



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 05:02 PM UTC

PDB ID : 7VSP / pdb_00007vsp
Title : Crystal structure of the tandem B-Box domains of COL2
Authors : Lv, X.; Liu, R.; Du, J.
Deposited on : 2021-10-27
Resolution : 2.10 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

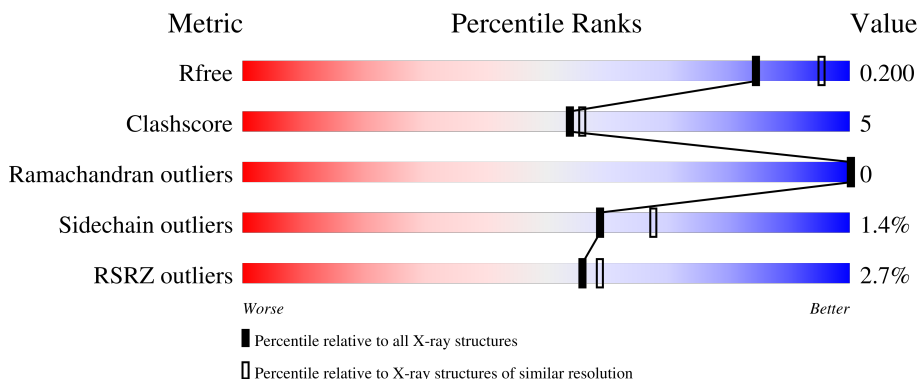
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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	121	0% 69% 5% 26%
1	B	121	2% 65% 7% 26%
1	C	121	3% 71% 5% 24%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2134 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 Zinc finger protein CONSTANS-LIKE 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	90	Total	C	N	O	S	Se	0	0	0
			666	393	133	126	11	3			
1	B	89	Total	C	N	O	S	Se	0	0	0
			657	389	129	125	11	3			
1	C	92	Total	C	N	O	S	Se	0	0	0
			681	404	135	128	11	3			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q96502
A	5	ALA	GLU	conflict	UNP Q96502
A	92	MSE	LEU	conflict	UNP Q96502
A	101	MSE	ILE	conflict	UNP Q96502
A	102	MSE	LEU	conflict	UNP Q96502
B	0	SER	-	expression tag	UNP Q96502
B	5	ALA	GLU	conflict	UNP Q96502
B	92	MSE	LEU	conflict	UNP Q96502
B	101	MSE	ILE	conflict	UNP Q96502
B	102	MSE	LEU	conflict	UNP Q96502
C	0	SER	-	expression tag	UNP Q96502
C	5	ALA	GLU	conflict	UNP Q96502
C	92	MSE	LEU	conflict	UNP Q96502
C	101	MSE	ILE	conflict	UNP Q96502
C	102	MSE	LEU	conflict	UNP Q96502

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

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

Continued on next page...

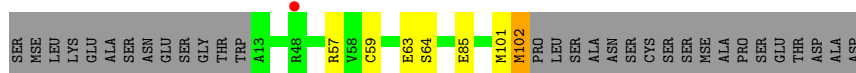
Continued from previous page...

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

- Molecule 3 is water.

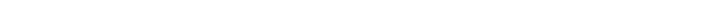
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	57	Total 57	O 57	0	0
3	B	25	Total 25	O 25	0	0
3	C	36	Total 36	O 36	0	0

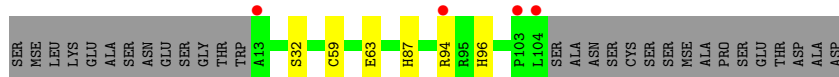
- Molecule 1: Zinc finger protein CONSTANS-LIKE 2



- Chain B: 2% 65% 7% 26%



- Chain C:  3% 71% 5% 24%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	36.93Å 44.48Å 99.77Å 90.00° 100.05° 90.00°	Depositor
Resolution (Å)	22.93 – 2.10 22.93 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.5 (22.93-2.10) 99.5 (22.93-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.67 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.179 , 0.200 0.179 , 0.200	Depositor DCC
R_{free} test set	964 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	27.8	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 59.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2134	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.02% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.64	0/672	0.78	0/904
1	B	0.71	0/664	0.84	2/895 (0.2%)
1	C	0.68	0/688	0.78	0/927
All	All	0.68	0/2024	0.80	2/2726 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	99	VAL	CA-C-N	-5.92	113.86	119.85
1	B	99	VAL	C-N-CA	-5.92	113.86	119.85

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	666	0	630	6	0
1	B	657	0	620	11	0
1	C	681	0	648	4	0
2	A	4	0	0	0	0
2	B	4	0	0	0	0
2	C	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	57	0	0	1	0
3	B	25	0	0	1	0
3	C	36	0	0	0	0
All	All	2134	0	1898	19	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 5.

All (19) 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:101:MSE:O	1:A:102:MSE:SE	2.42	0.88
1:B:44:HIS:CD2	1:B:55:ARG:HD3	2.19	0.77
1:A:101:MSE:C	1:A:102:MSE:SE	2.91	0.63
1:C:87:HIS:O	1:C:94:ARG:HD2	1.99	0.63
1:B:44:HIS:NE2	1:B:55:ARG:HD3	2.22	0.54
1:A:59:CYS:O	1:A:63:GLU:HA	2.10	0.51
1:B:91:PRO:O	1:B:95:ARG:HG3	2.13	0.49
1:C:59:CYS:O	1:C:63:GLU:HA	2.13	0.49
1:B:48:ARG:HH11	1:B:48:ARG:HB2	1.78	0.48
1:C:87:HIS:CD2	1:C:96:HIS:HB2	2.50	0.47
1:B:28:CYS:HB2	1:B:55:ARG:HG2	1.95	0.47
1:B:55:ARG:NH2	3:B:301:HOH:O	2.48	0.45
1:A:57:ARG:NE	3:A:301:HOH:O	2.36	0.45
1:A:85:GLU:CD	1:B:72:LYS:HE2	2.43	0.44
1:B:48:ARG:HH11	1:B:48:ARG:CB	2.31	0.43
1:B:40:ASP:OD1	1:B:55:ARG:HD2	2.19	0.42
1:A:64:SER:HB3	1:C:32:SER:OG	2.20	0.41
1:B:48:ARG:CB	1:B:48:ARG:NH1	2.84	0.41
1:B:101:MSE:HE3	1:B:101:MSE:HB3	1.51	0.41

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	88/121 (73%)	86 (98%)	2 (2%)	0	100	100
1	B	87/121 (72%)	86 (99%)	1 (1%)	0	100	100
1	C	90/121 (74%)	88 (98%)	2 (2%)	0	100	100
All	All	265/363 (73%)	260 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	70/91 (77%)	69 (99%)	1 (1%)	59	67
1	B	70/91 (77%)	68 (97%)	2 (3%)	37	42
1	C	72/91 (79%)	72 (100%)	0	100	100
All	All	212/273 (78%)	209 (99%)	3 (1%)	59	67

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	MSE
1	B	48	ARG
1	B	101	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	44	HIS
1	B	97	GLN
1	C	97	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 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, 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	87/121 (71%)	-0.09	1 (1%) 78 80	24, 31, 47, 68	0
1	B	86/121 (71%)	0.20	2 (2%) 61 64	24, 35, 56, 78	0
1	C	89/121 (73%)	0.17	4 (4%) 38 40	24, 32, 51, 63	0
All	All	262/363 (72%)	0.09	7 (2%) 56 59	24, 32, 55, 78	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	103	PRO	3.4
1	C	104	LEU	3.4
1	B	15	ALA	3.2
1	C	13	ALA	2.6
1	C	94	ARG	2.2
1	C	103	PRO	2.2
1	A	48	ARG	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 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,

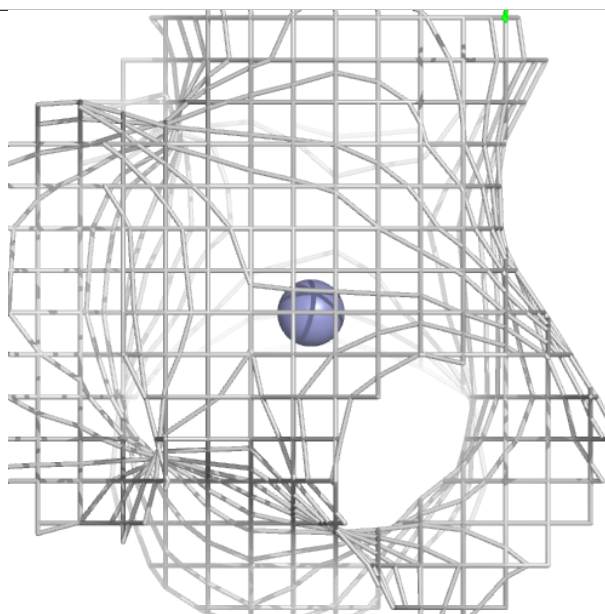
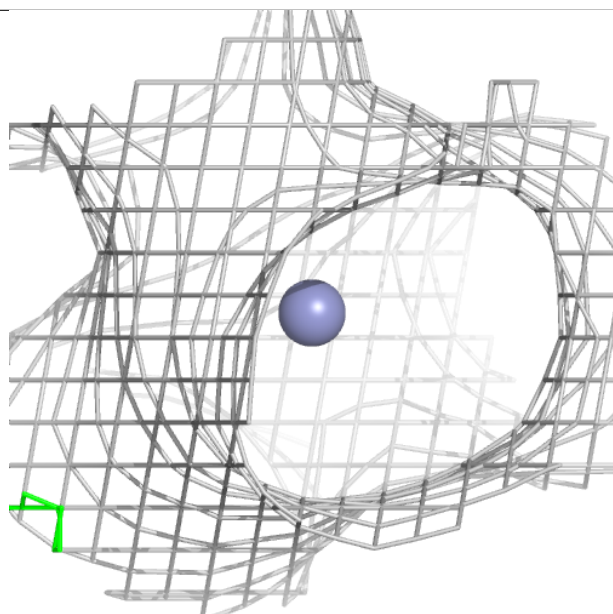
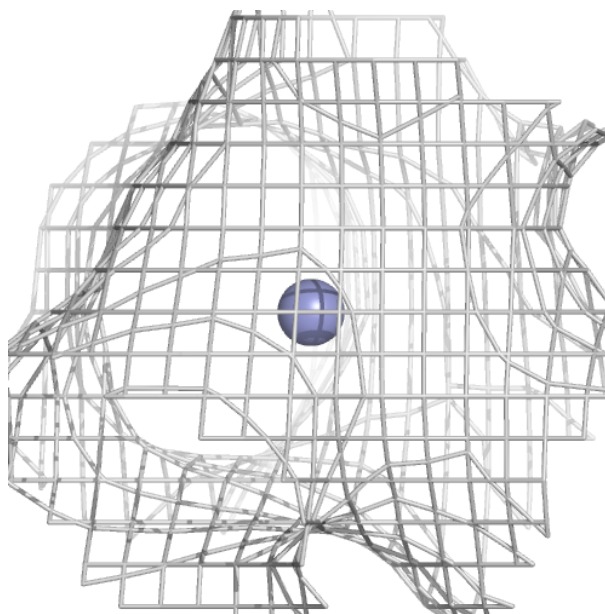
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	ZN	B	201	1/1	0.99	0.02	50,50,50,50	0
2	ZN	C	203	1/1	0.99	0.01	28,28,28,28	0
2	ZN	C	204	1/1	0.99	0.01	27,27,27,27	0
2	ZN	A	204	1/1	1.00	0.04	25,25,25,25	0
2	ZN	A	201	1/1	1.00	0.01	33,33,33,33	0
2	ZN	B	202	1/1	1.00	0.01	27,27,27,27	0
2	ZN	B	203	1/1	1.00	0.01	33,33,33,33	0
2	ZN	B	204	1/1	1.00	0.02	26,26,26,26	0
2	ZN	C	201	1/1	1.00	0.02	34,34,34,34	0
2	ZN	C	202	1/1	1.00	0.01	26,26,26,26	0
2	ZN	A	202	1/1	1.00	0.02	27,27,27,27	0
2	ZN	A	203	1/1	1.00	0.02	27,27,27,27	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.

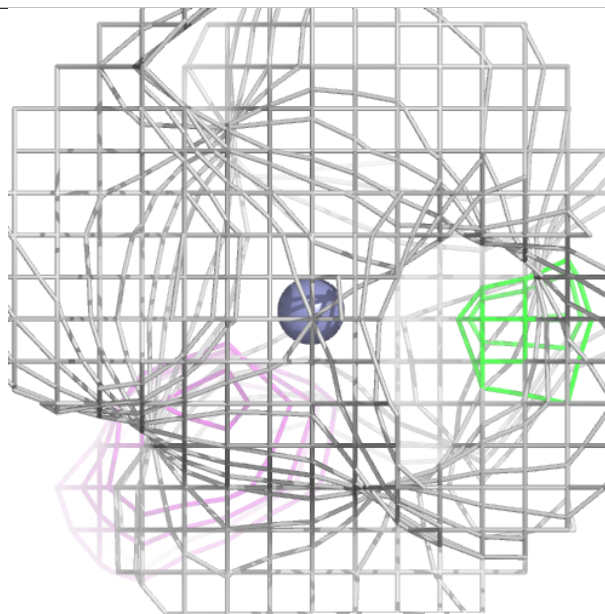
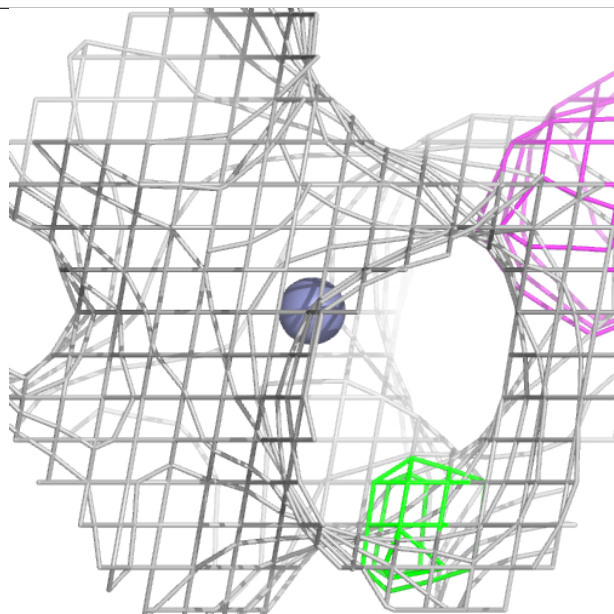
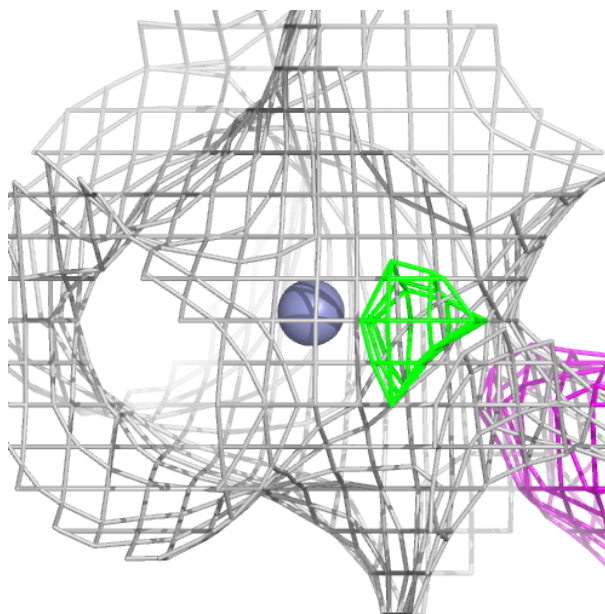
Electron density around ZN B 201:

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



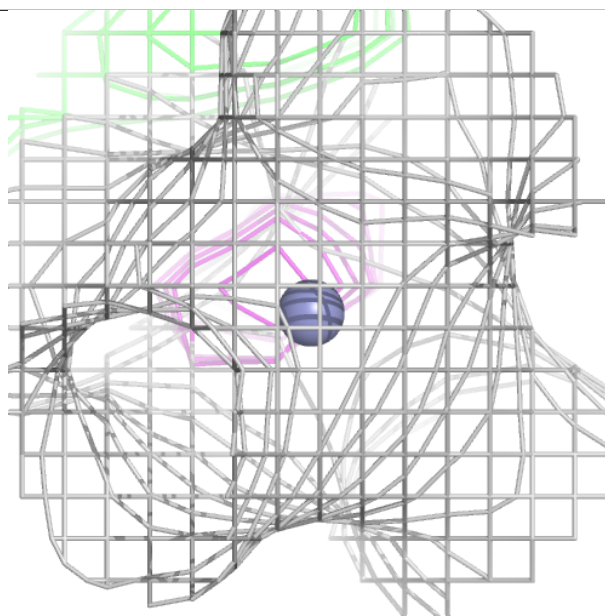
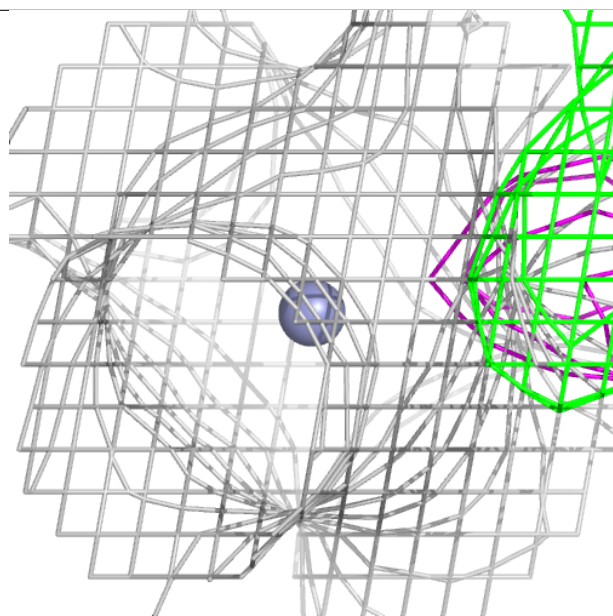
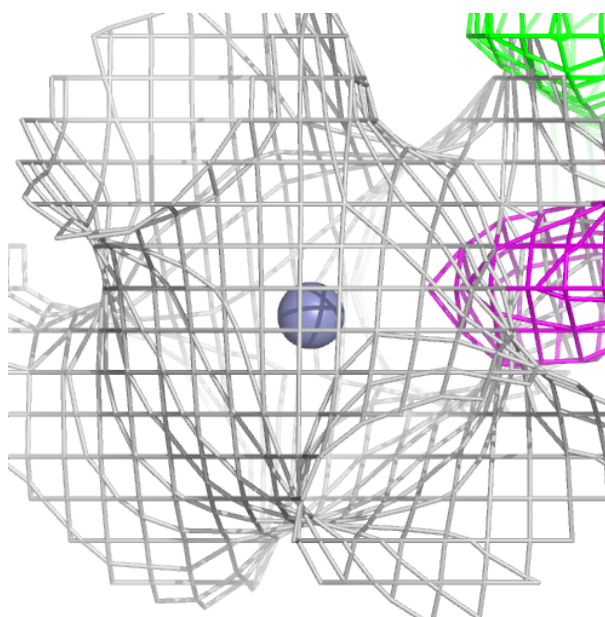
Electron density around ZN C 203:

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



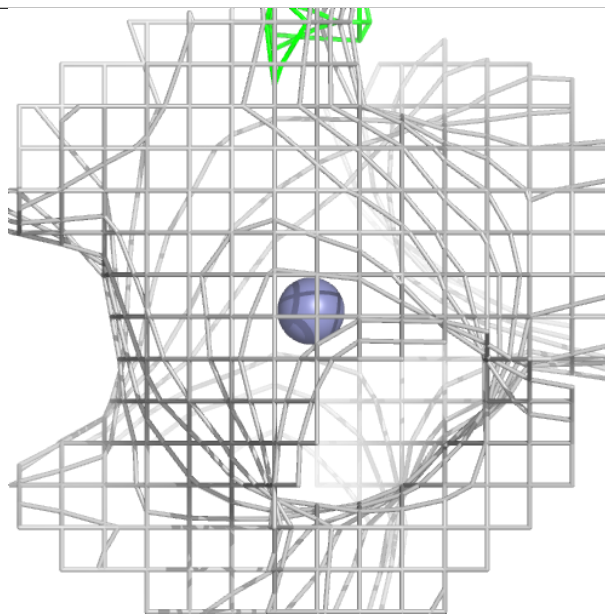
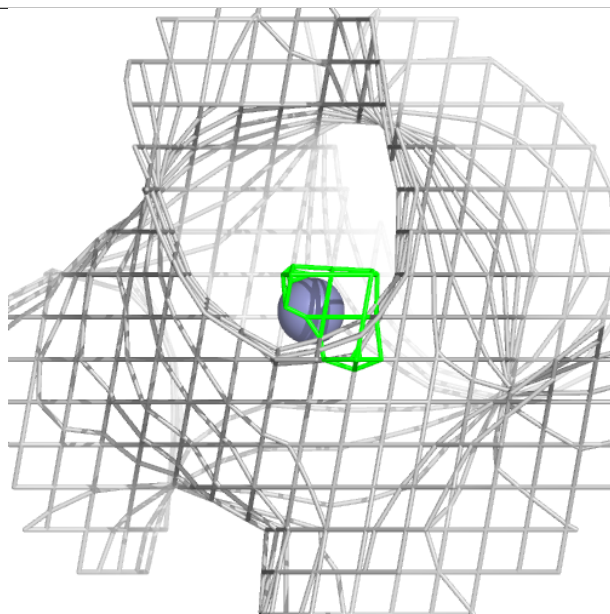
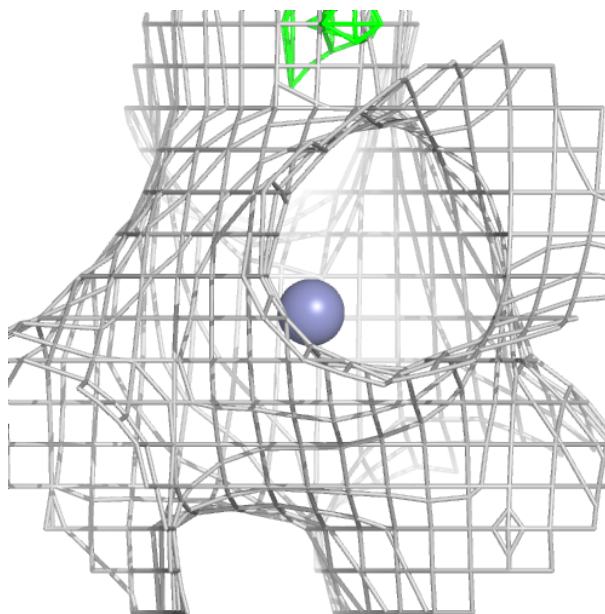
Electron density around ZN C 204:

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



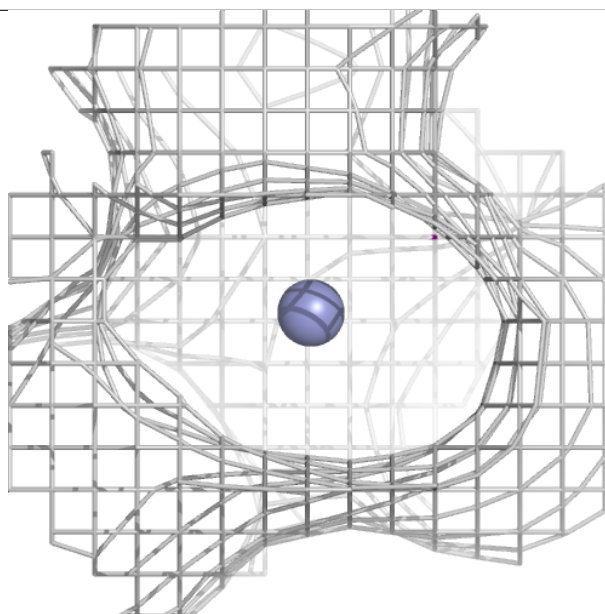
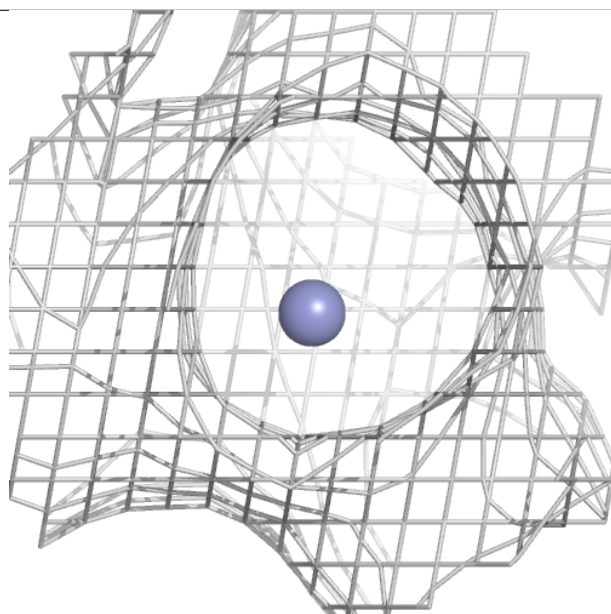
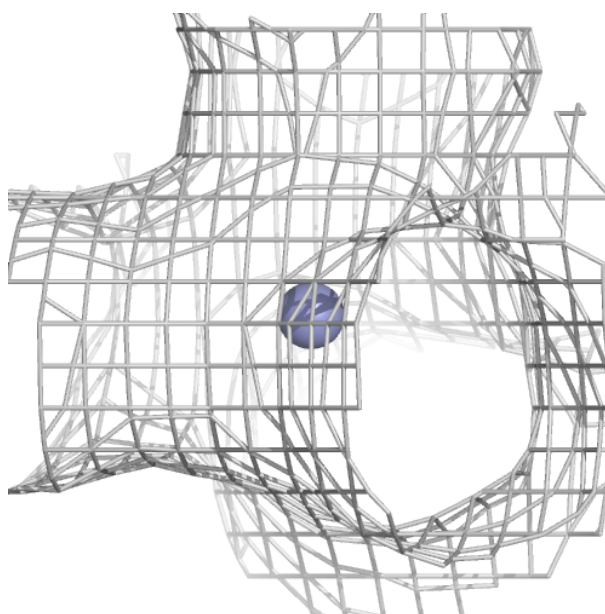
Electron density around ZN A 204:

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



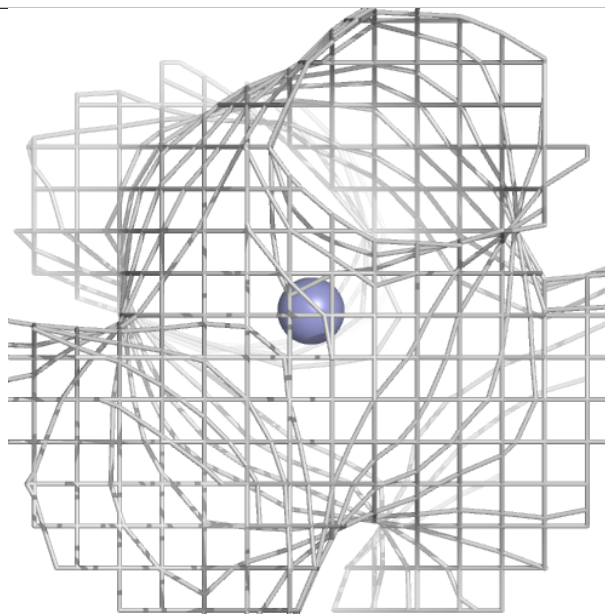
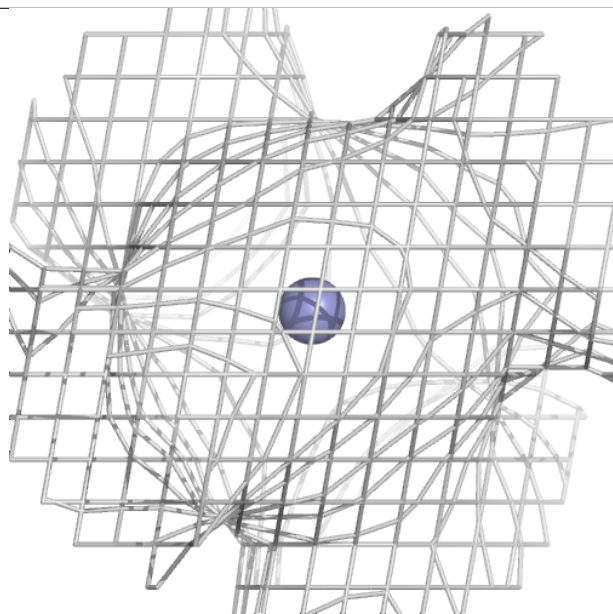
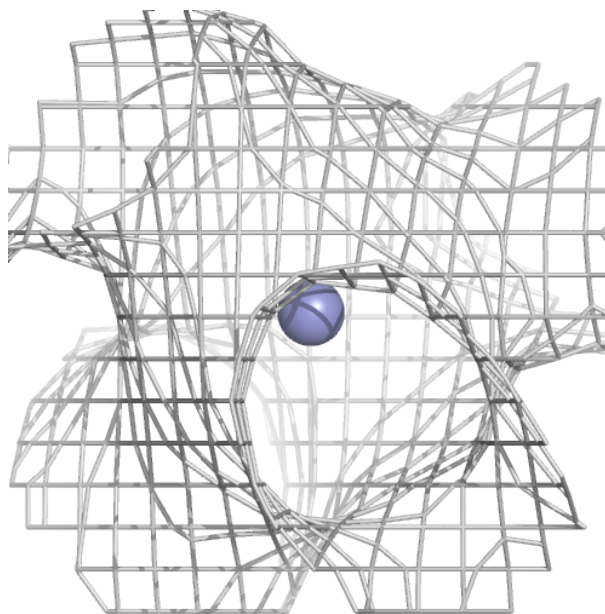
Electron density around ZN A 201:

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



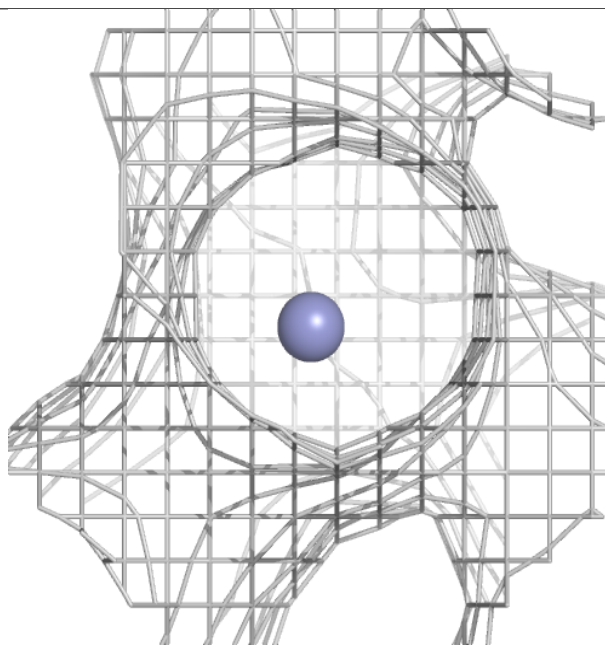
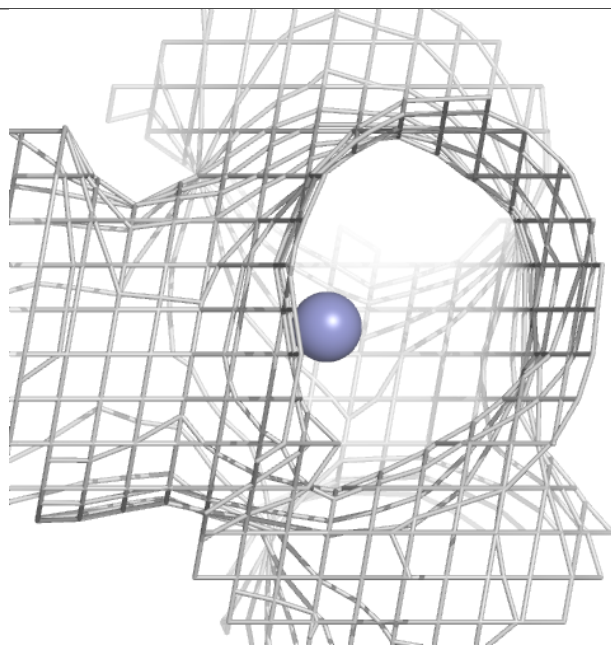
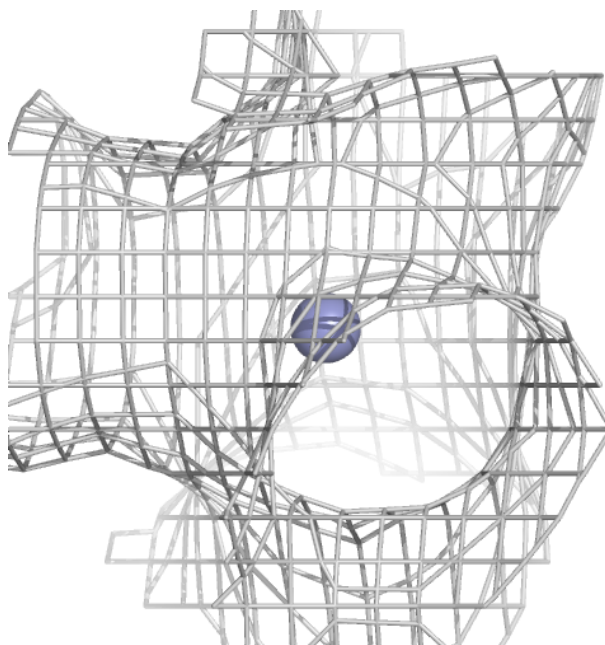
Electron density around ZN B 202:

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



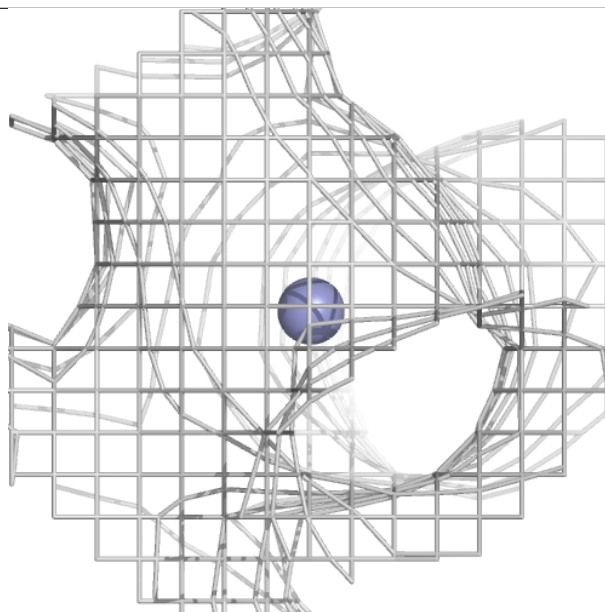
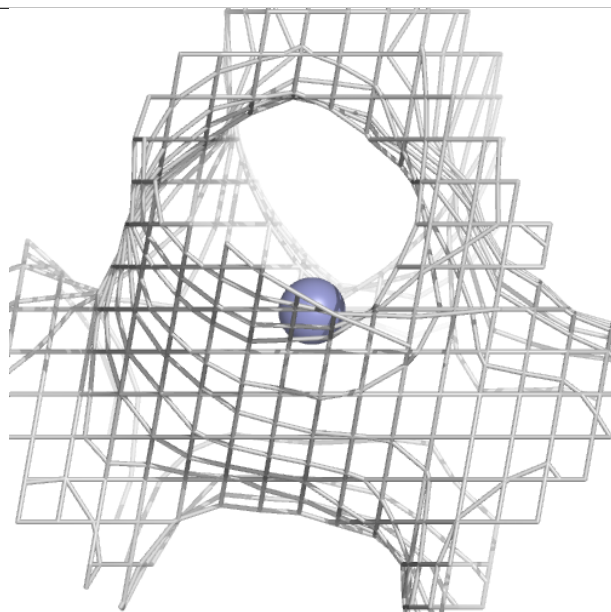
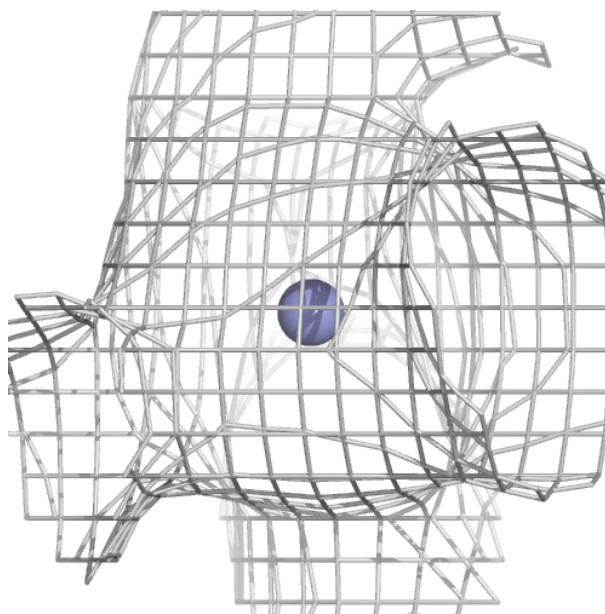
Electron density around ZN B 203:

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



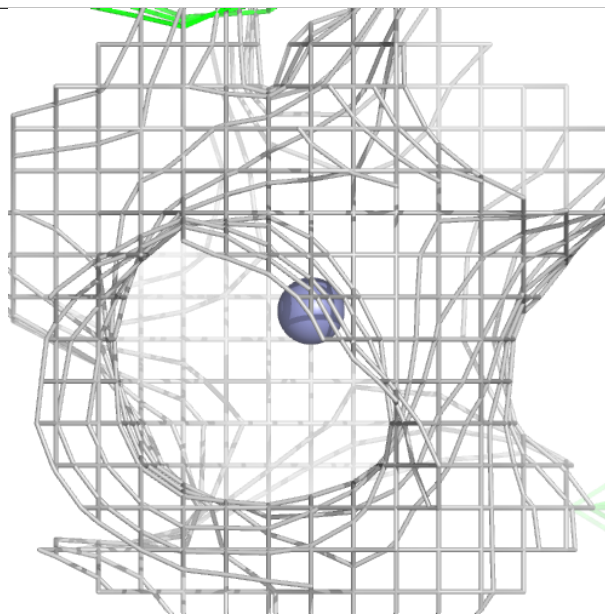
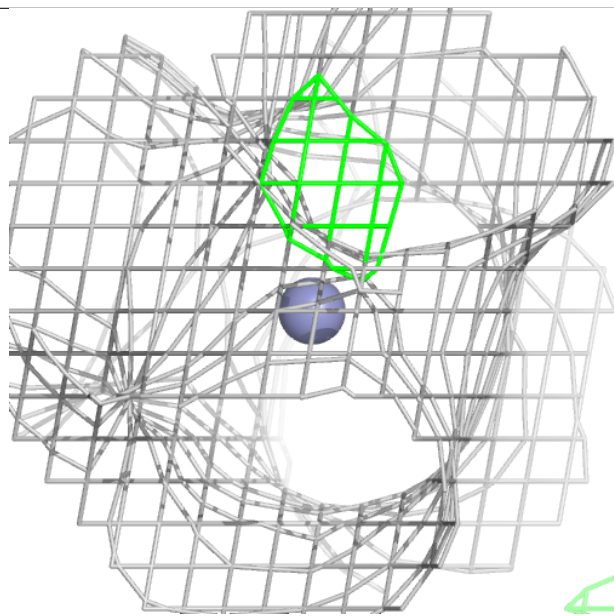
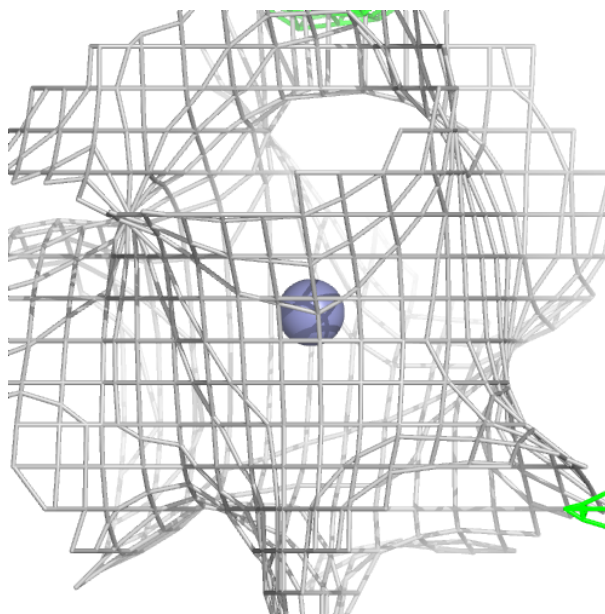
Electron density around ZN B 204:

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



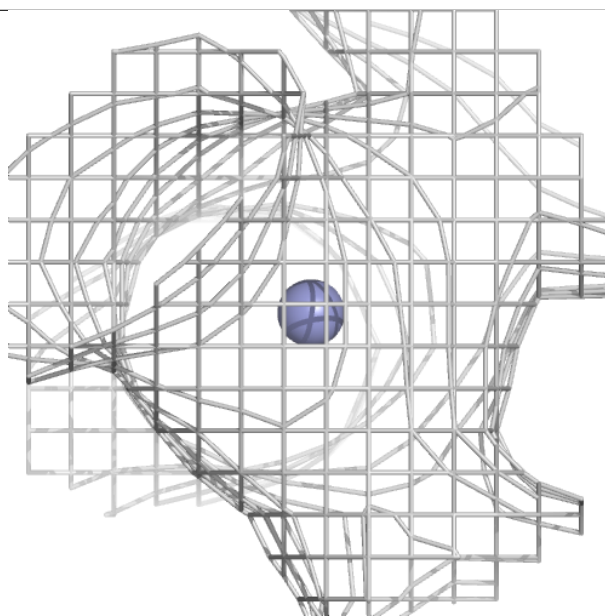
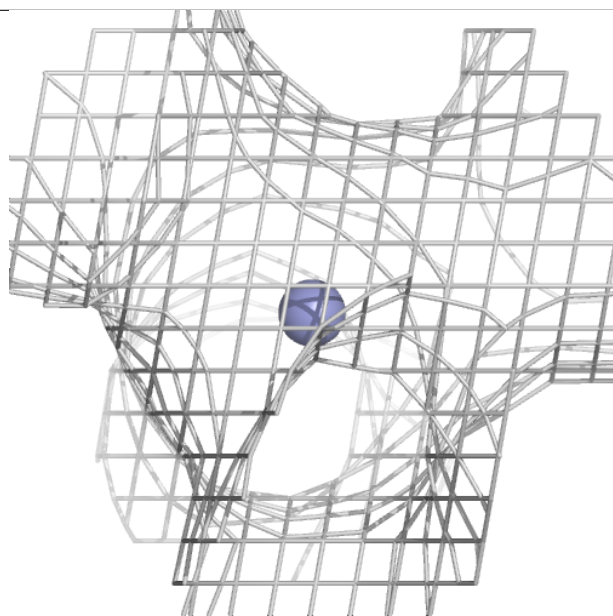
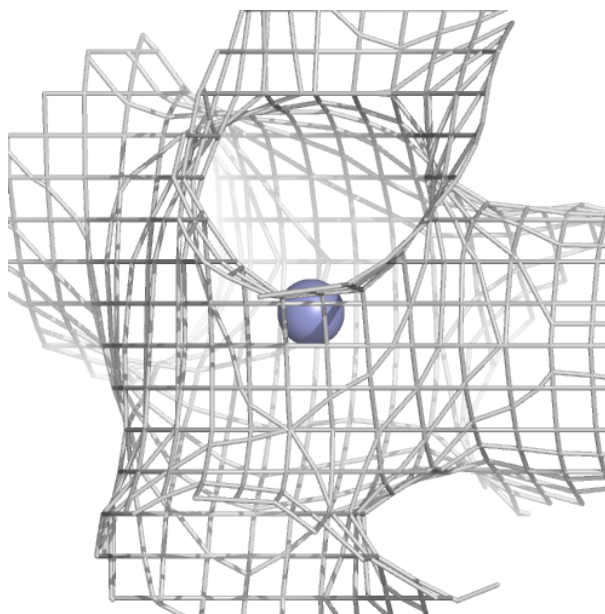
Electron density around ZN C 201:

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



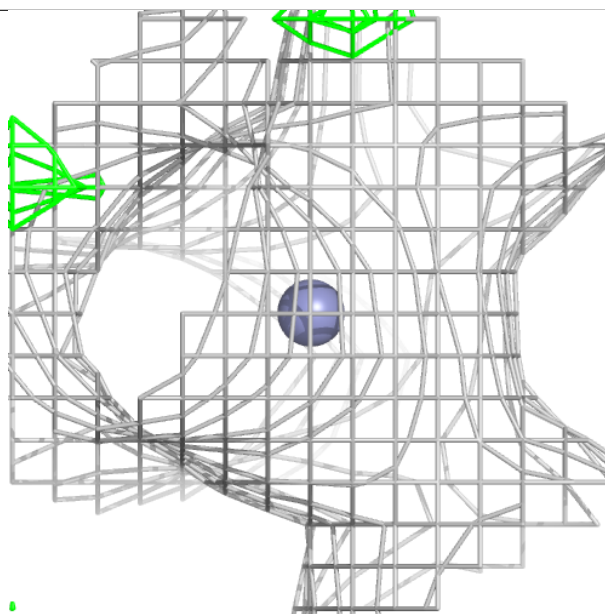
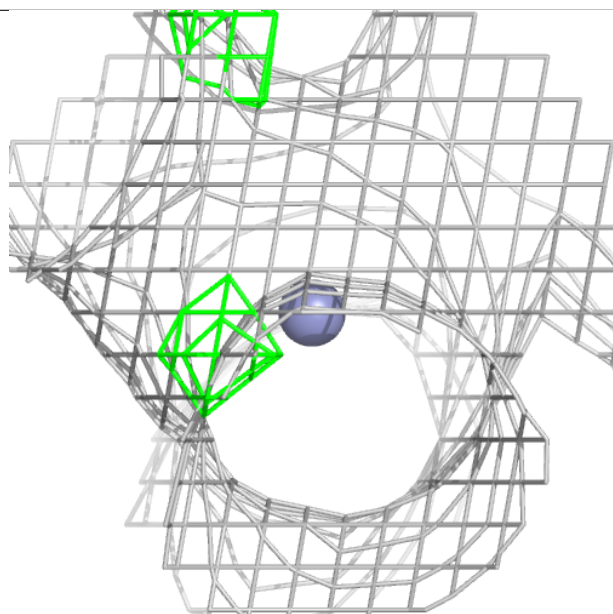
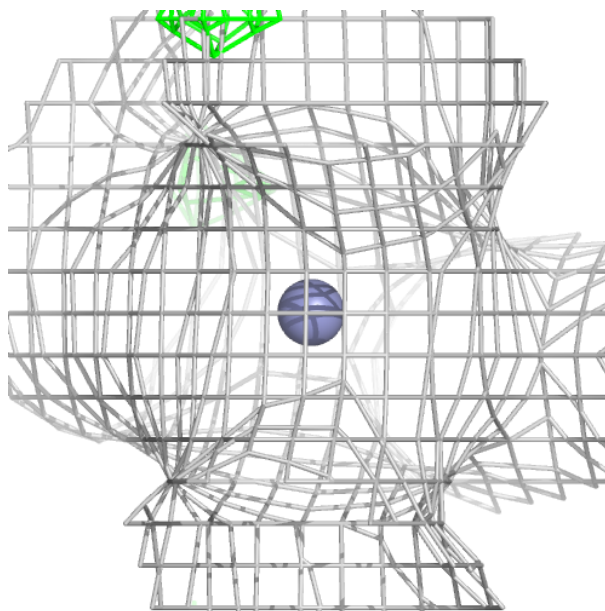
Electron density around ZN C 202:

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



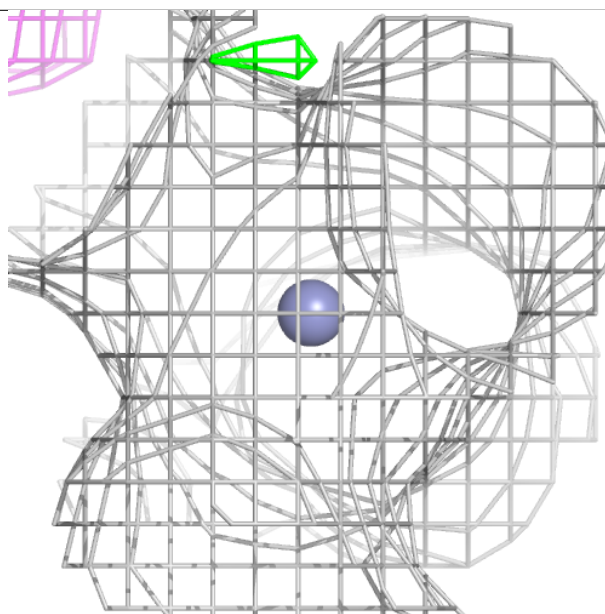
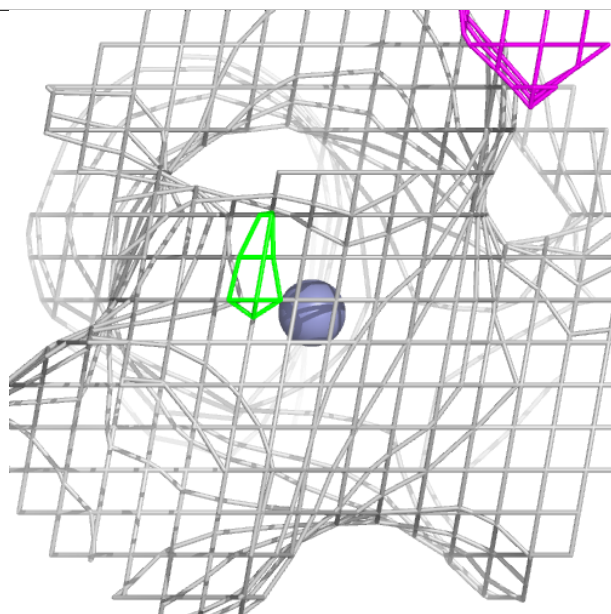
Electron density around ZN A 202:

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



Electron density around ZN A 203:

$2mF_o-DF_c$ (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.