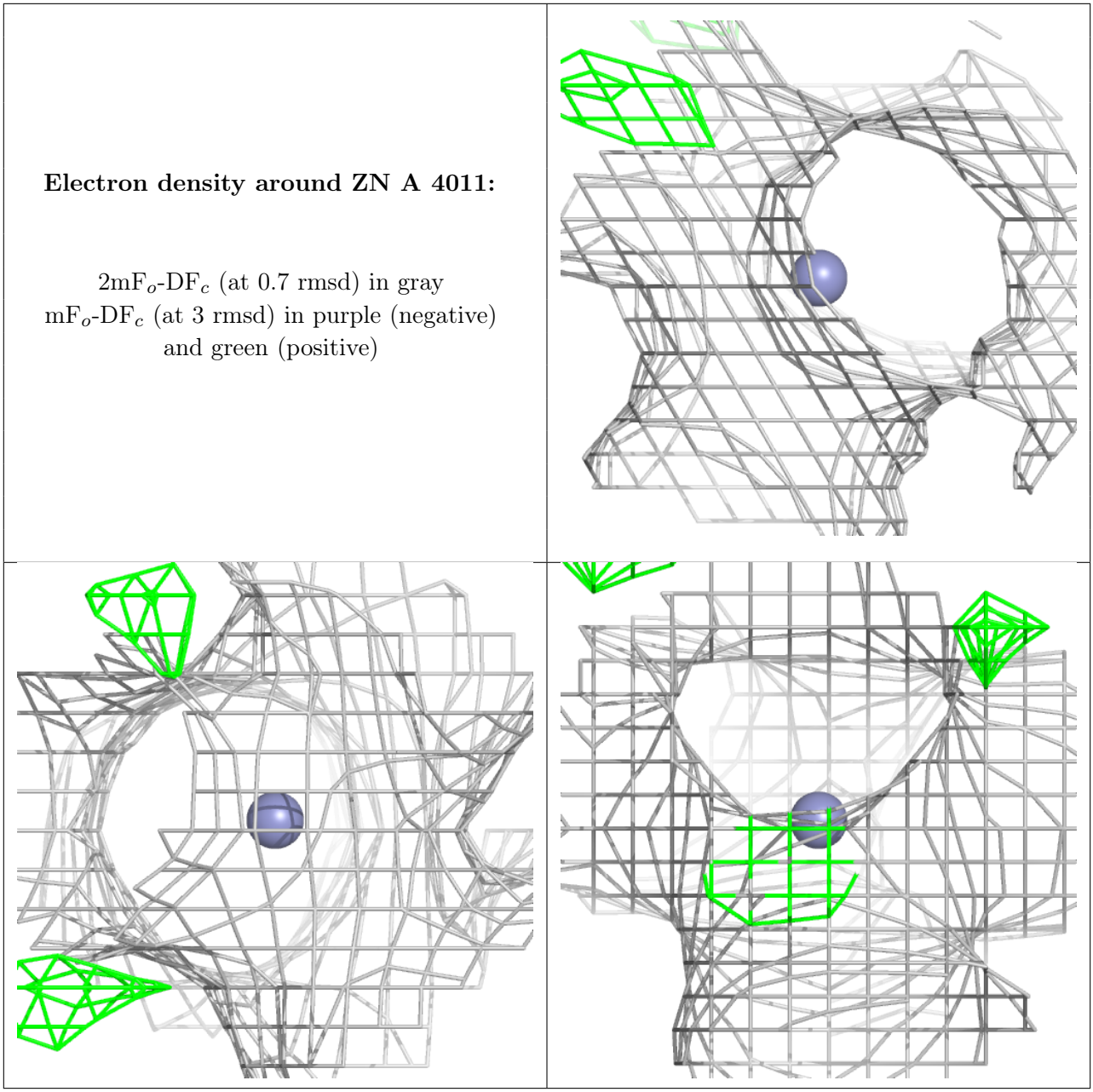






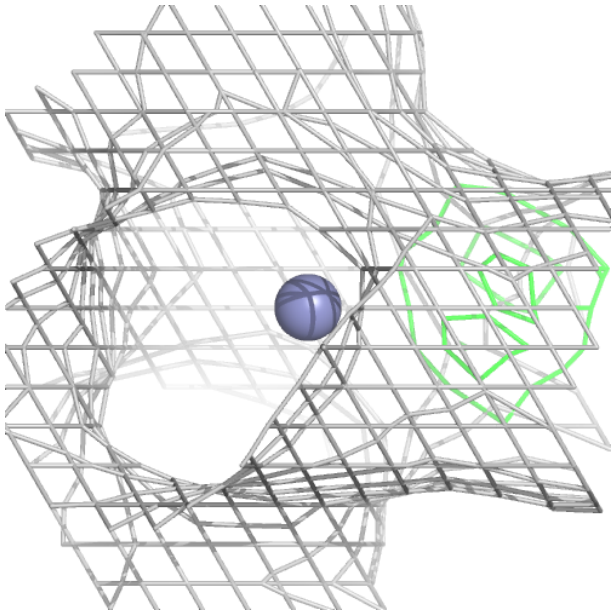
Electron density around ZN A 4011:

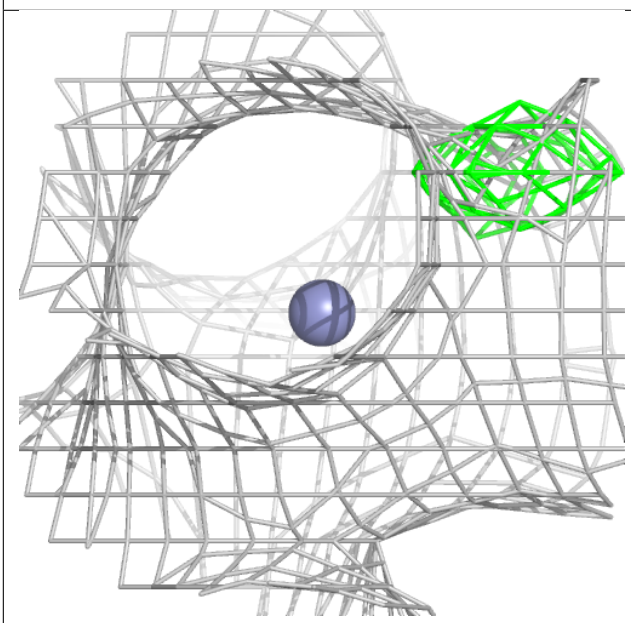
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

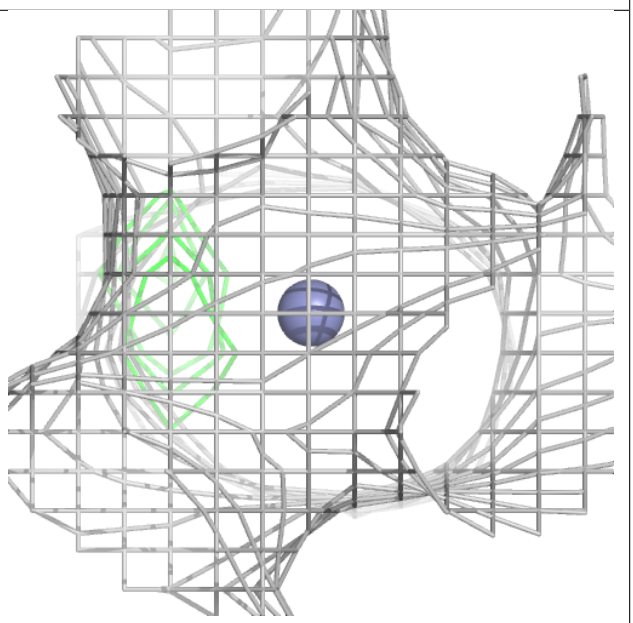


Electron density around ZN a 4012:

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

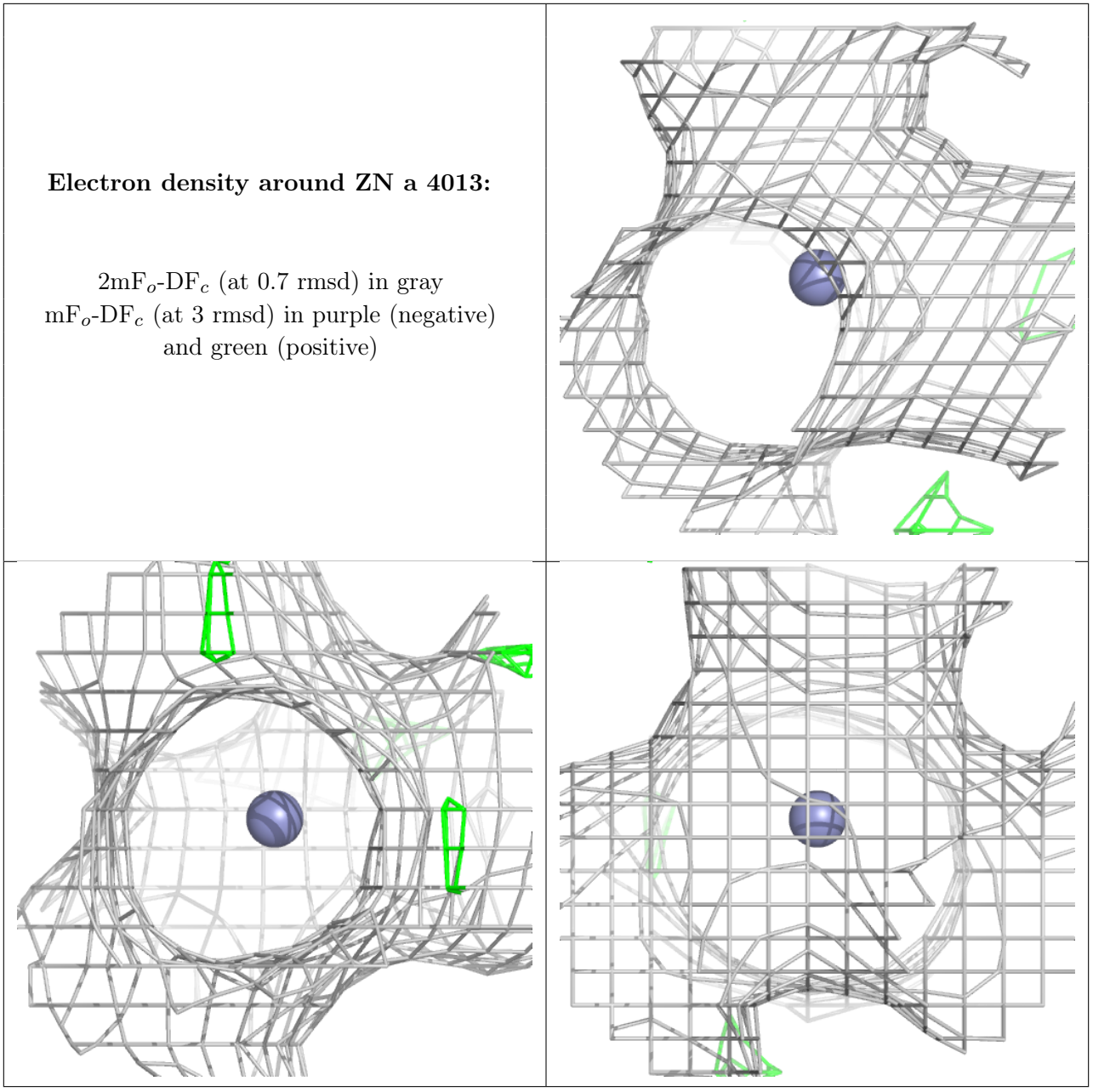






Electron density around ZN a 4013:

$2mF_o$-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.