wwPDB X-ray Structure Validation Summary Report (i)

Aug 24, 2023 – 04:04 PM EDT

PDB ID	:	8D40
Title	:	Crystal structure of human CELSR1 EC1-4
Authors	:	Tamilselvan, E.; Sotomayor, M.
Deposited on	:	2022-06-01
Resolution	:	3.55 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:

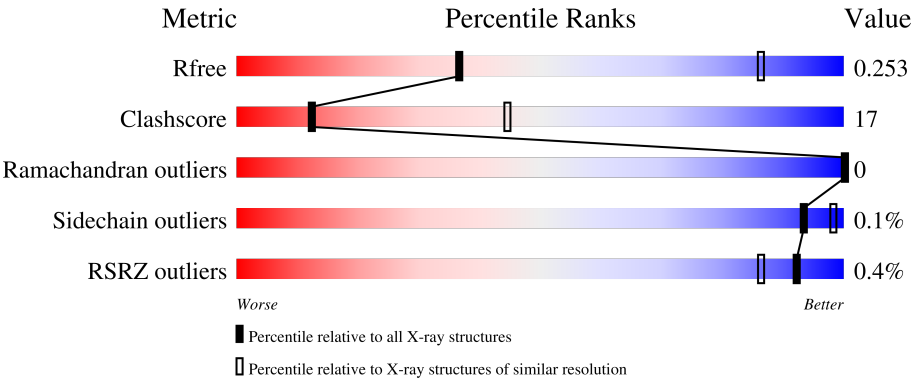
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.55 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	1020 (3.62-3.50)
Clashscore	141614	1100 (3.62-3.50)
Ramachandran outliers	138981	1065 (3.62-3.50)
Sidechain outliers	138945	1066 (3.62-3.50)
RSRZ outliers	127900	1009 (3.64-3.48)

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	457	65% 26% · 8%
1	B	457	64% 29% 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6673 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 Cadherin EGF LAG seven-pass G-type receptor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3307	2066	565	672	4			
1	B	427	Total	C	N	O	S	0	0	0
			3340	2085	571	680	4			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q9NYQ6
A	449	LEU	-	expression tag	UNP Q9NYQ6
A	450	GLU	-	expression tag	UNP Q9NYQ6
A	451	HIS	-	expression tag	UNP Q9NYQ6
A	452	HIS	-	expression tag	UNP Q9NYQ6
A	453	HIS	-	expression tag	UNP Q9NYQ6
A	454	HIS	-	expression tag	UNP Q9NYQ6
A	455	HIS	-	expression tag	UNP Q9NYQ6
A	456	HIS	-	expression tag	UNP Q9NYQ6
B	0	MET	-	initiating methionine	UNP Q9NYQ6
B	449	LEU	-	expression tag	UNP Q9NYQ6
B	450	GLU	-	expression tag	UNP Q9NYQ6
B	451	HIS	-	expression tag	UNP Q9NYQ6
B	452	HIS	-	expression tag	UNP Q9NYQ6
B	453	HIS	-	expression tag	UNP Q9NYQ6
B	454	HIS	-	expression tag	UNP Q9NYQ6
B	455	HIS	-	expression tag	UNP Q9NYQ6
B	456	HIS	-	expression tag	UNP Q9NYQ6

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Na 1 1	0	0
2	B	2	Total Na 2 2	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	8	Total Ca 8 8	0	0
3	B	7	Total Ca 7 7	0	0

- Molecule 4 is water.

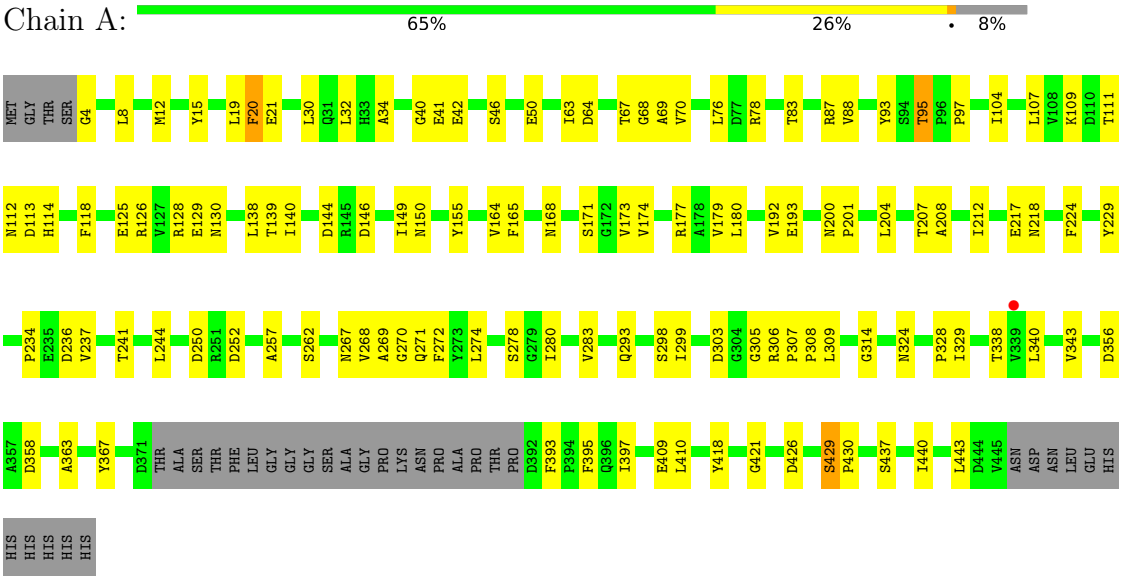
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	4	Total O 4 4	0	0
4	B	4	Total O 4 4	0	0

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.

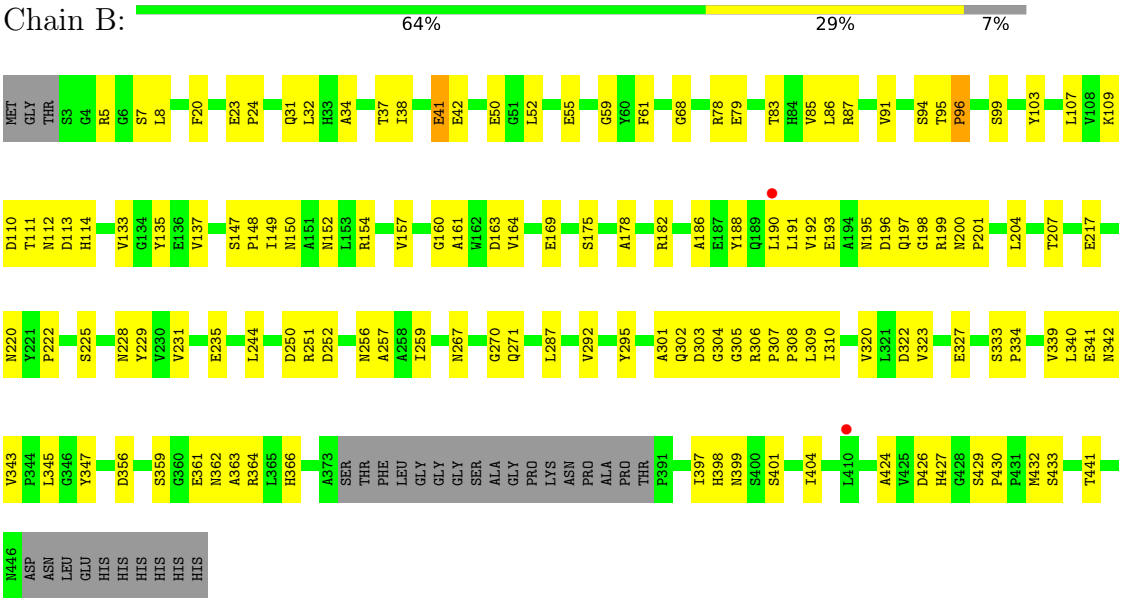
• Molecule 1: Cadherin EGF LAG seven-pass G-type receptor 1

Chain A:



• Molecule 1: Cadherin EGF LAG seven-pass G-type receptor 1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	305.65Å 90.36Å 94.95Å 90.00° 96.17° 90.00°	Depositor
Resolution (Å)	47.24 – 3.55 47.20 – 3.55	Depositor EDS
% Data completeness (in resolution range)	98.2 (47.24-3.55) 98.3 (47.20-3.55)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.49 (at 3.57Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.216 , 0.252 0.217 , 0.253	Depositor DCC
R_{free} test set	1483 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	102.7	Xtriage
Anisotropy	0.690	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 91.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6673	wwPDB-VP
Average B, all atoms (Å^2)	133.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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: CA, NA

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	0.73	0/3375	1.02	3/4603 (0.1%)
1	B	0.73	0/3409	1.06	4/4650 (0.1%)
All	All	0.73	0/6784	1.04	7/9253 (0.1%)

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	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	113	ASP	CB-CA-C	-7.08	96.23	110.40
1	B	96	PRO	N-CA-C	-6.87	94.24	112.10
1	A	293	GLN	CB-CA-C	-6.39	97.62	110.40
1	A	20	PHE	CB-CA-C	6.33	123.05	110.40
1	B	31	GLN	CB-CA-C	5.51	121.42	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	429	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	137	VAL	Peptide
1	B	59	GLY	Peptide

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	3307	0	3152	110	0
1	B	3340	0	3183	110	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	8	0	0	2	0
3	B	7	0	0	0	0
4	A	4	0	0	0	0
4	B	4	0	0	0	0
All	All	6673	0	6335	220	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 17.

The worst 5 of 220 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:B:333:SER:OG	1:B:334:PRO:HD3	1.41	1.17
1:A:19:LEU:HD12	1:A:20:PHE:O	1.48	1.11
1:A:306:ARG:HB2	1:A:307:PRO:HD3	1.21	1.09
1:A:274:LEU:CD2	1:A:299:ILE:HD11	1.81	1.09
1:A:306:ARG:HB2	1:A:307:PRO:CD	1.84	1.05

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	418/457 (92%)	363 (87%)	55 (13%)	0	100	100
1	B	423/457 (93%)	351 (83%)	72 (17%)	0	100	100
All	All	841/914 (92%)	714 (85%)	127 (15%)	0	100	100

There are no Ramachandran outliers to report.

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	368/395 (93%)	368 (100%)	0	100	100
1	B	372/395 (94%)	371 (100%)	1 (0%)	92	97
All	All	740/790 (94%)	739 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	41	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	267	ASN
1	B	255	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	195	ASN
1	B	152	ASN
1	B	197	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 18 ligands modelled in this entry, 18 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

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	422/457 (92%)	-0.01	1 (0%)	95	91	94, 133, 171, 213	0
1	B	427/457 (93%)	0.02	2 (0%)	91	83	78, 130, 172, 219	0
All	All	849/914 (92%)	0.01	3 (0%)	92	86	78, 131, 172, 219	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	339	VAL	2.2
1	B	190	LEU	2.1
1	B	410	LEU	2.1

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
3	CA	B	507	1/1	0.87	0.16	151,151,151,151	0
2	NA	A	501	1/1	0.91	0.15	111,111,111,111	0

Continued on next page...

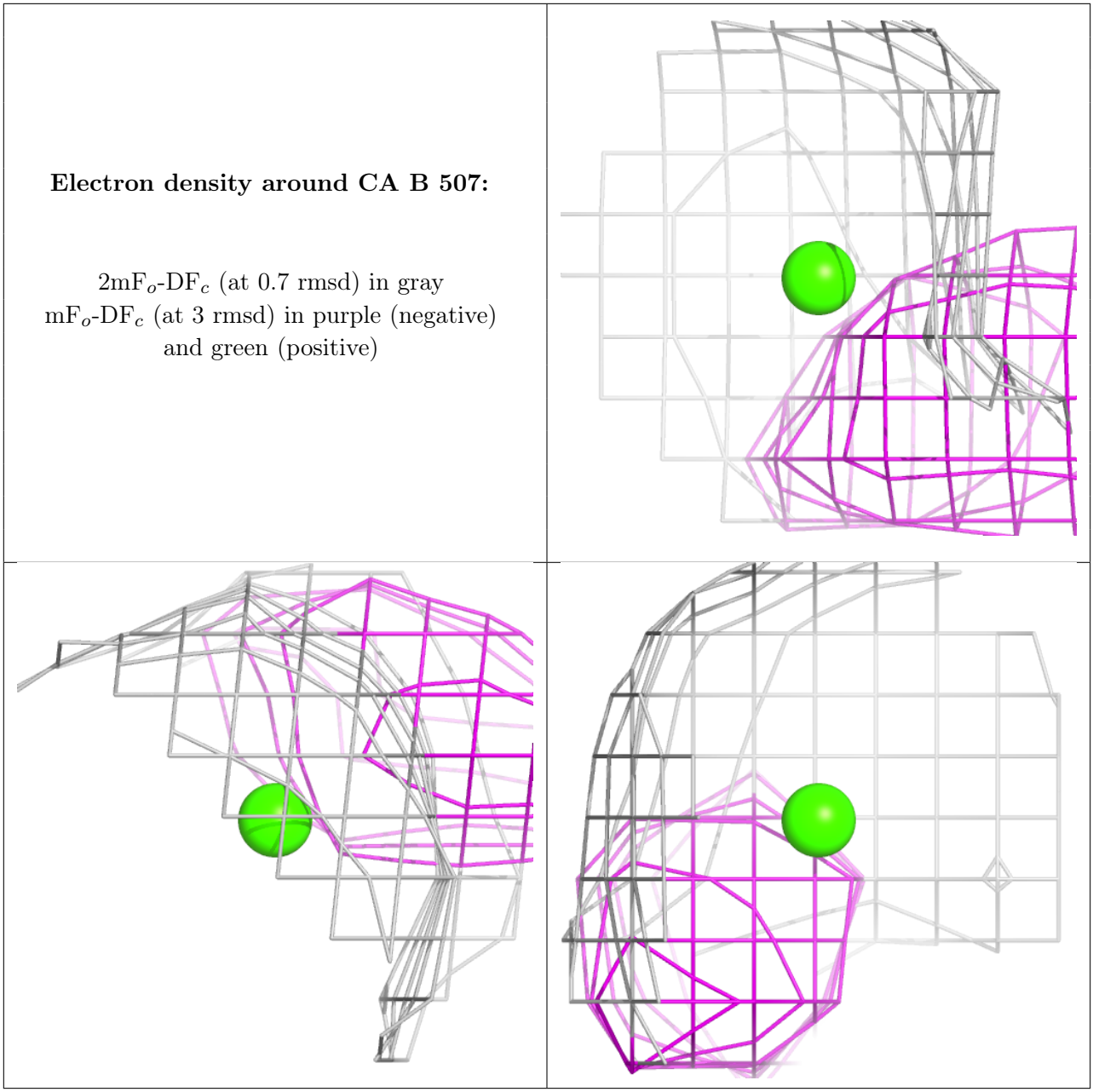
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CA	B	503	1/1	0.92	0.17	125,125,125,125	0
2	NA	B	504	1/1	0.92	0.22	95,95,95,95	0
3	CA	A	502	1/1	0.93	0.18	125,125,125,125	0
2	NA	B	502	1/1	0.94	0.12	103,103,103,103	0
3	CA	A	504	1/1	0.94	0.12	135,135,135,135	0
3	CA	A	507	1/1	0.96	0.06	175,175,175,175	0
3	CA	B	508	1/1	0.97	0.21	86,86,86,86	0
3	CA	A	503	1/1	0.98	0.15	121,121,121,121	0
3	CA	A	508	1/1	0.98	0.13	109,109,109,109	0
3	CA	A	509	1/1	0.98	0.14	110,110,110,110	0
3	CA	B	509	1/1	0.98	0.20	93,93,93,93	0
3	CA	B	505	1/1	0.99	0.17	142,142,142,142	0
3	CA	B	506	1/1	0.99	0.19	122,122,122,122	0
3	CA	A	505	1/1	0.99	0.20	121,121,121,121	0
3	CA	B	501	1/1	0.99	0.17	108,108,108,108	0
3	CA	A	506	1/1	0.99	0.15	122,122,122,122	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.

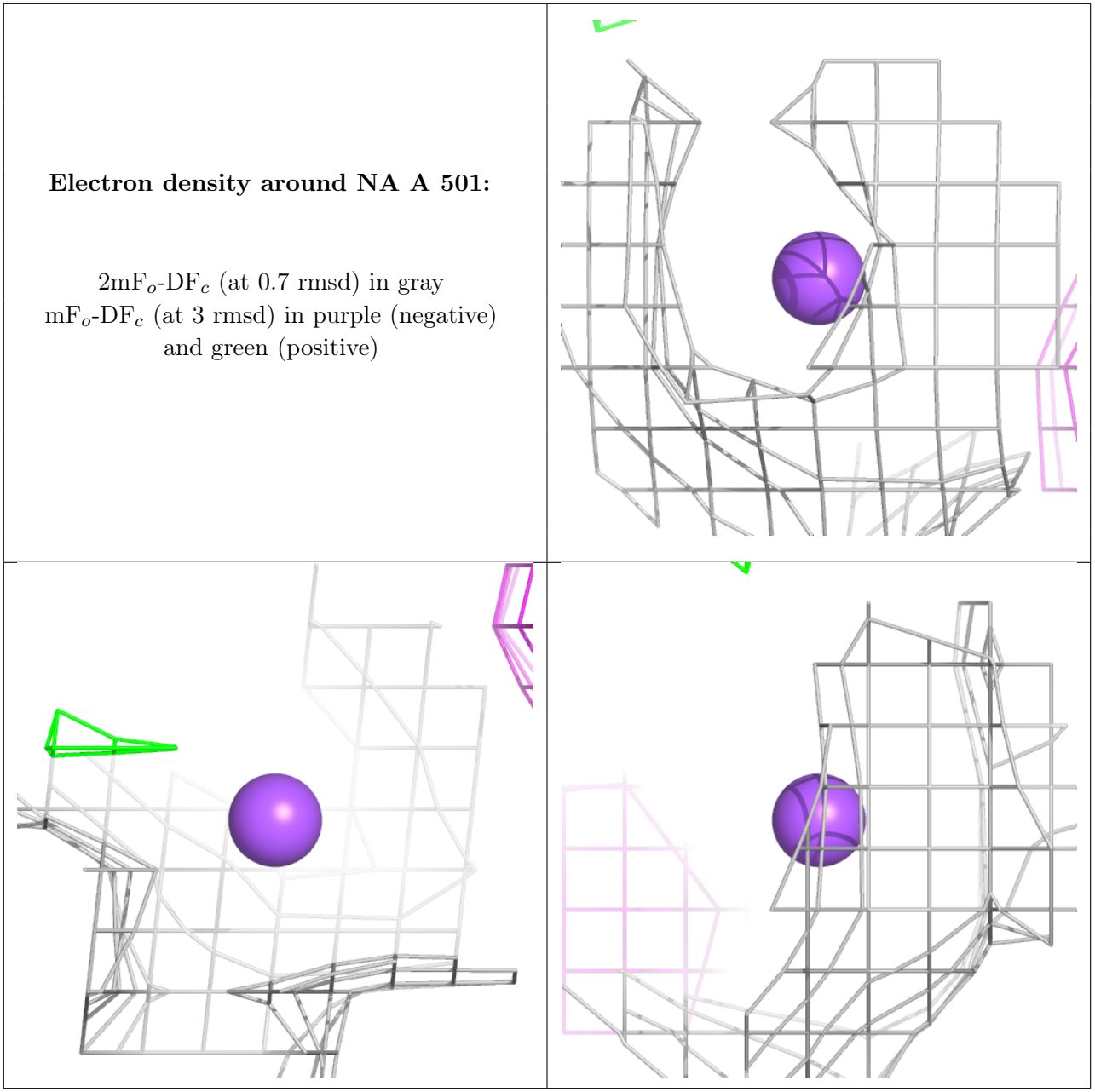
Electron density around CA B 507:

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



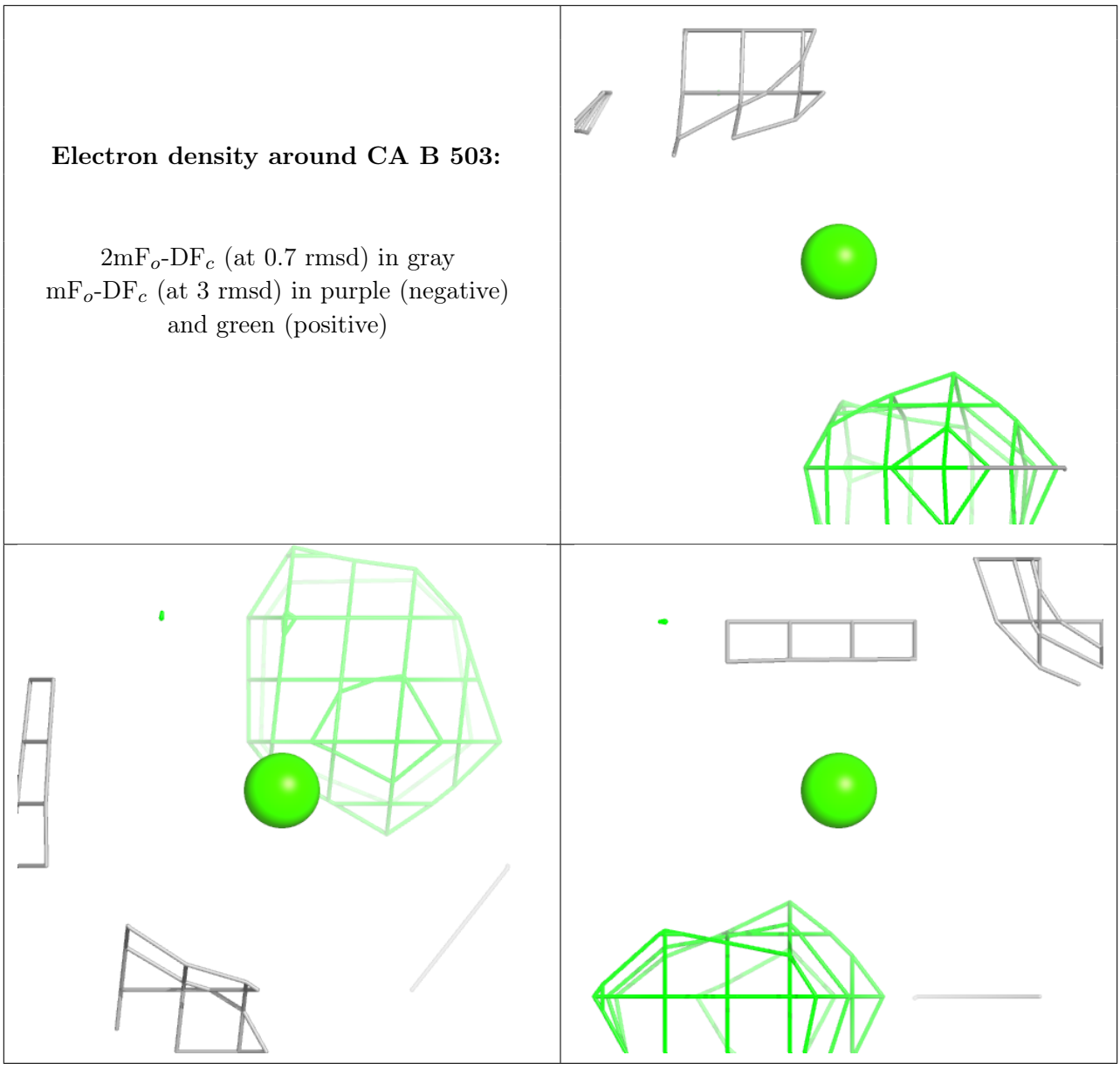
Electron density around NA A 501:

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



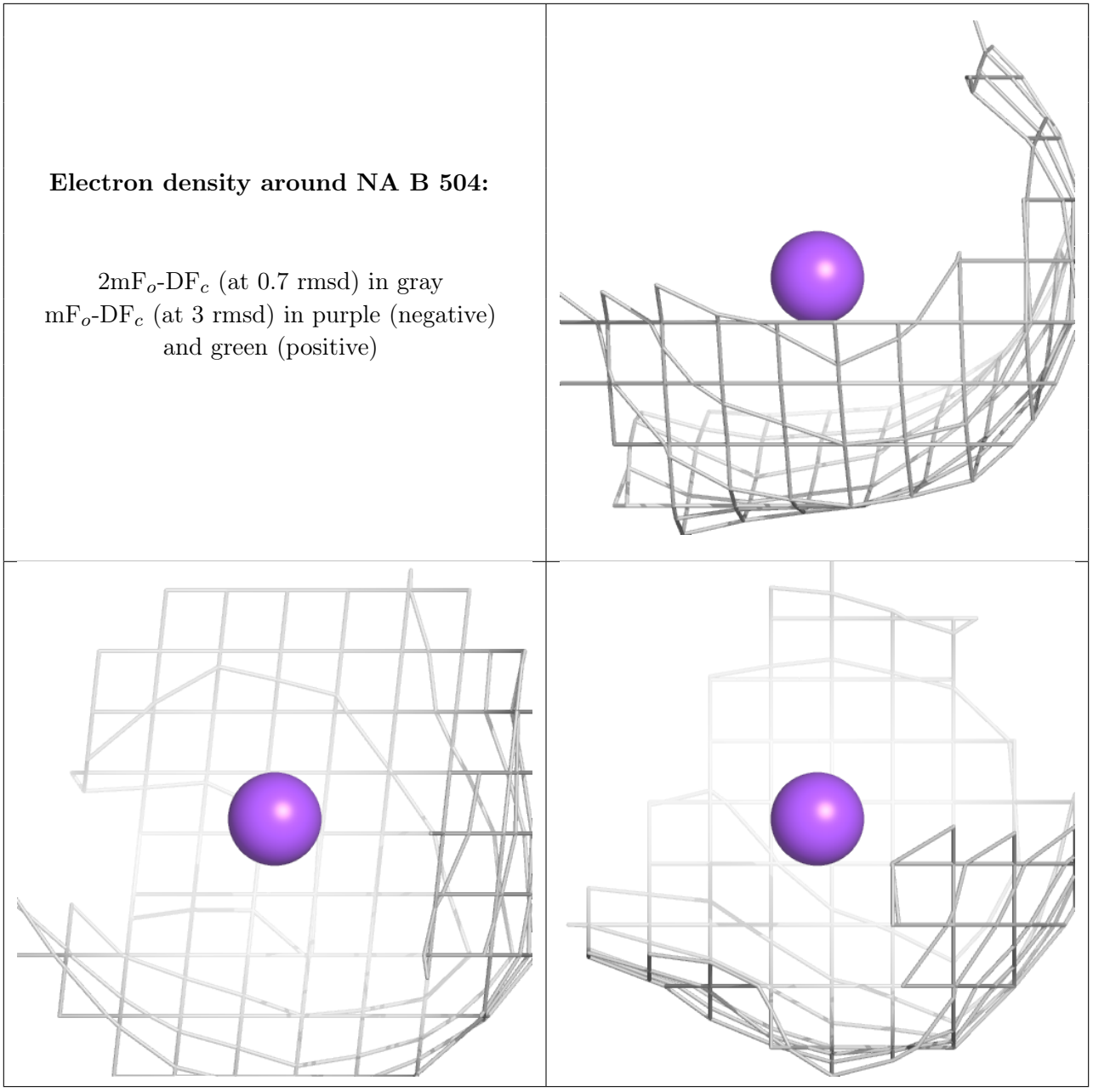
Electron density around CA B 503:

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



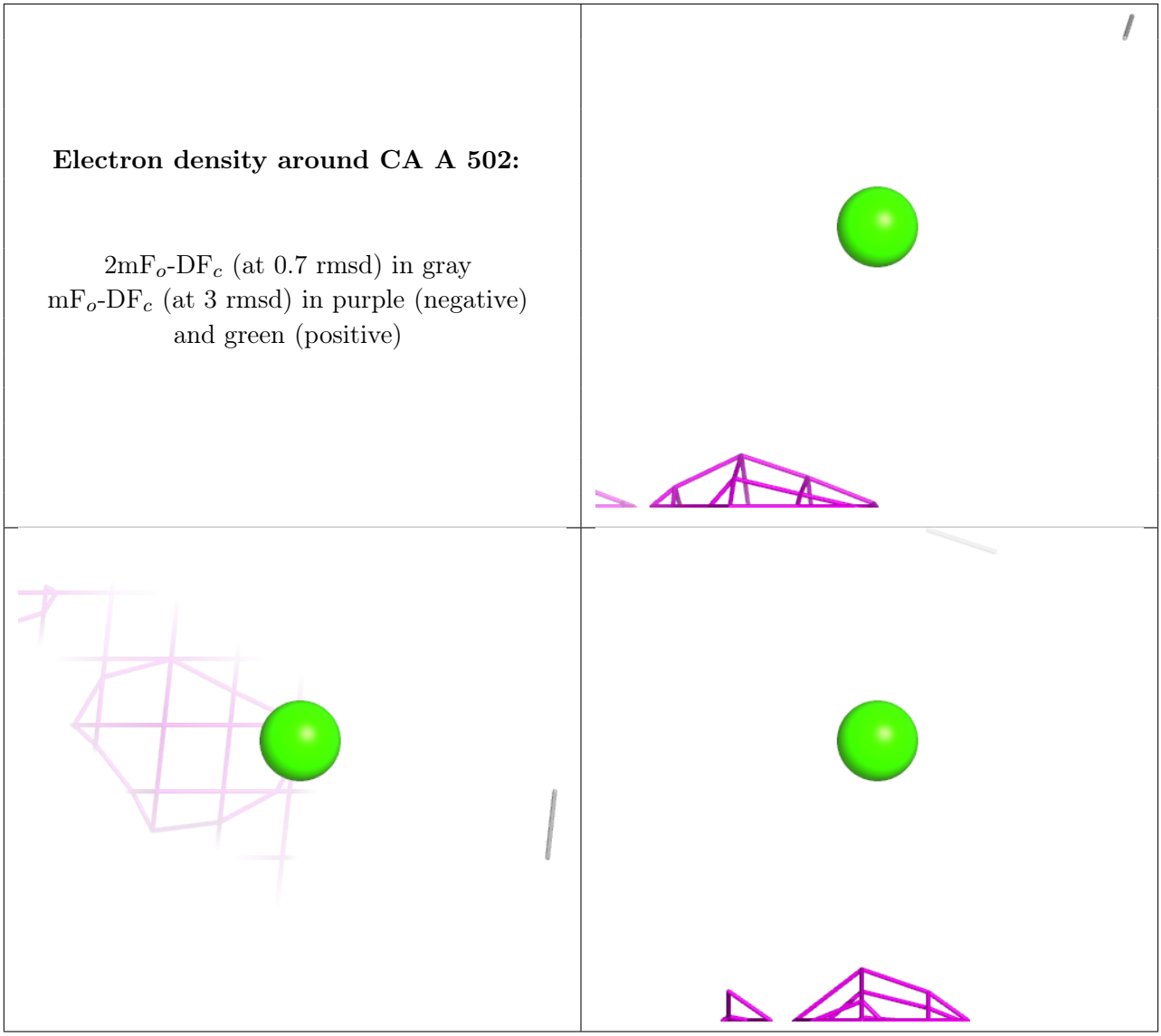
Electron density around NA B 504:

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



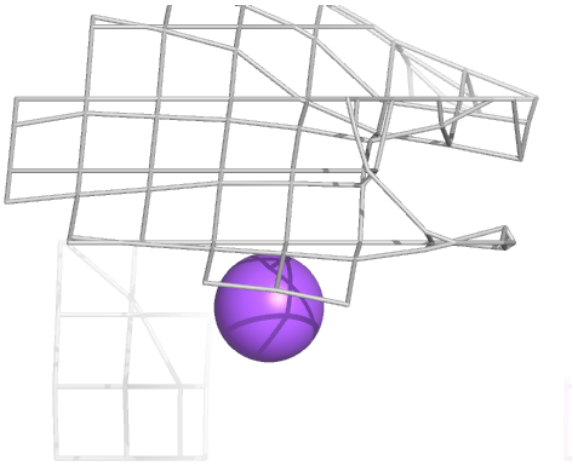
Electron density around CA A 502:

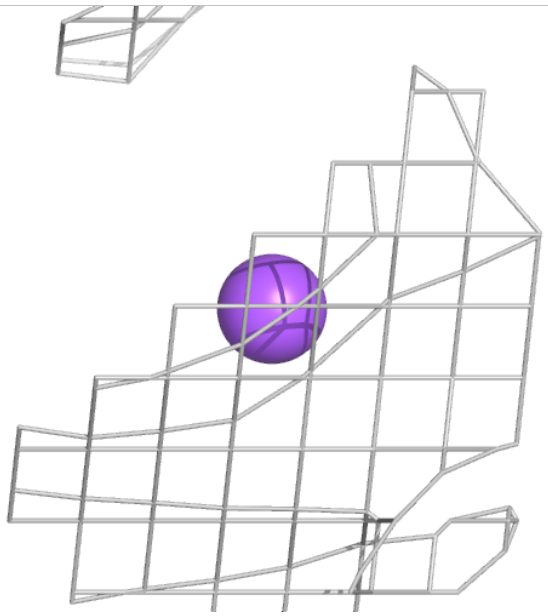
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

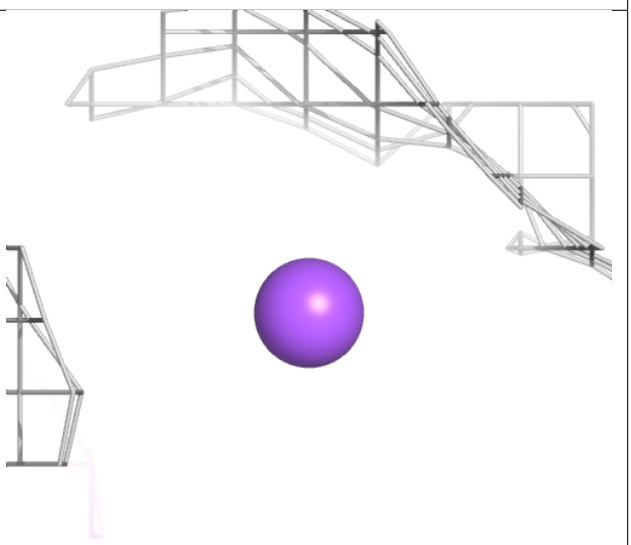


Electron density around NA B 502:

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

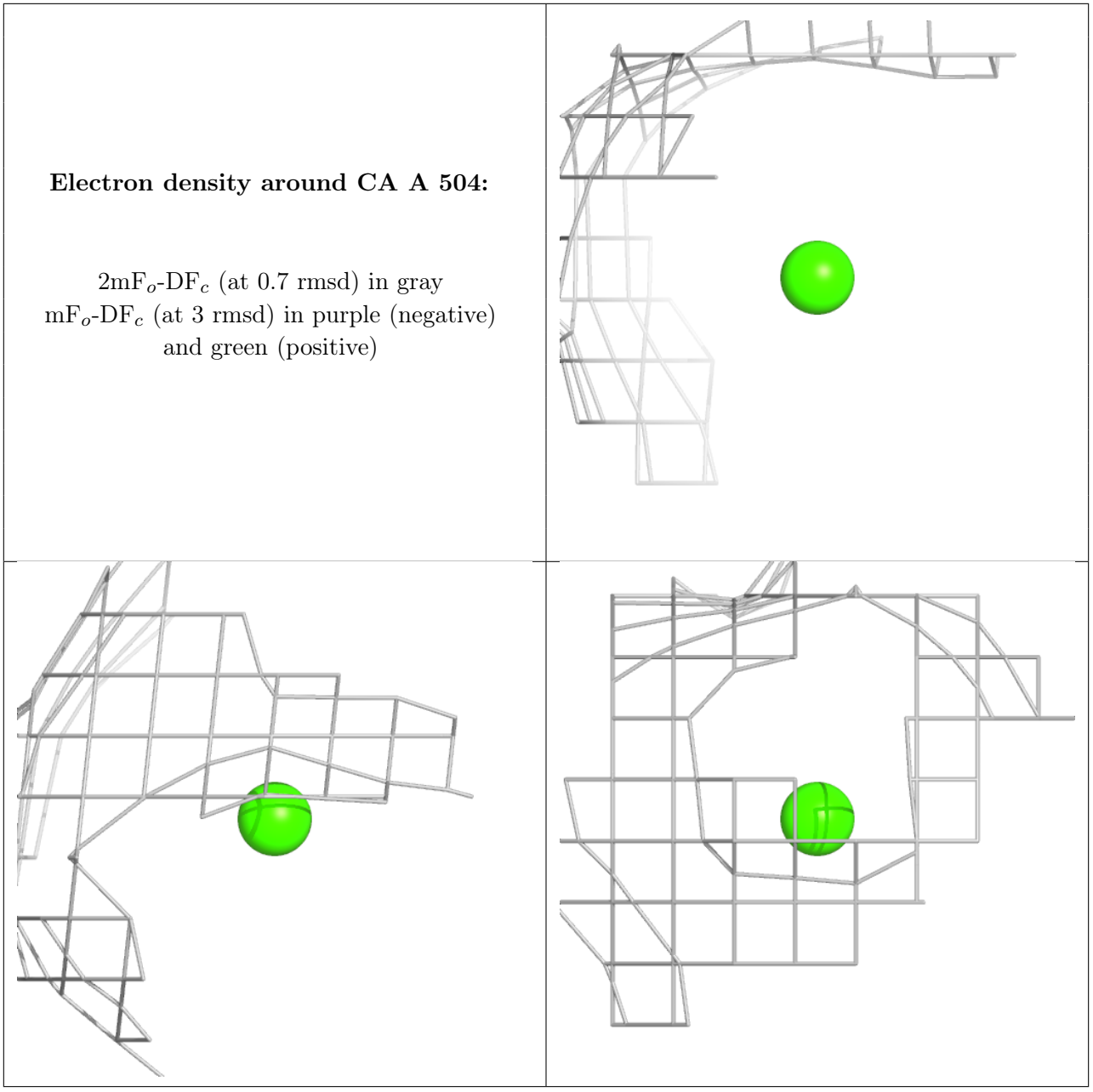






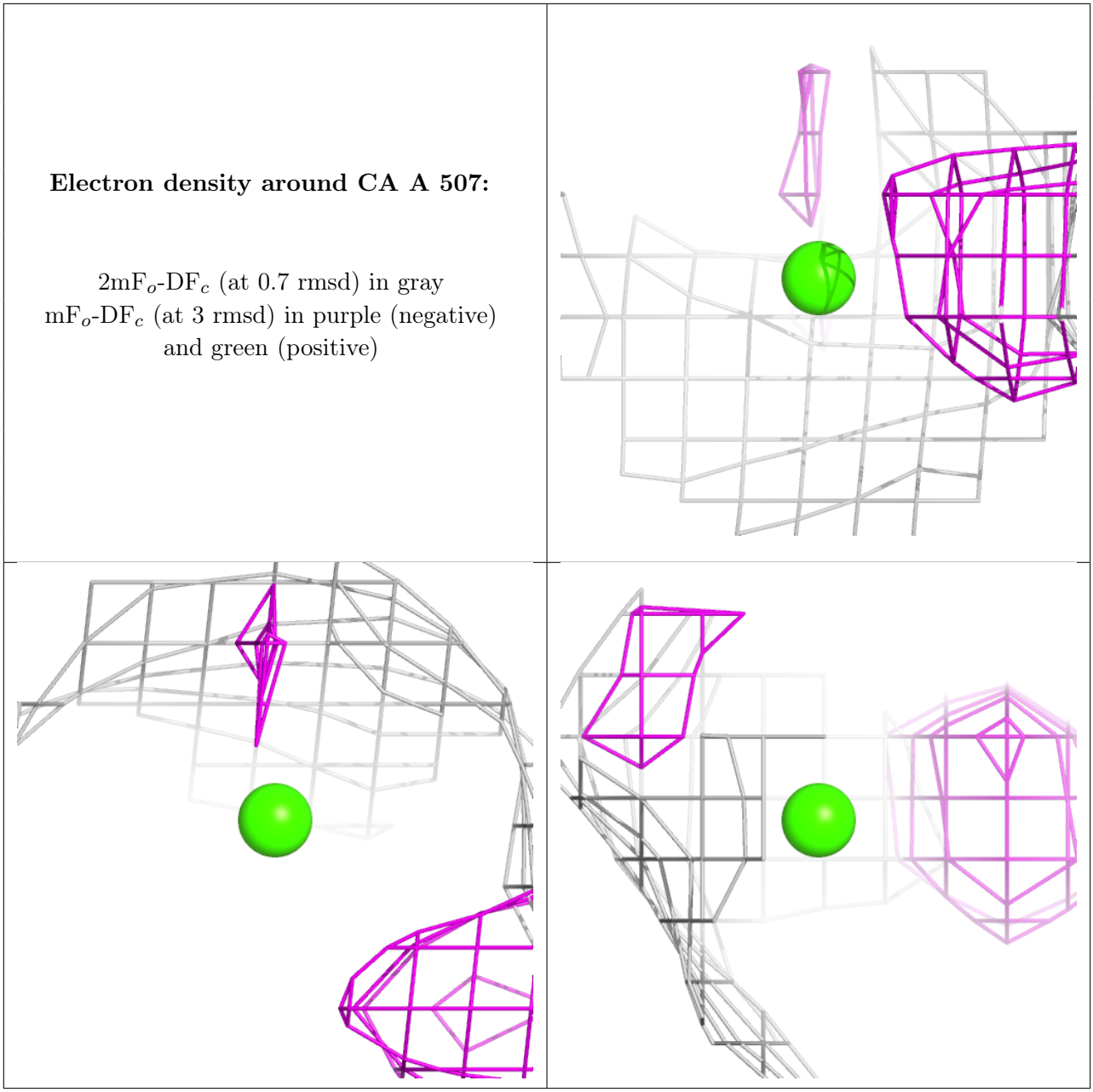
Electron density around CA A 504:

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



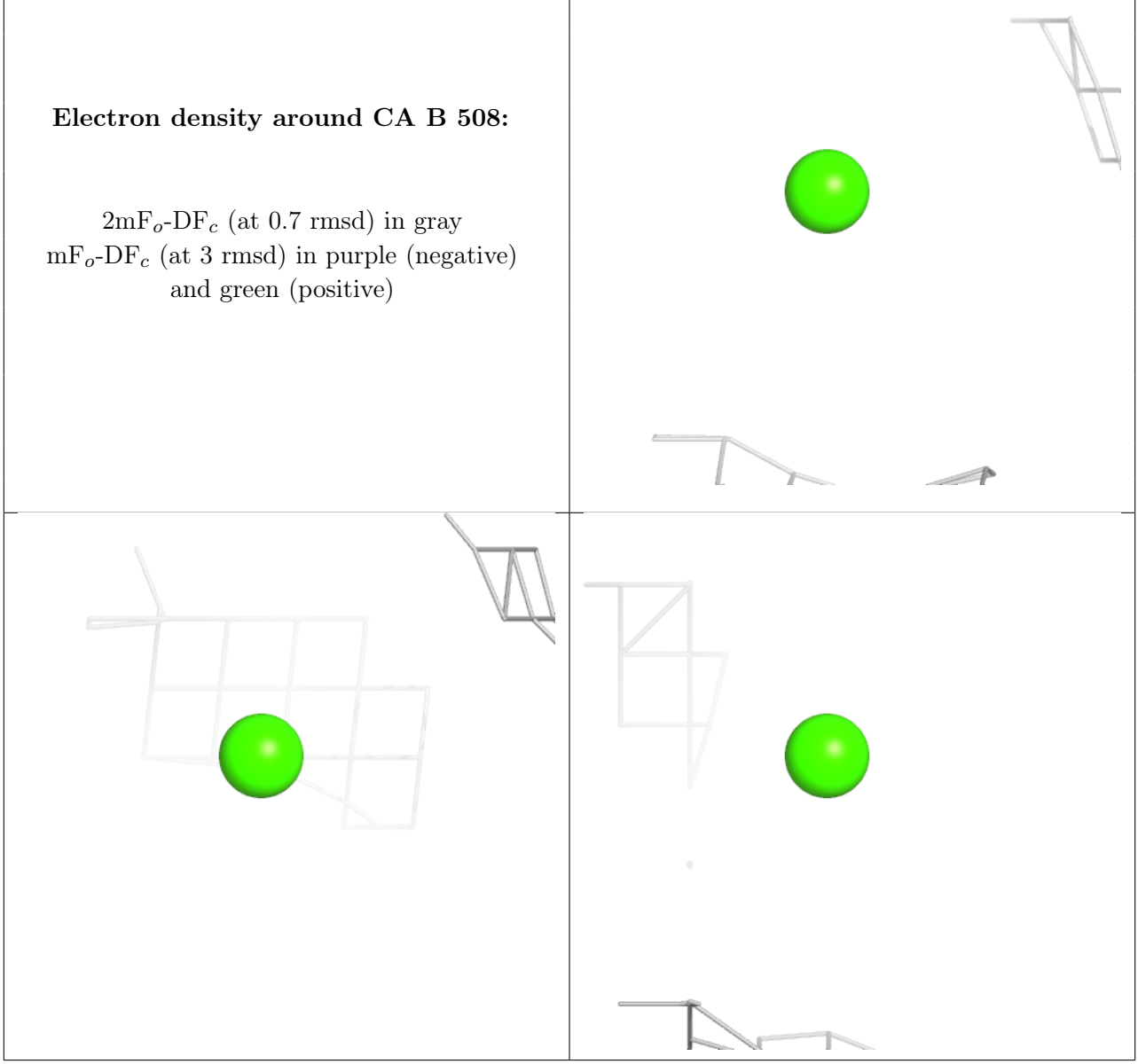
Electron density around CA A 507:

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



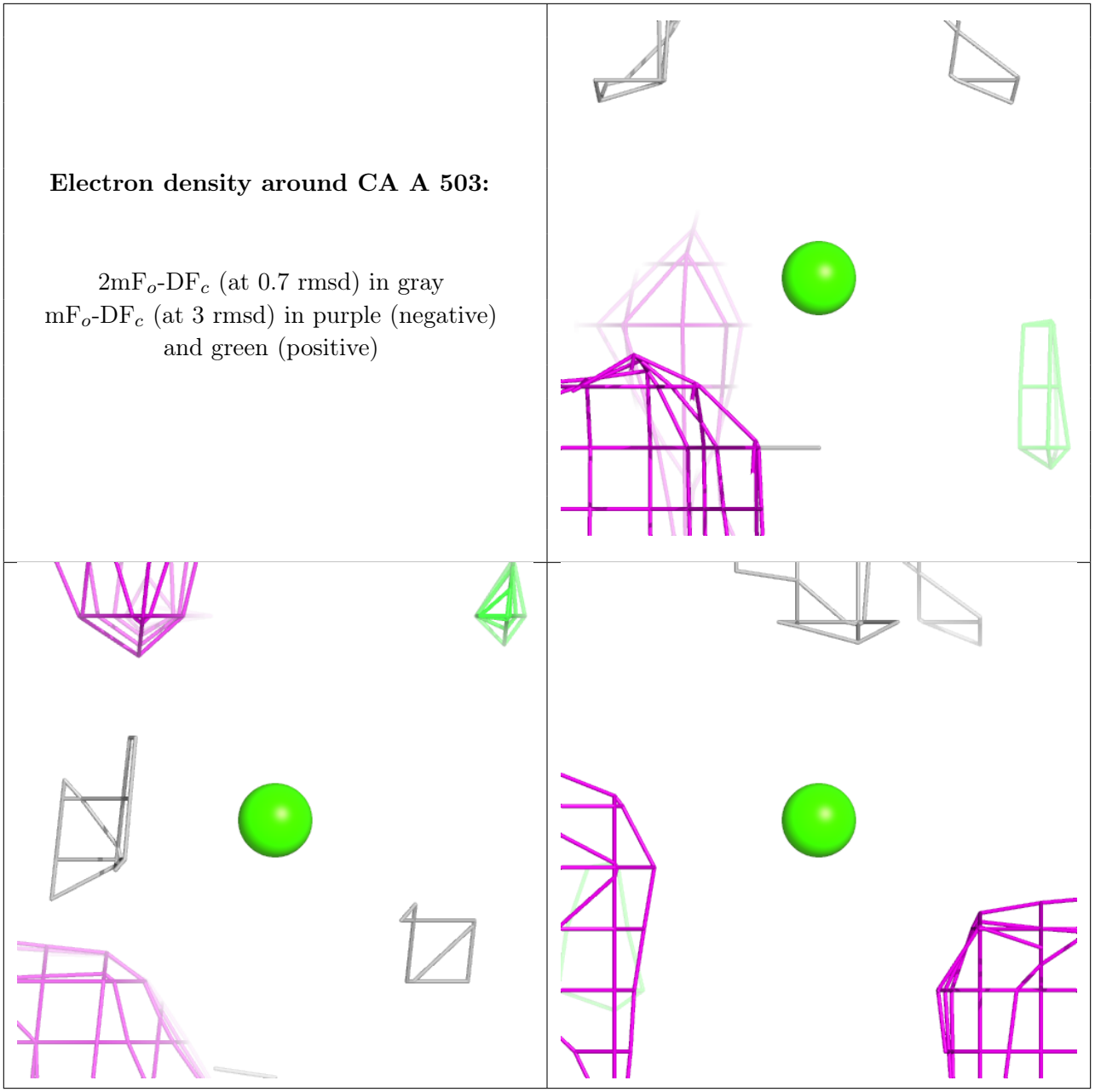
Electron density around CA B 508:

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



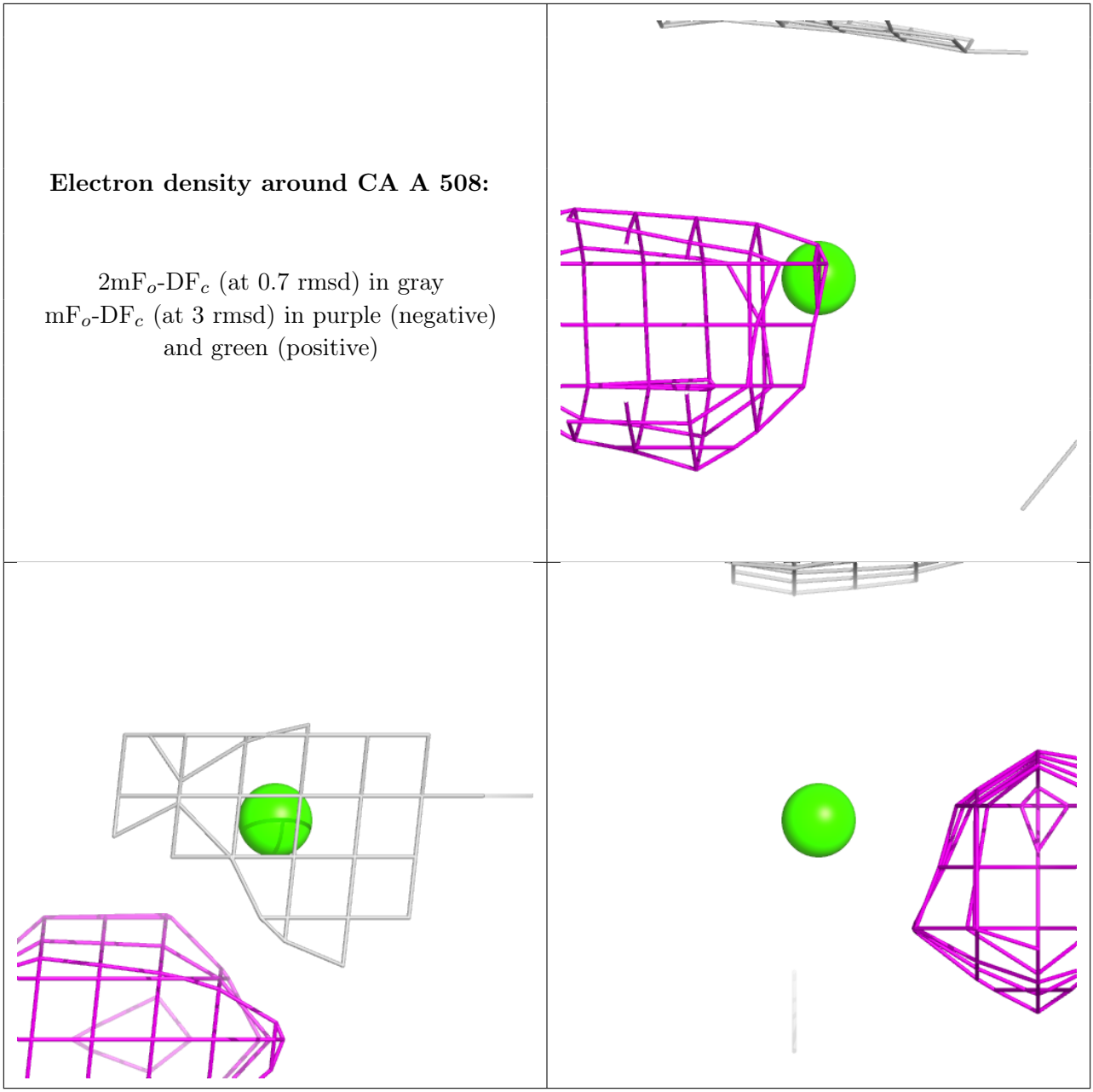
Electron density around CA A 503:

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



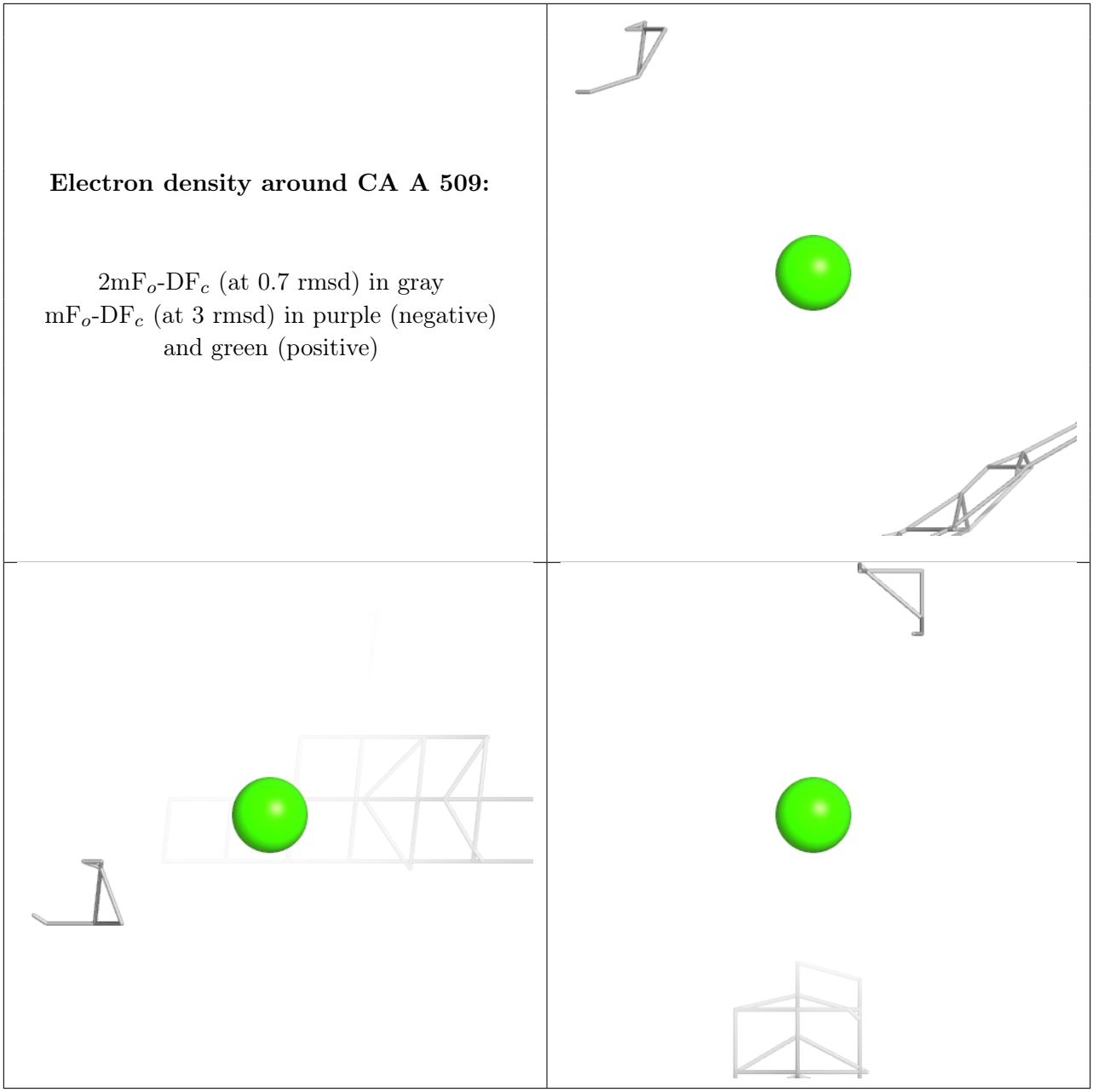
Electron density around CA A 508:

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



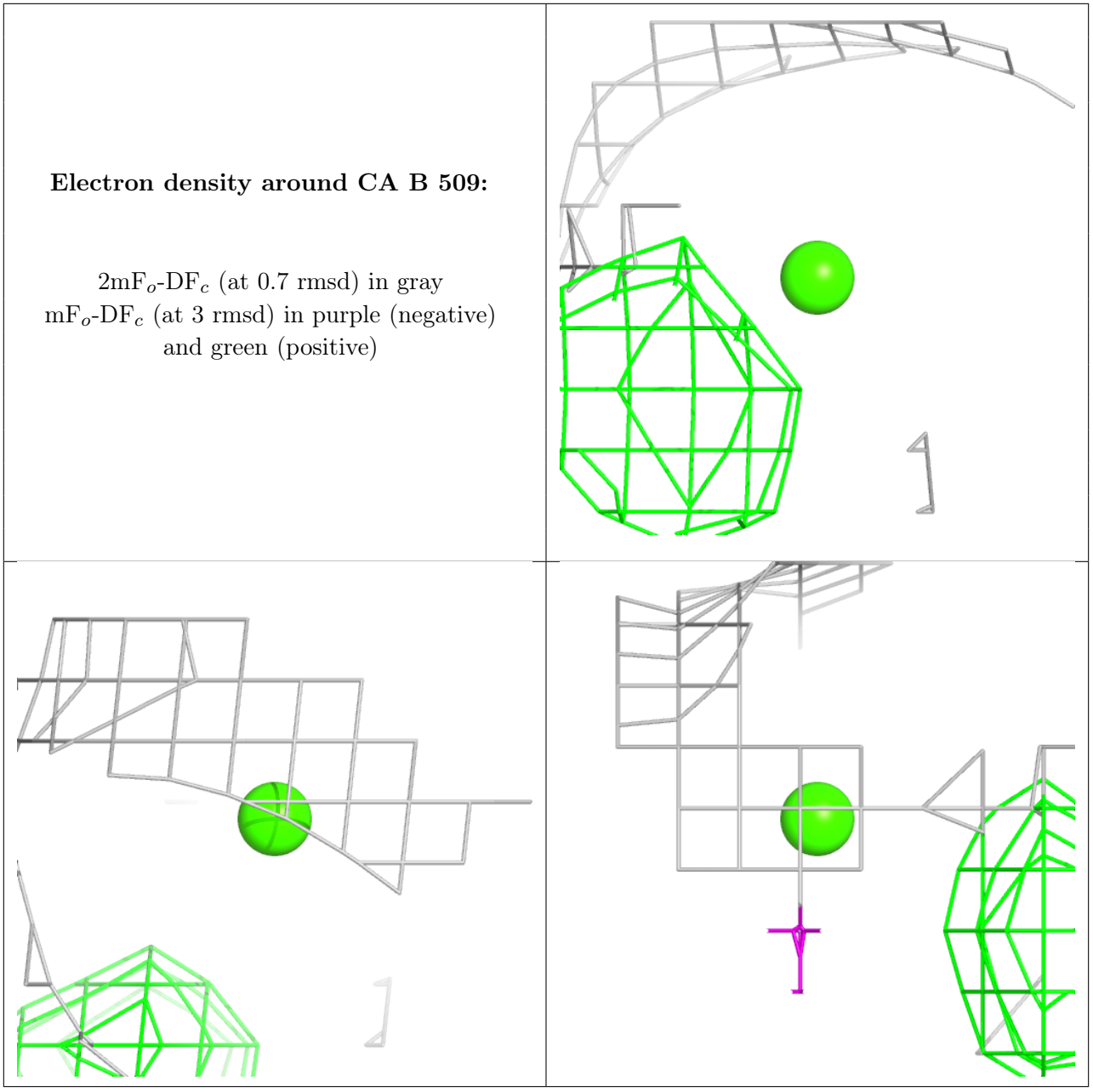
Electron density around CA A 509:

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



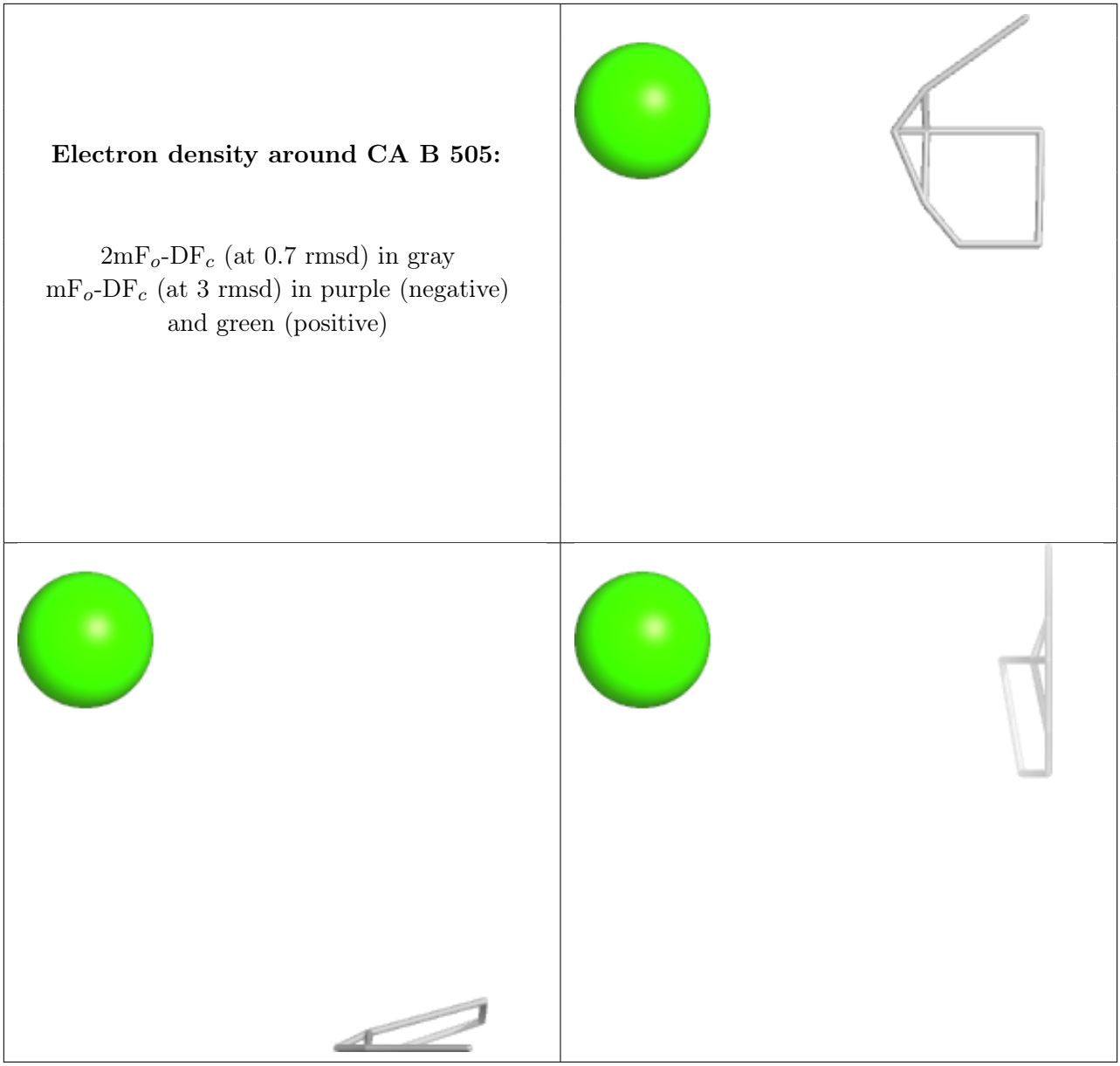
Electron density around CA B 509:

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



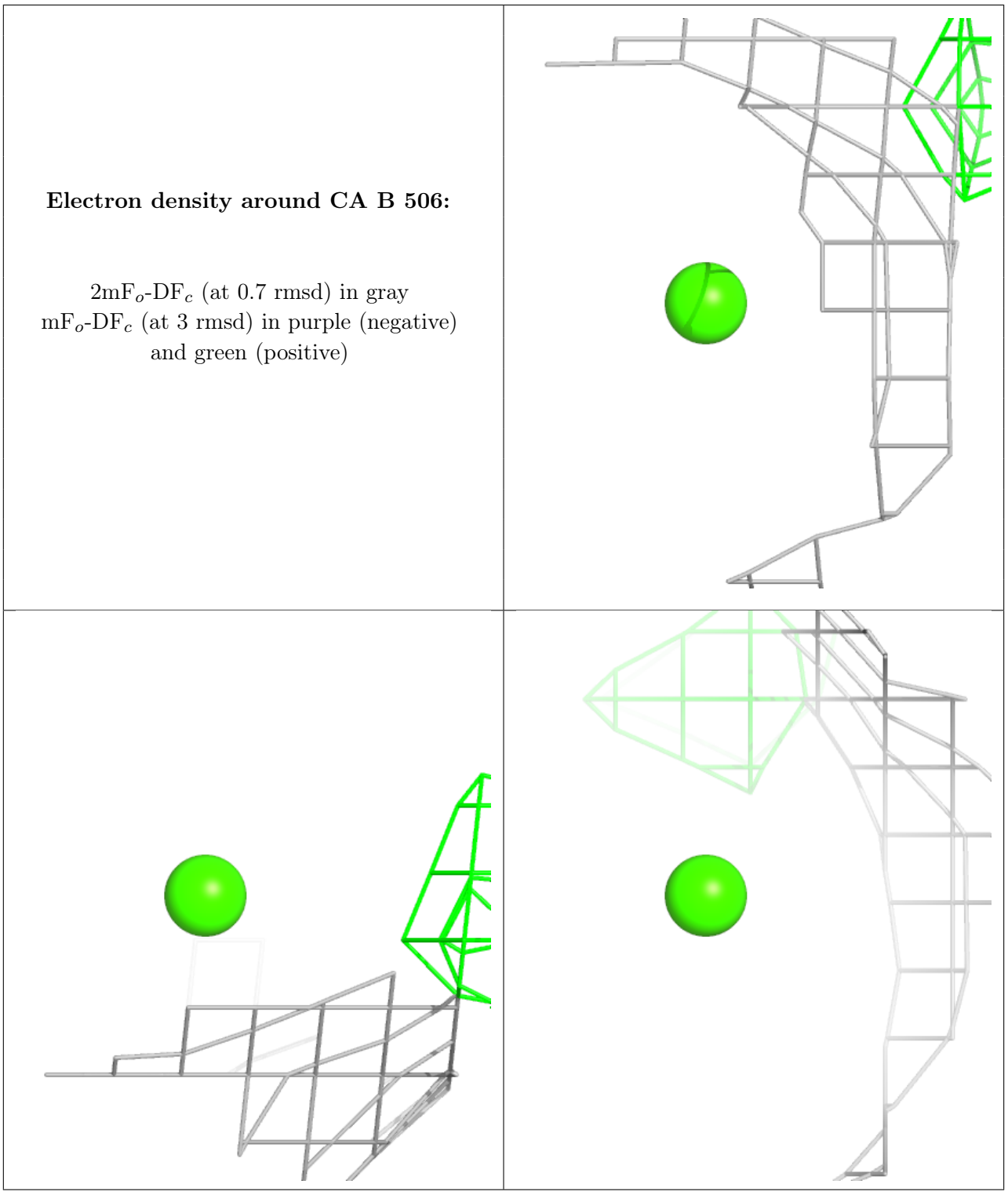
Electron density around CA B 505:

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



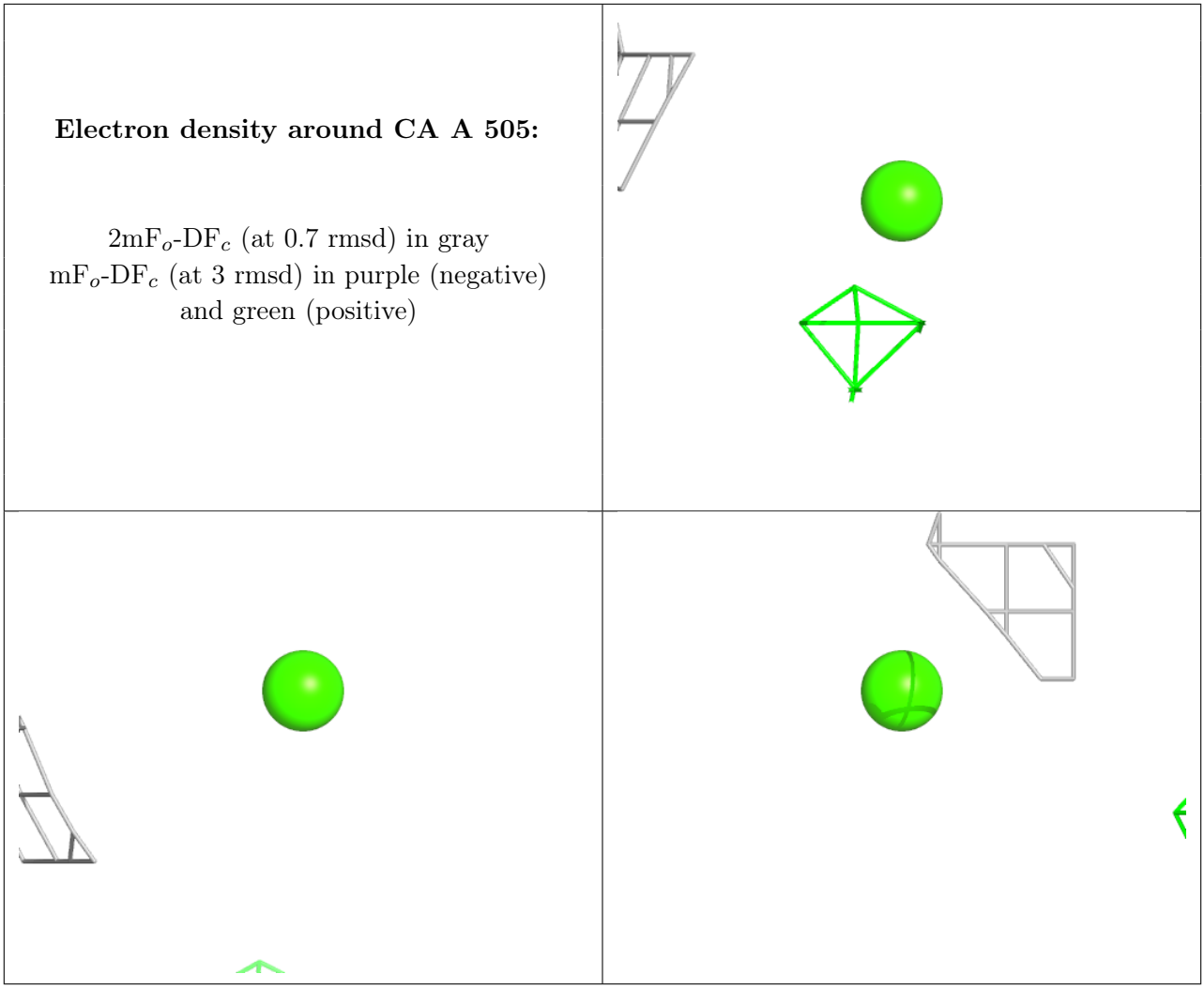
Electron density around CA B 506:

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



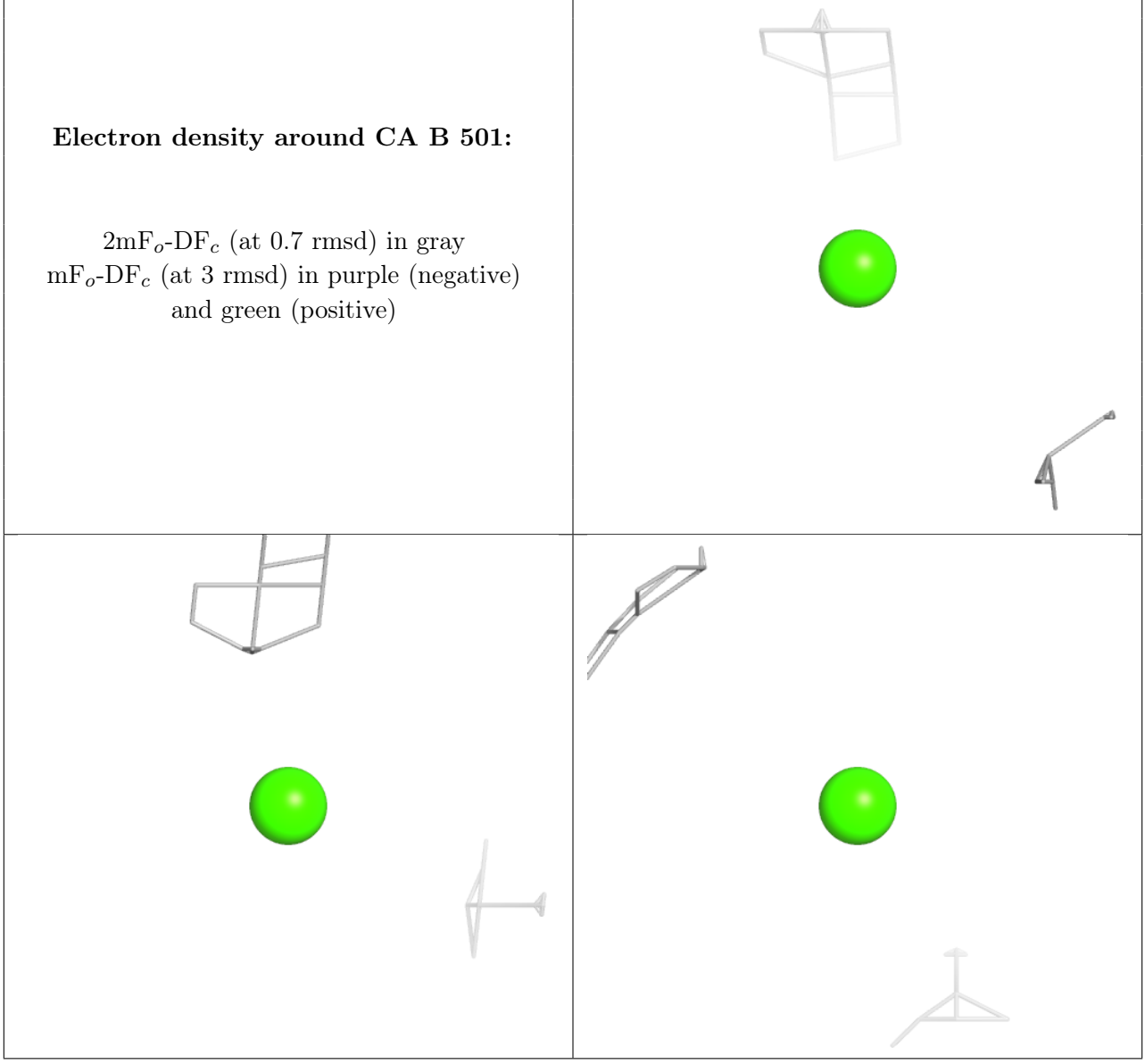
Electron density around CA A 505:

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



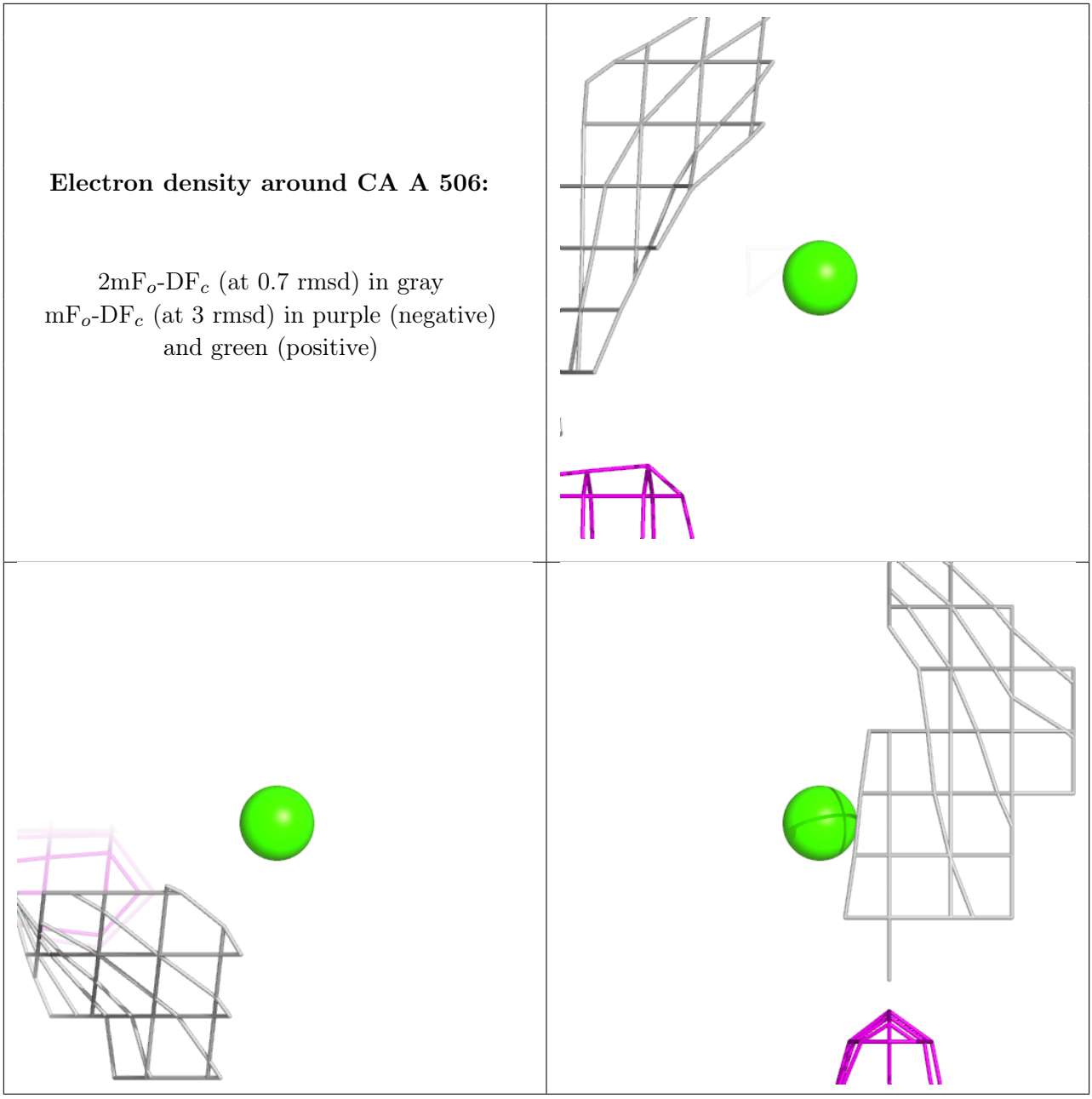
Electron density around CA B 501:

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



Electron density around CA A 506:

$2mF_o$-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.