wwPDB X-ray Structure Validation Summary Report (i)

Nov 27, 2023 – 04:10 PM JST

PDB ID	:	7YPR
Title	:	Structural basis of a superoxide dismutase from a tardigrade, Ramazzottius varieornatus strain YOKOZUNA-1.
Authors	:	Sim, K.-S.; Fukuda, Y.; Inoue, T.
Deposited on	:	2022-08-04
Resolution	:	2.10 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 (i) 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 (1)) were used in the production of this report:

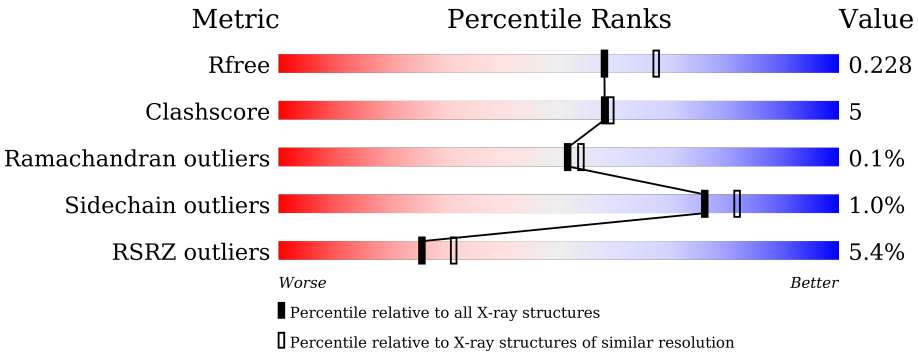
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson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

1 Overall quality at a glance (i)

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}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 <=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	194	76% 9% 15%
1	B	194	77% 7% 15%
1	C	194	80% 5% 15%
1	D	194	76% 9% 15%
1	E	194	18% 71% 14% 15%
1	F	194	6% 74% 11% 15%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7914 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 Superoxide dismutase [Cu-Zn].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	165	Total	C	N	O	S	0	6	0
			1248	780	224	239	5			
1	B	164	Total	C	N	O	S	0	7	0
			1247	778	221	243	5			
1	C	165	Total	C	N	O	S	0	2	0
			1217	764	216	234	3			
1	D	164	Total	C	N	O	S	7	10	0
			1265	790	226	246	3			
1	E	165	Total	C	N	O	S	0	7	0
			1251	787	223	238	3			
1	F	165	Total	C	N	O	S	0	6	0
			1236	772	220	239	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85
B	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85
C	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85
D	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85
E	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85
F	87	HIS	VAL	engineered mutation	UNP A0A1D1VU85

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

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		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cu 1 1	0	0
3	B	1	Total Cu 1 1	0	0
3	C	1	Total Cu 1 1	0	0
3	D	1	Total Cu 1 1	0	0
3	E	1	Total Cu 1 1	0	0
3	F	1	Total Cu 1 1	0	0

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

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	100	Total O 102 102	0	2
5	B	88	Total O 93 93	0	6
5	C	103	Total O 105 105	0	2
5	D	75	Total O 76 76	0	1

Continued on next page...

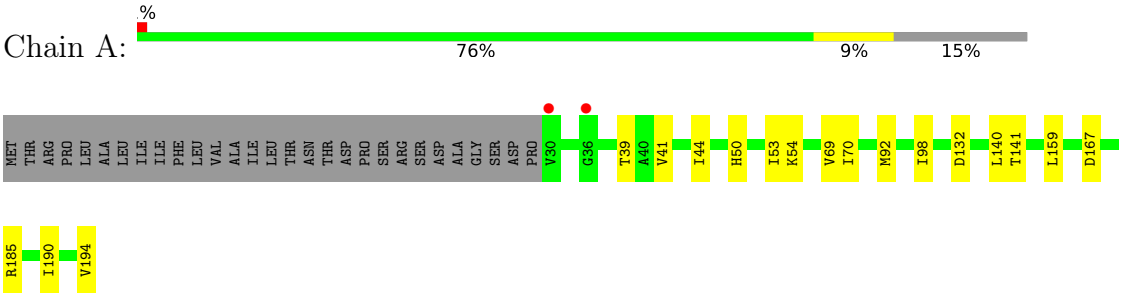
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	29	Total O 29 29	0	0
5	F	32	Total O 32 32	0	0

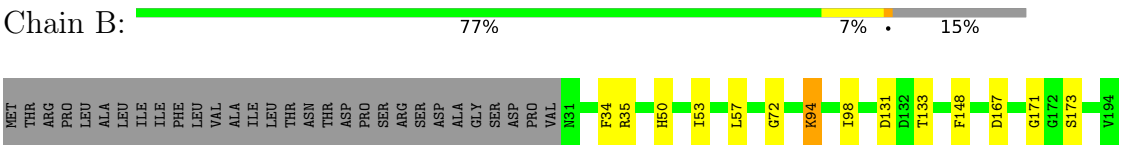
3 Residue-property plots (i)

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.

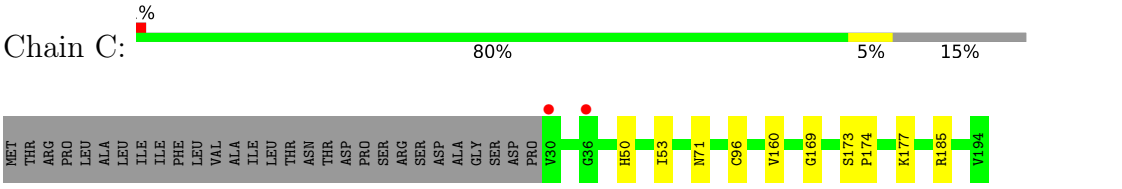
• Molecule 1: Superoxide dismutase [Cu-Zn]



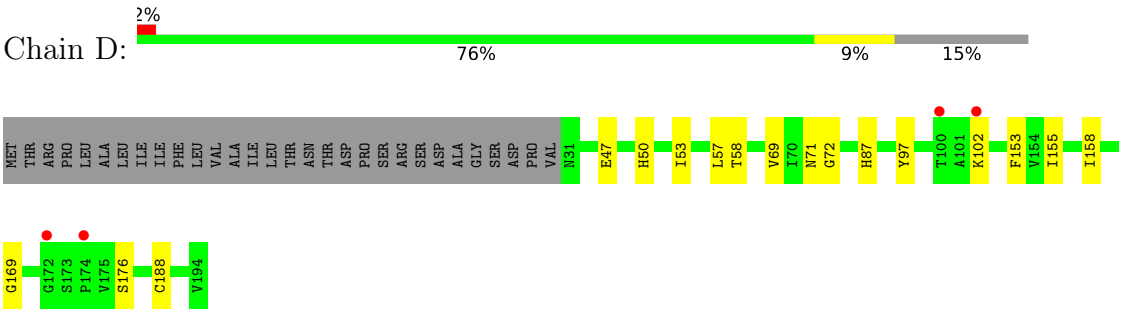
• Molecule 1: Superoxide dismutase [Cu-Zn]



• Molecule 1: Superoxide dismutase [Cu-Zn]

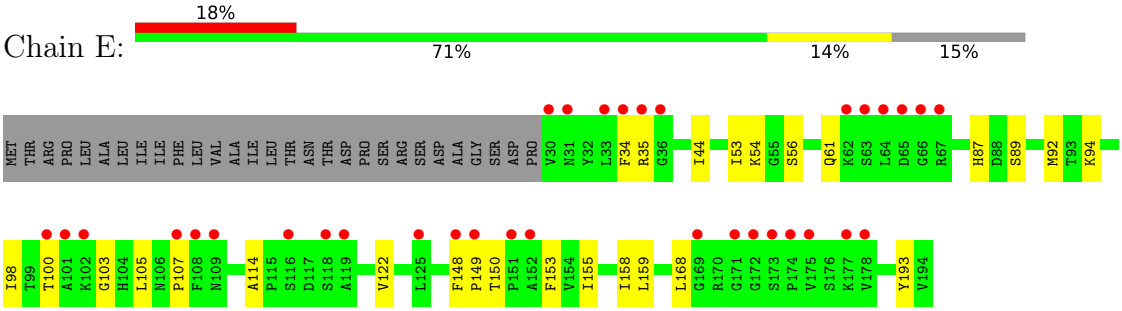


• Molecule 1: Superoxide dismutase [Cu-Zn]



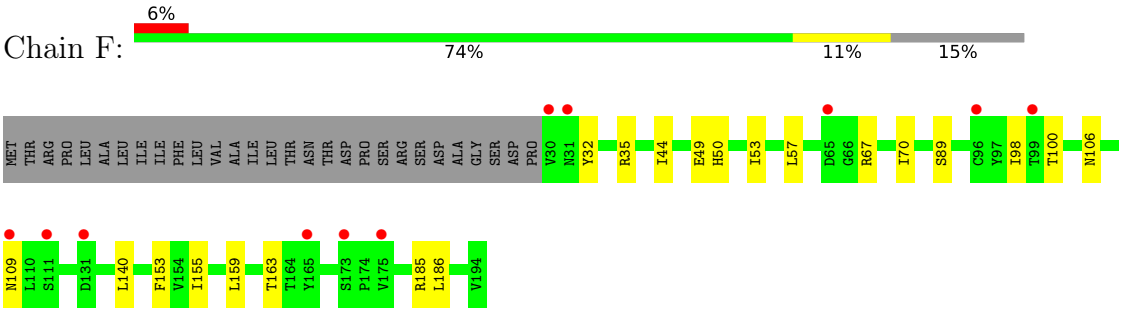
• Molecule 1: Superoxide dismutase [Cu-Zn]

Chain E:



• Molecule 1: Superoxide dismutase [Cu-Zn]

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	105.83Å 105.83Å 76.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.87 – 2.10 45.82 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.87-2.10) 99.7 (45.82-2.10)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.180 , 0.228 0.180 , 0.228	Depositor DCC
R_{free} test set	2800 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	38.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.001 for -h,-k,l 0.028 for h,-h-k,-l 0.004 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7914	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 18.40% 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: K, CU, 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	0.39	0/1272	0.66	1/1726 (0.1%)
1	B	0.40	0/1271	0.64	0/1725
1	C	0.43	0/1242	0.71	0/1686
1	D	0.37	0/1289	0.64	0/1750
1	E	0.51	2/1280 (0.2%)	0.72	4/1737 (0.2%)
1	F	0.36	0/1260	0.60	0/1710
All	All	0.41	2/7614 (0.0%)	0.66	5/10334 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	103[A]	GLY	C-N	-8.97	1.13	1.34
1	E	103[B]	GLY	C-N	-8.97	1.13	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	103[A]	GLY	C-N-CA	9.18	144.66	121.70
1	E	103[B]	GLY	C-N-CA	9.18	144.66	121.70
1	E	103[A]	GLY	O-C-N	-6.05	113.01	122.70
1	E	103[B]	GLY	O-C-N	-6.05	113.01	122.70
1	A	167	ASP	CB-CG-OD1	5.16	122.94	118.30

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	1248	0	1228	10	0
1	B	1247	0	1216	6	0
1	C	1217	0	1203	5	0
1	D	1265	0	1246	14	0
1	E	1251	0	1244	25	0
1	F	1236	0	1214	13	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
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	B	1	0	0	0	0
5	A	102	0	0	1	0
5	B	93	0	0	0	0
5	C	105	0	0	0	0
5	D	76	0	0	2	0
5	E	29	0	0	0	0
5	F	32	0	0	0	0
All	All	7914	0	7351	72	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.

The worst 5 of 72 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:F:70:ILE:HD12	1:F:140:LEU:HD11	1.06	1.05
1:F:70:ILE:HD12	1:F:140:LEU:CD1	1.91	1.00

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:69:VAL:HG12	1:D:71[B]:ASN:HD21	1.46	0.81
1:E:105[B]:LEU:HD22	1:E:122:VAL:CG1	2.12	0.80
1:D:50:HIS:HB3	1:D:53:ILE:HD12	1.78	0.66

There are no symmetry-related clashes.

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	169/194 (87%)	162 (96%)	7 (4%)	0	100 100
1	B	169/194 (87%)	162 (96%)	7 (4%)	0	100 100
1	C	165/194 (85%)	160 (97%)	5 (3%)	0	100 100
1	D	172/194 (89%)	168 (98%)	4 (2%)	0	100 100
1	E	170/194 (88%)	160 (94%)	8 (5%)	2 (1%)	13 8
1	F	169/194 (87%)	164 (97%)	5 (3%)	0	100 100
All	All	1014/1164 (87%)	976 (96%)	36 (4%)	2 (0%)	51 49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	100[A]	THR
1	E	100[B]	THR

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	134/153 (88%)	133 (99%)	1 (1%)	84 88
1	B	134/153 (88%)	131 (98%)	3 (2%)	52 57
1	C	130/153 (85%)	129 (99%)	1 (1%)	81 86
1	D	135/153 (88%)	135 (100%)	0	100 100
1	E	133/153 (87%)	131 (98%)	2 (2%)	65 71
1	F	132/153 (86%)	131 (99%)	1 (1%)	81 86
All	All	798/918 (87%)	790 (99%)	8 (1%)	76 82

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	67	ARG
1	E	56	SER
1	C	96	CYS
1	B	173	SER
1	E	35	ARG

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

Mol	Chain	Res	Type
1	C	139	ASN
1	E	109	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 13 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	103[B]:GLY	C	104:HIS	N	1.13
1	E	103[A]:GLY	C	104:HIS	N	1.13

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	165/194 (85%)	-0.06	2 (1%) 79 82	24, 33, 51, 71	0
1	B	164/194 (84%)	0.14	0 100 100	24, 34, 54, 71	0
1	C	165/194 (85%)	0.02	2 (1%) 79 82	25, 35, 57, 72	0
1	D	164/194 (84%)	-0.00	4 (2%) 59 64	26, 39, 59, 69	0
1	E	165/194 (85%)	0.91	34 (20%) 1 1	38, 51, 72, 80	1 (0%)
1	F	165/194 (85%)	0.31	11 (6%) 17 22	39, 53, 69, 90	0
All	All	988/1164 (84%)	0.22	53 (5%) 25 31	24, 41, 65, 90	1 (0%)

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	30	VAL	6.3
1	E	178	VAL	5.4
1	E	36	GLY	4.5
1	E	174	PRO	4.4
1	E	173	SER	4.4

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 (i)

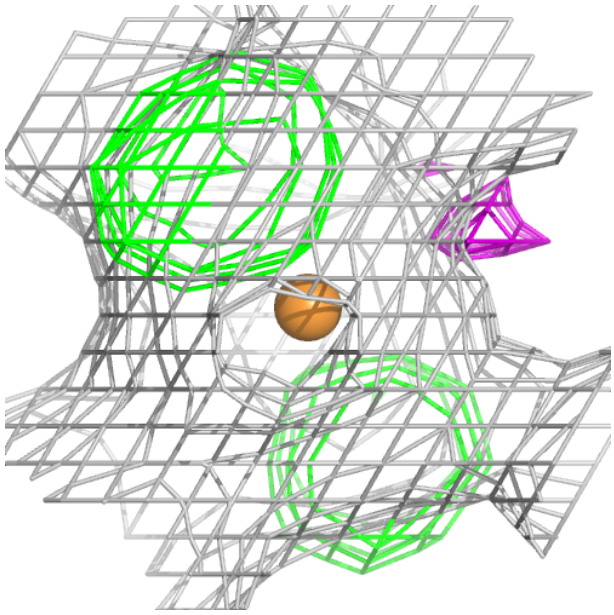
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

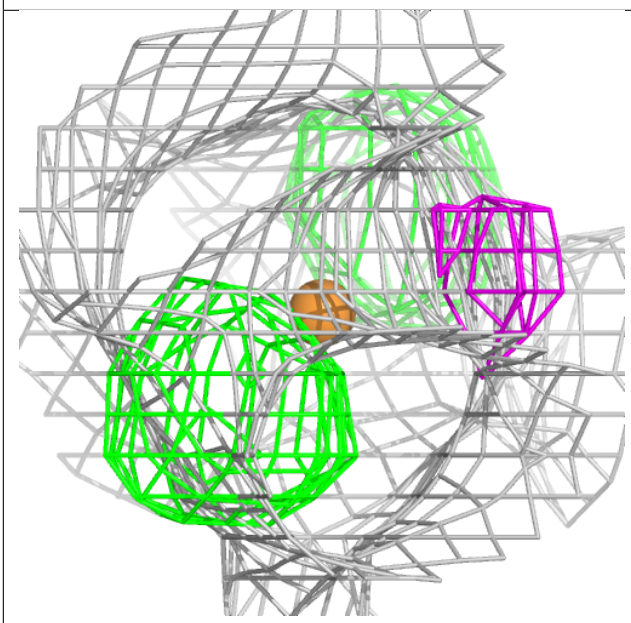
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CU	B	202	1/1	0.93	0.12	49,49,49,49	0
4	K	B	203	1/1	0.93	0.10	58,58,58,58	0
3	CU	D	202	1/1	0.96	0.07	53,53,53,53	0
3	CU	E	202	1/1	0.97	0.04	65,65,65,65	0
3	CU	F	202	1/1	0.98	0.08	46,46,46,46	1
2	ZN	D	201	1/1	0.99	0.11	37,37,37,37	0
3	CU	C	202	1/1	0.99	0.13	42,42,42,42	0
3	CU	A	202	1/1	0.99	0.10	39,39,39,39	0
2	ZN	A	201	1/1	1.00	0.12	31,31,31,31	0
2	ZN	E	201	1/1	1.00	0.05	47,47,47,47	0
2	ZN	F	201	1/1	1.00	0.08	49,49,49,49	0
2	ZN	B	201	1/1	1.00	0.16	27,27,27,27	0
2	ZN	C	201	1/1	1.00	0.13	30,30,30,30	0

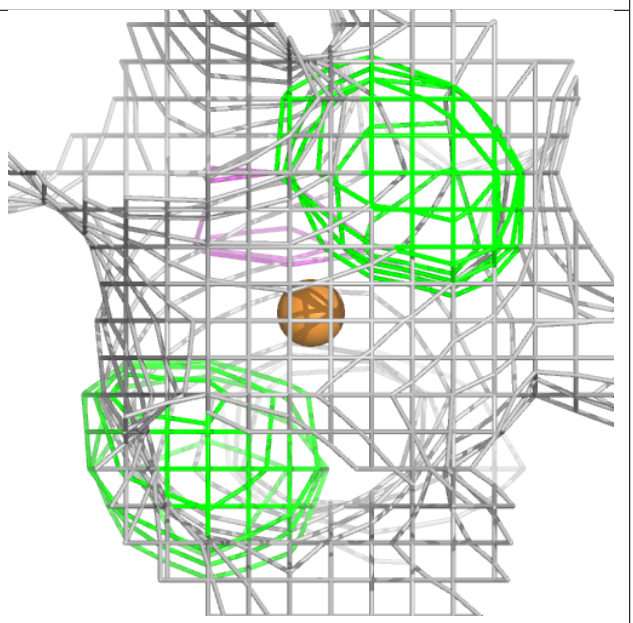
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CU 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)

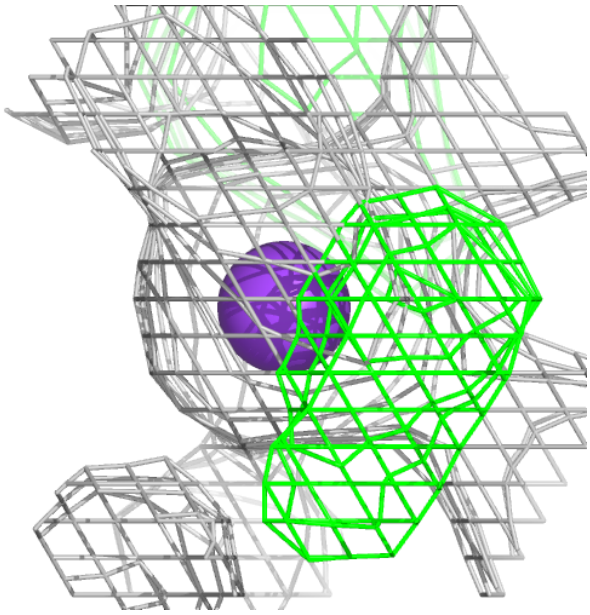


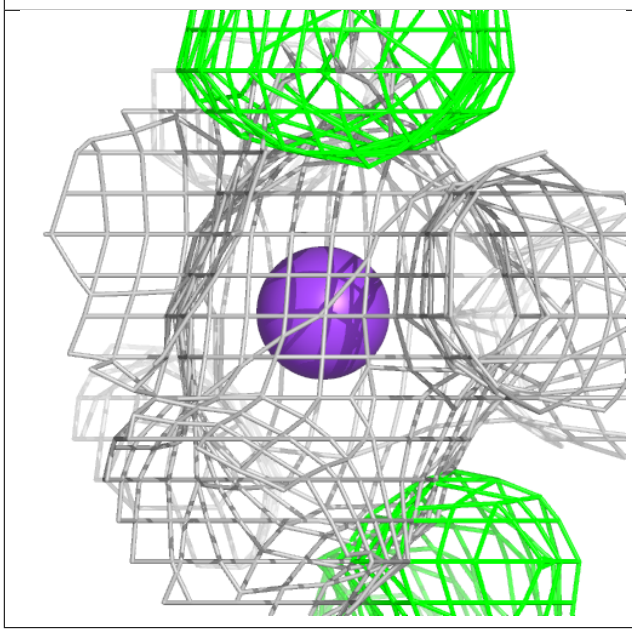


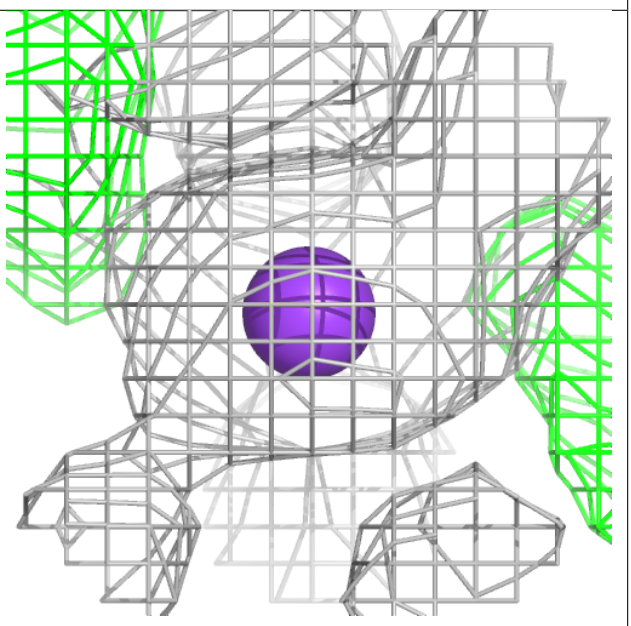


Electron density around K B 203:

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

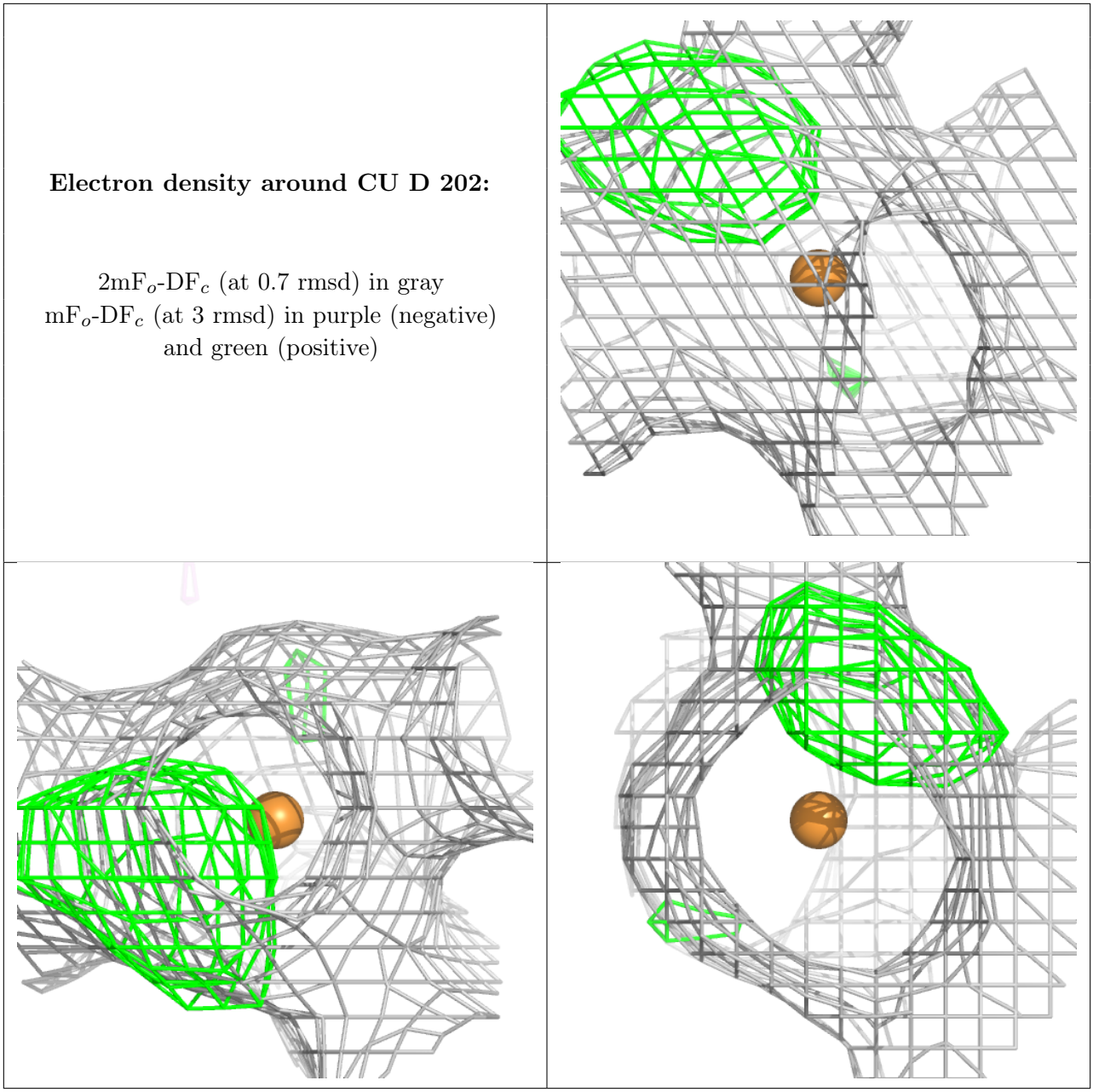






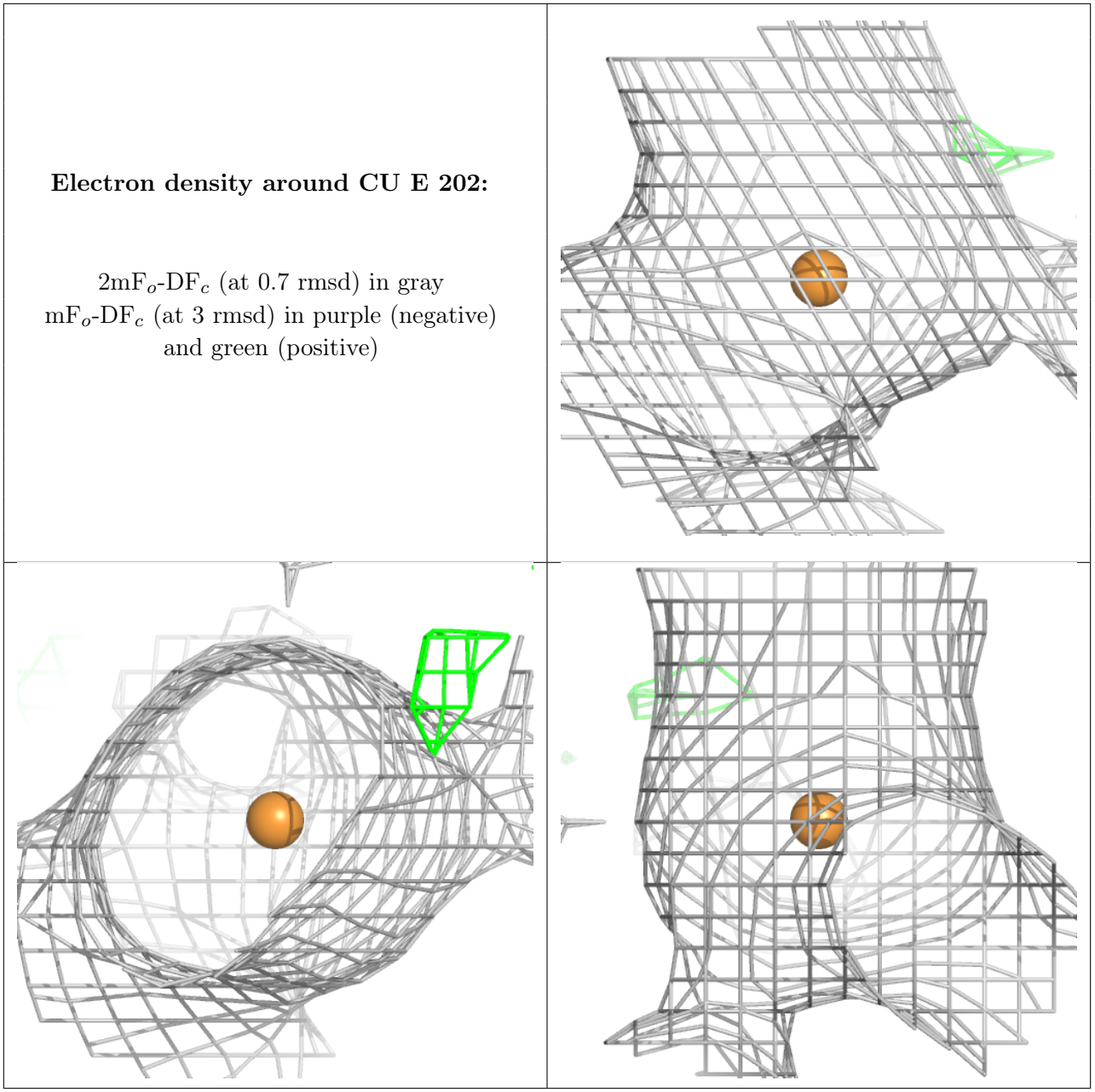
Electron density around CU D 202:

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



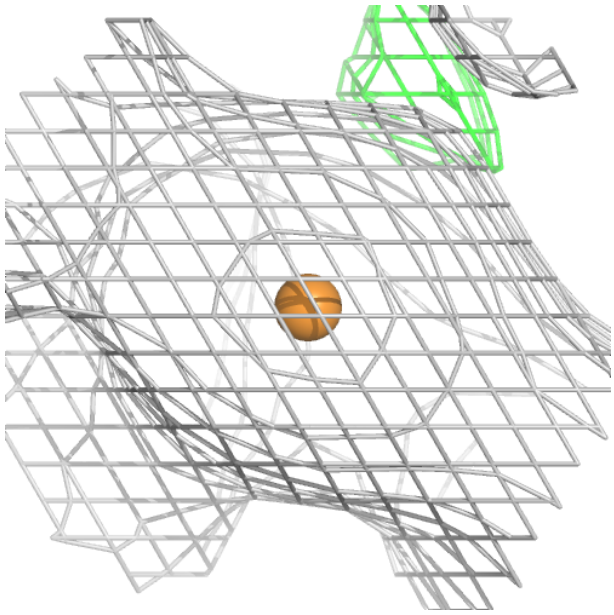
Electron density around CU E 202:

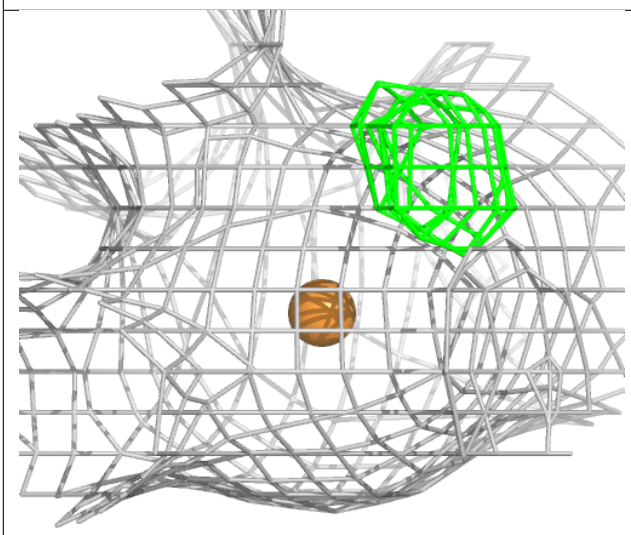
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

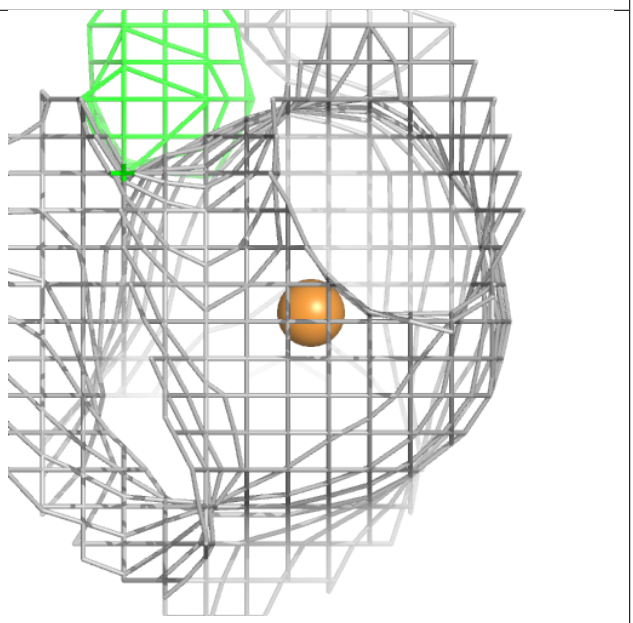


Electron density around CU F 202:

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

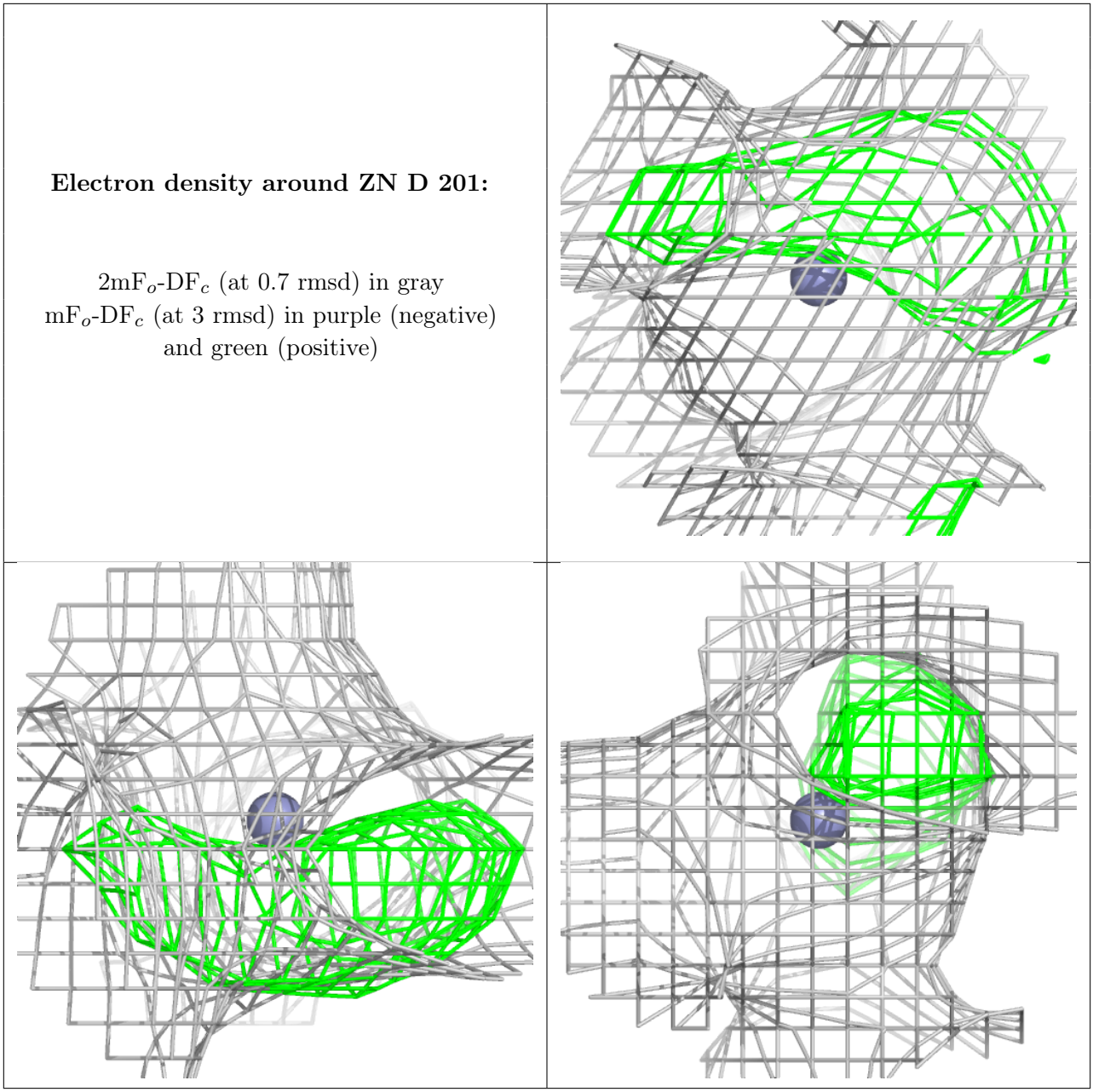






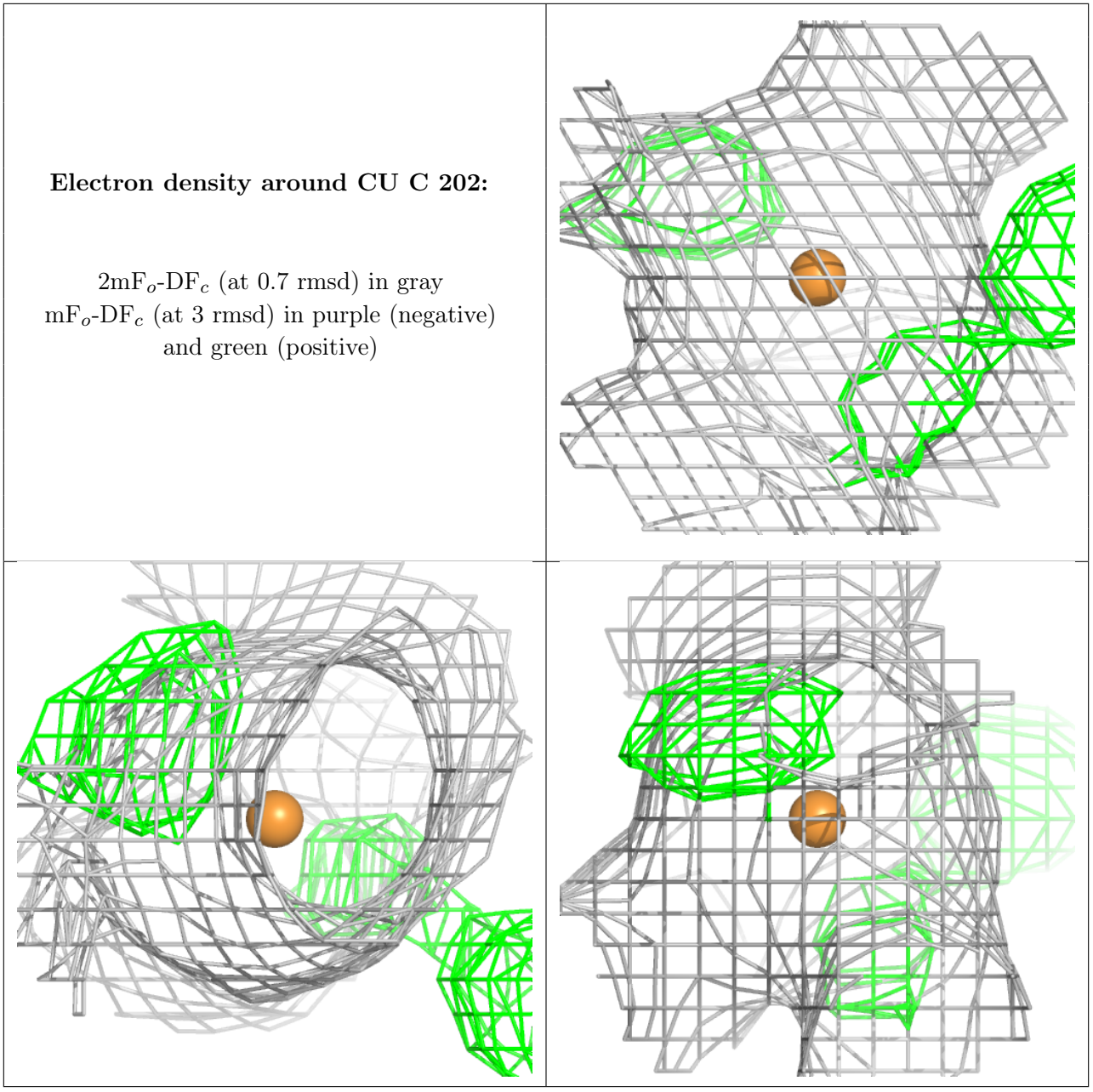
Electron density around ZN D 201:

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



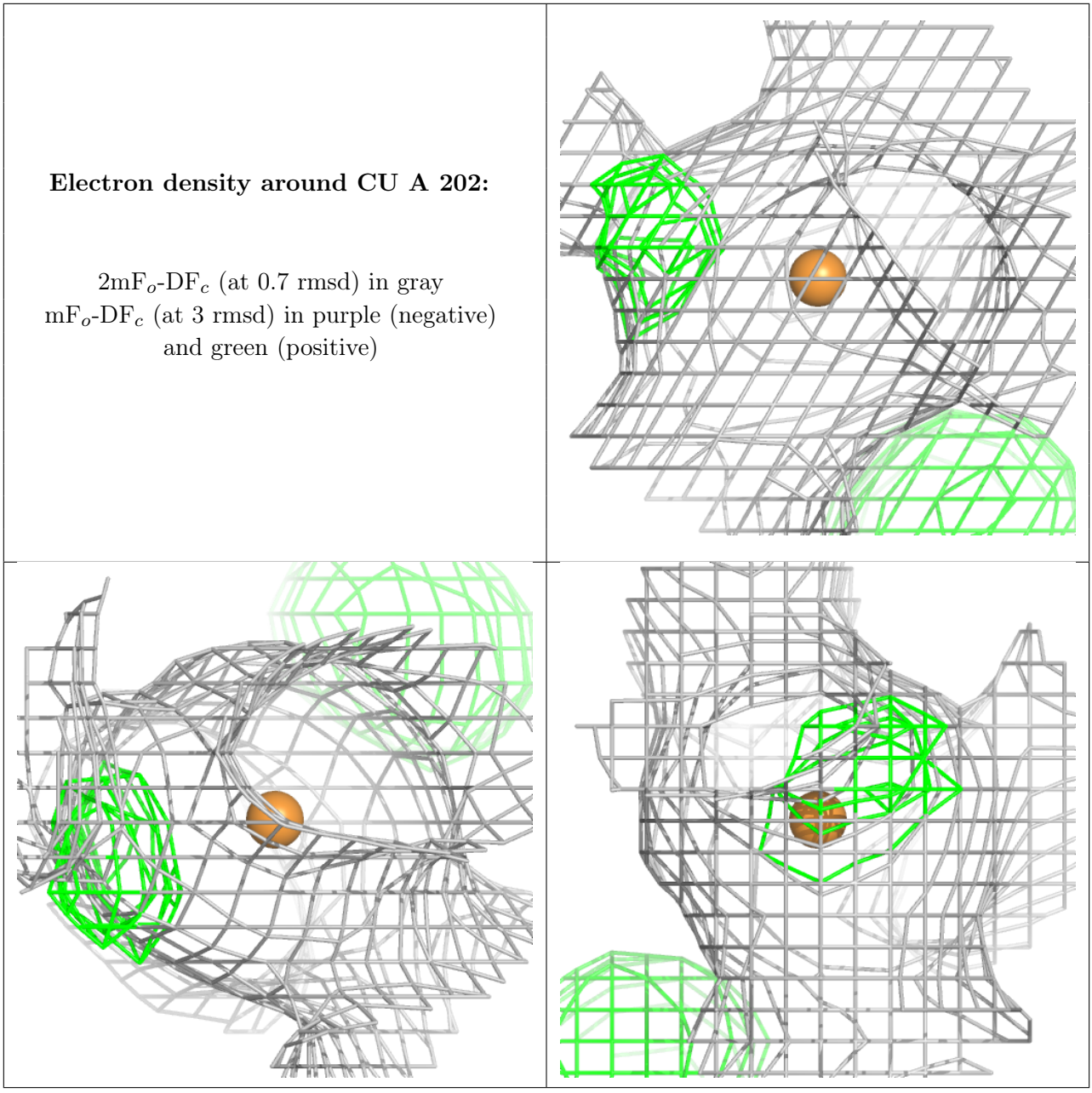
Electron density around CU C 202:

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



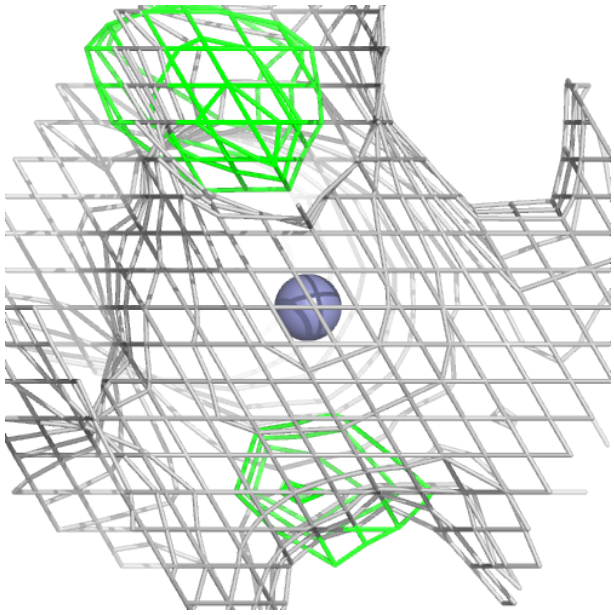
Electron density around CU A 202:

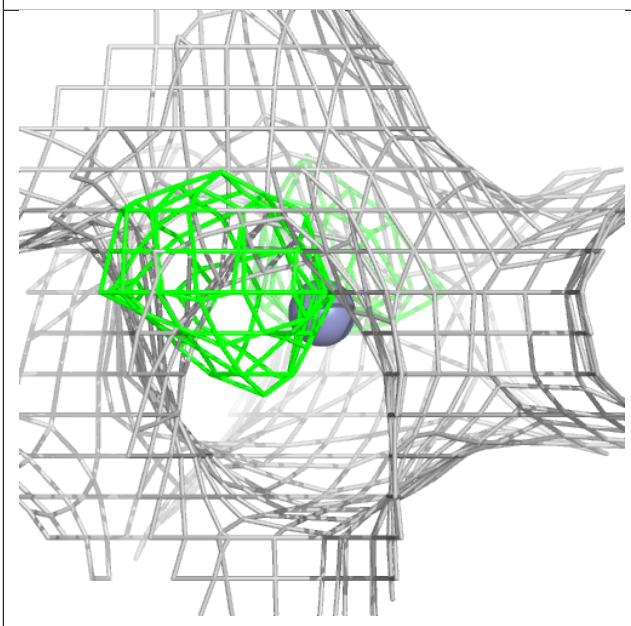
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

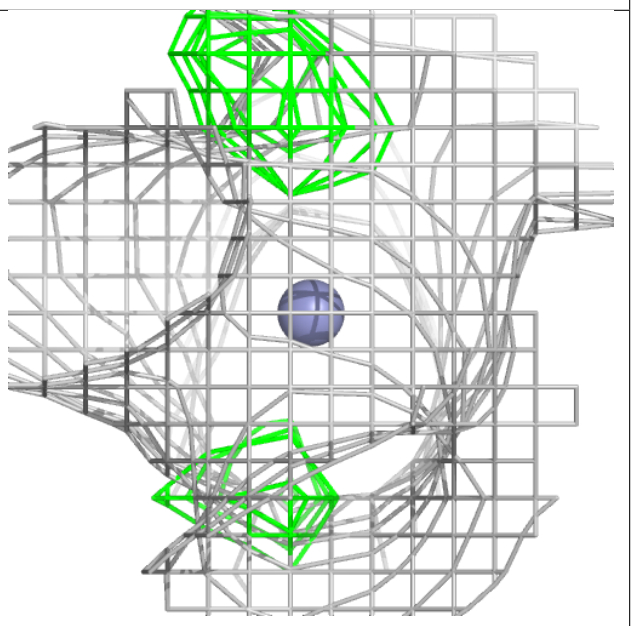


Electron density around ZN A 201:

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

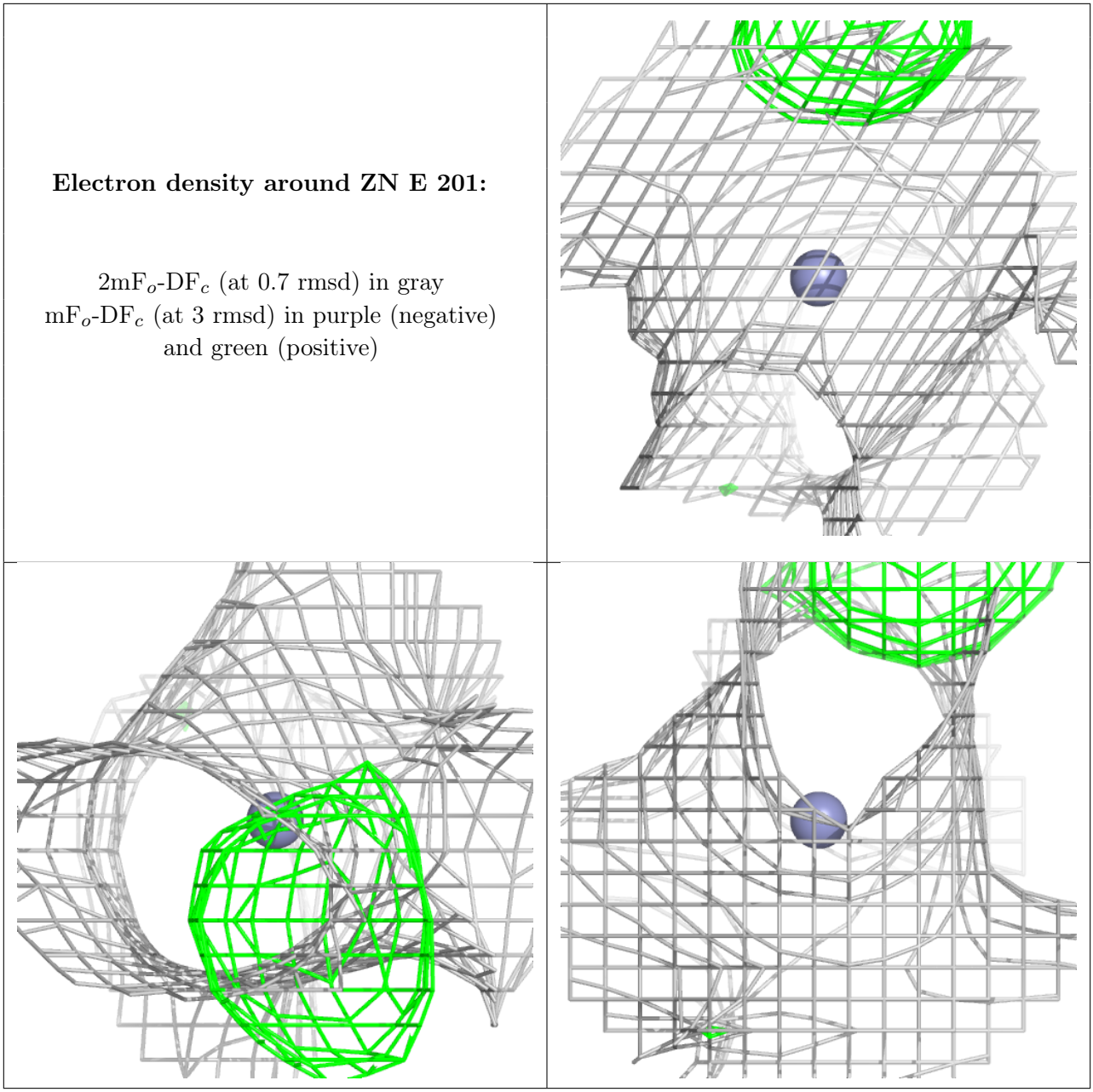






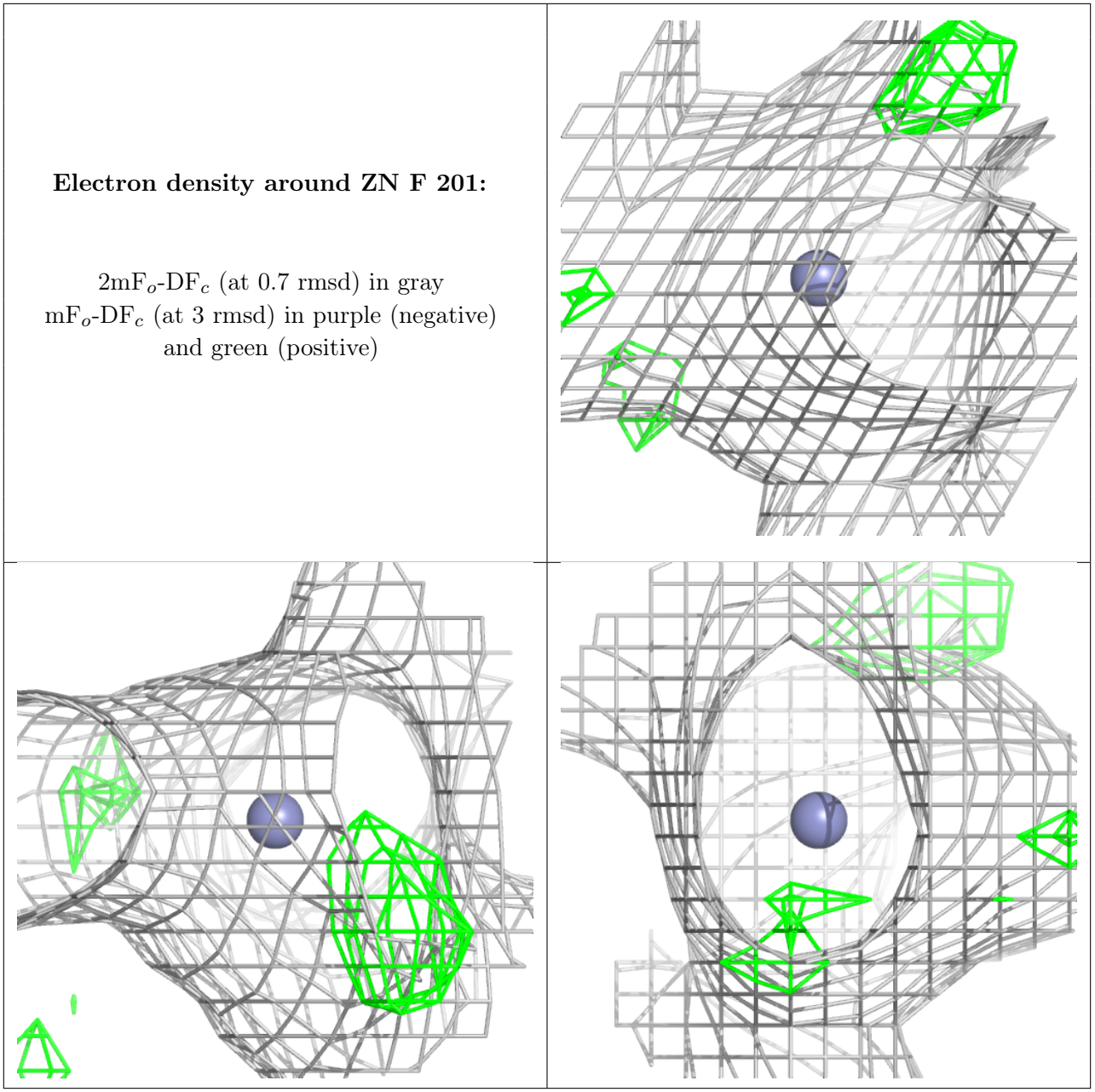
Electron density around ZN E 201:

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



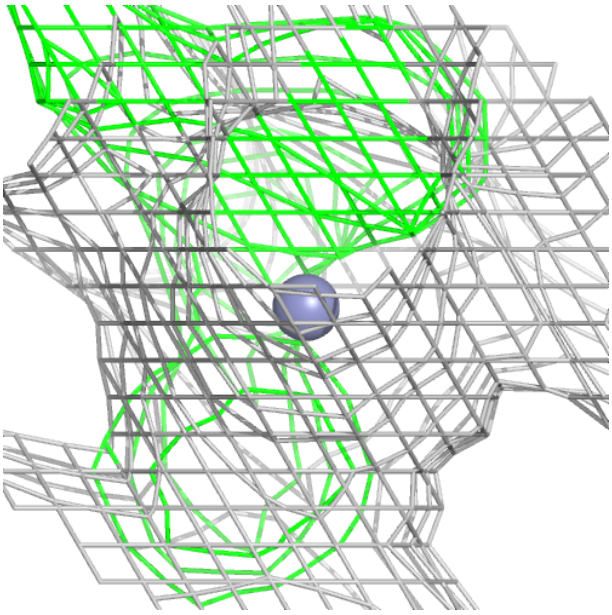
Electron density around ZN F 201:

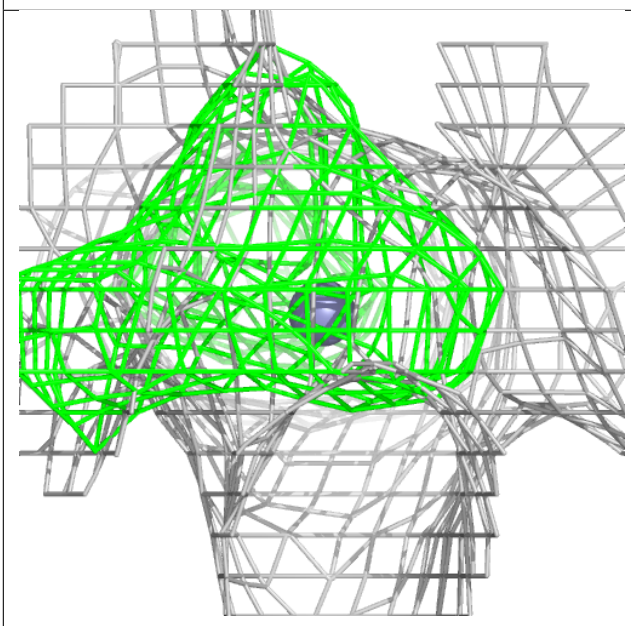
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

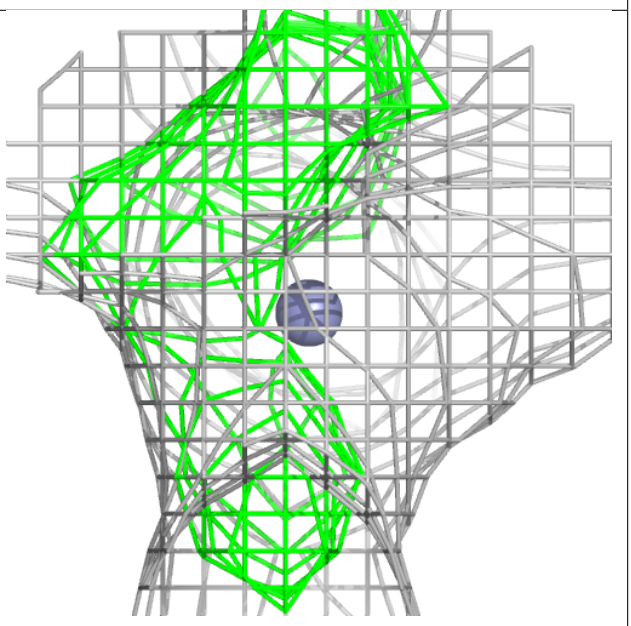


Electron density around ZN B 201:

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

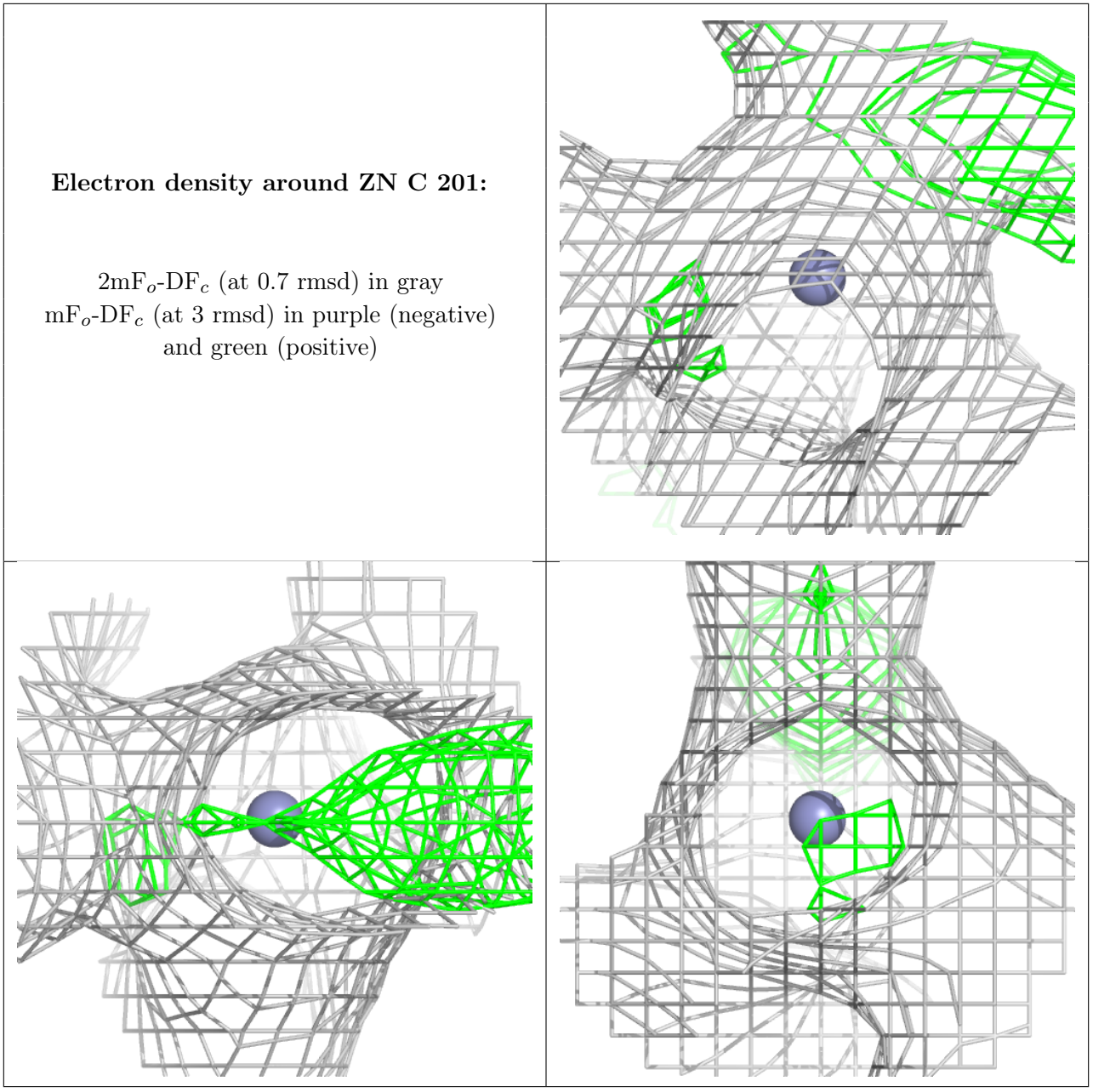






Electron density around ZN C 201:

$2mF_o$-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

There are no such residues in this entry.