wwPDB X-ray Structure Validation Summary Report (i)

Aug 19, 2024 – 01:57 PM EDT

PDB ID	:	8TRI
Title	:	Crystal Structure of Mouse Cadherin-23 EC25-MAD28 F2894A
Authors	:	Ashraf, Q.; Sotomayor, M.
Deposited on	:	2023-08-09
Resolution	:	3.72 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 (i)) were used in the production of this report:

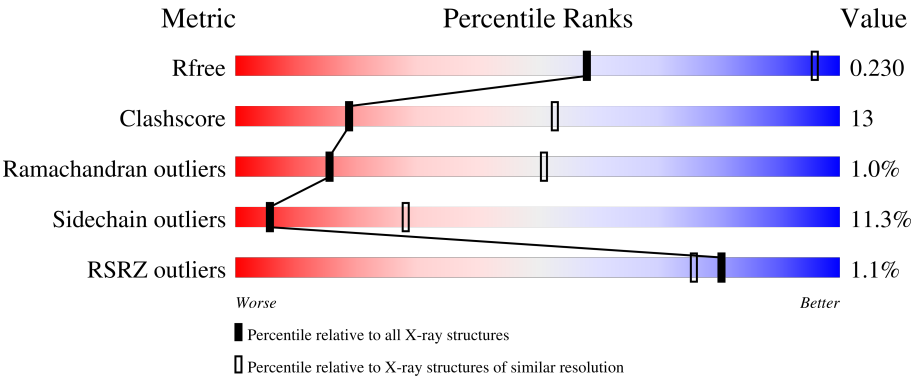
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.72 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	1089 (3.90-3.54)
Clashscore	141614	1012 (3.88-3.56)
Ramachandran outliers	138981	1114 (3.90-3.54)
Sidechain outliers	138945	1110 (3.90-3.54)
RSRZ outliers	127900	1020 (3.90-3.54)

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	474	61% 28% 5% • •
1	B	474	59% 31% 6% • •
1	C	474	3% 55% 29% • • 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10517 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-23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	453	Total	C	N	O	S	0	0	0
			3584	2266	616	697	5			
1	B	454	Total	C	N	O	S	0	0	0
			3590	2269	617	699	5			
1	C	419	Total	C	N	O	S	0	0	0
			3316	2093	572	647	4			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2577	MET	-	initiating methionine	UNP Q99PF4
A	2578	ALA	-	expression tag	UNP Q99PF4
A	2579	SER	-	expression tag	UNP Q99PF4
A	2894	ALA	PHE	engineered mutation	UNP Q99PF4
A	3043	LEU	-	expression tag	UNP Q99PF4
A	3044	GLU	-	expression tag	UNP Q99PF4
A	3045	HIS	-	expression tag	UNP Q99PF4
A	3046	HIS	-	expression tag	UNP Q99PF4
A	3047	HIS	-	expression tag	UNP Q99PF4
A	3048	HIS	-	expression tag	UNP Q99PF4
A	3049	HIS	-	expression tag	UNP Q99PF4
A	3050	HIS	-	expression tag	UNP Q99PF4
B	2577	MET	-	initiating methionine	UNP Q99PF4
B	2578	ALA	-	expression tag	UNP Q99PF4
B	2579	SER	-	expression tag	UNP Q99PF4
B	2894	ALA	PHE	engineered mutation	UNP Q99PF4
B	3043	LEU	-	expression tag	UNP Q99PF4
B	3044	GLU	-	expression tag	UNP Q99PF4
B	3045	HIS	-	expression tag	UNP Q99PF4
B	3046	HIS	-	expression tag	UNP Q99PF4
B	3047	HIS	-	expression tag	UNP Q99PF4
B	3048	HIS	-	expression tag	UNP Q99PF4
B	3049	HIS	-	expression tag	UNP Q99PF4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	3050	HIS	-	expression tag	UNP Q99PF4
C	2577	MET	-	initiating methionine	UNP Q99PF4
C	2578	ALA	-	expression tag	UNP Q99PF4
C	2579	SER	-	expression tag	UNP Q99PF4
C	2894	ALA	PHE	engineered mutation	UNP Q99PF4
C	3043	LEU	-	expression tag	UNP Q99PF4
C	3044	GLU	-	expression tag	UNP Q99PF4
C	3045	HIS	-	expression tag	UNP Q99PF4
C	3046	HIS	-	expression tag	UNP Q99PF4
C	3047	HIS	-	expression tag	UNP Q99PF4
C	3048	HIS	-	expression tag	UNP Q99PF4
C	3049	HIS	-	expression tag	UNP Q99PF4
C	3050	HIS	-	expression tag	UNP Q99PF4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	6	Total Ca 6 6	0	0
2	B	6	Total Ca 6 6	0	0
2	C	6	Total Ca 6 6	0	0

- Molecule 3 is water.

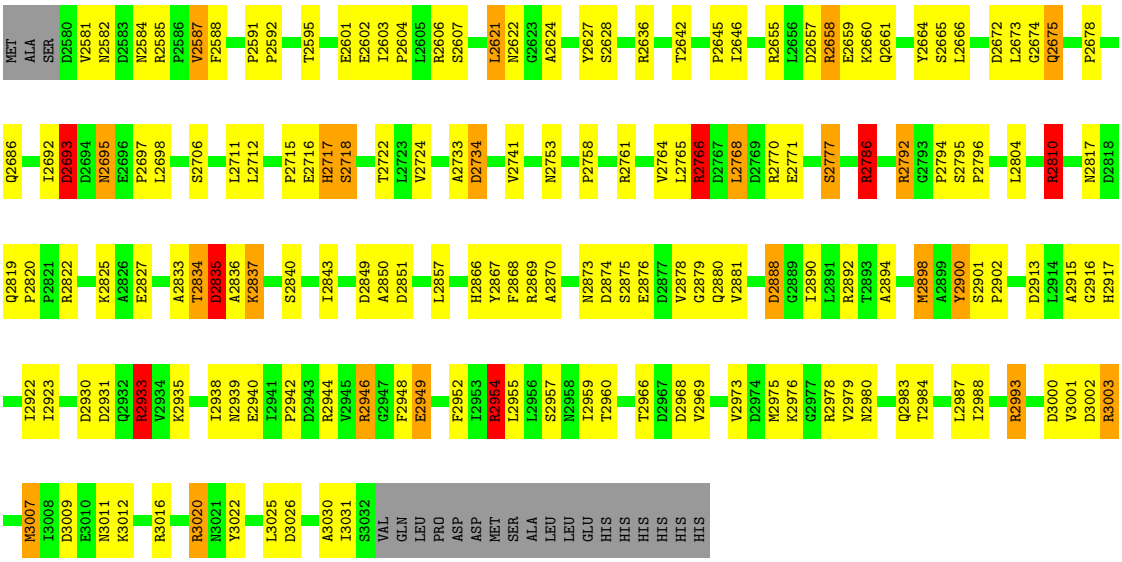
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	7	Total O 7 7	0	0
3	B	2	Total O 2 2	0	0

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.

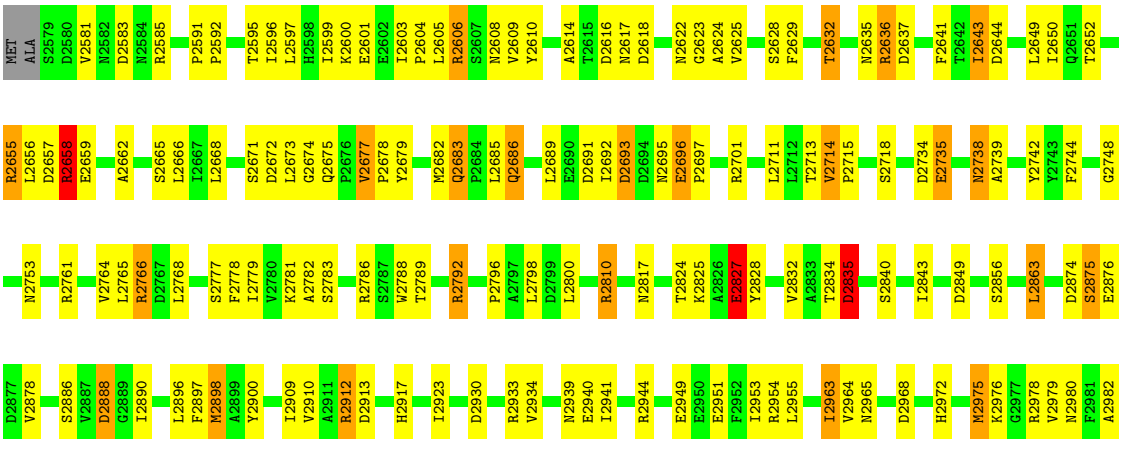
- Molecule 1: Cadherin-23

Chain A:



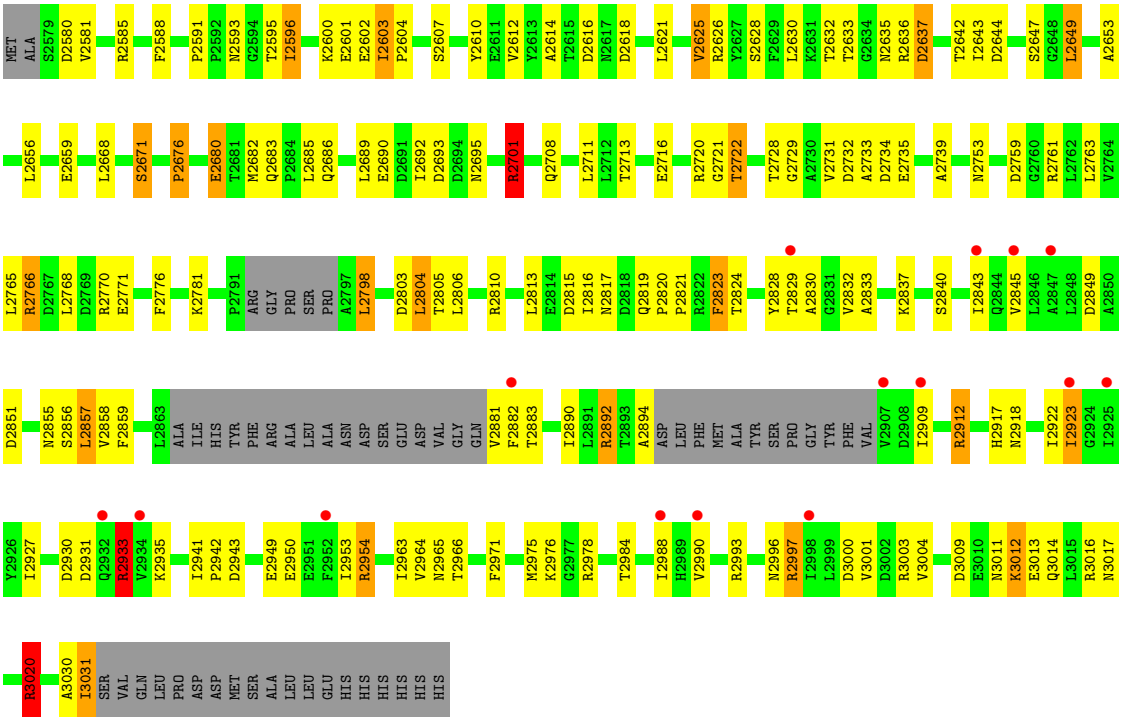
- Molecule 1: Cadherin-23

Chain B:



• Molecule 1: Cadherin-23

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	146.99Å 164.66Å 183.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.56 – 3.72 49.51 – 3.72	Depositor EDS
% Data completeness (in resolution range)	98.7 (49.56-3.72) 98.7 (49.51-3.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 3.67Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, R_{free}	0.187 , 0.229 0.192 , 0.230	Depositor DCC
R_{free} test set	2329 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	98.8	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 92.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10517	wwPDB-VP
Average B, all atoms (Å^2)	126.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.10 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1709e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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	0.84	4/3654 (0.1%)	0.91	7/4982 (0.1%)
1	B	0.90	7/3660 (0.2%)	0.91	8/4990 (0.2%)
1	C	0.79	4/3373 (0.1%)	0.83	1/4594 (0.0%)
All	All	0.85	15/10687 (0.1%)	0.89	16/14566 (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	18
1	B	0	20
1	C	0	17
All	All	0	55

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	2949	GLU	CD-OE1	10.64	1.37	1.25
1	C	2680	GLU	CD-OE1	9.03	1.35	1.25
1	B	3013	GLU	CD-OE1	8.23	1.34	1.25
1	A	2827	GLU	CD-OE2	8.19	1.34	1.25
1	B	2827	GLU	CD-OE2	7.83	1.34	1.25

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2949	GLU	CB-CA-C	-8.80	92.80	110.40
1	A	2933	ARG	NE-CZ-NH2	-7.45	116.58	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2834	THR	CA-CB-OG1	-7.39	93.48	109.00
1	A	2761	ARG	NE-CZ-NH1	-6.46	117.07	120.30
1	A	2655	ARG	NE-CZ-NH1	-6.22	117.19	120.30

There are no chirality outliers.

5 of 55 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2585	ARG	Sidechain
1	A	2606	ARG	Sidechain
1	A	2636	ARG	Sidechain
1	A	2658	ARG	Sidechain
1	A	2675	GLN	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	3584	0	3527	92	0
1	B	3590	0	3532	92	0
1	C	3316	0	3280	94	0
2	A	6	0	0	0	0
2	B	6	0	0	0	0
2	C	6	0	0	0	0
3	A	7	0	0	1	0
3	B	2	0	0	0	0
All	All	10517	0	10339	274	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 13.

The worst 5 of 274 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:2657:ASP:OD1	1:B:2659:GLU:HG2	1.66	0.95

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:2603:ILE:HB	1:C:2604:PRO:HD2	1.52	0.91
1:C:2817:ASN:ND2	1:C:2855:ASN:HA	1.91	0.86
1:A:2604:PRO:HG2	1:A:2607:SER:HB3	1.59	0.83
1:C:2857:LEU:C	1:C:2857:LEU:HD13	1.98	0.83

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	451/474 (95%)	405 (90%)	41 (9%)	5 (1%)	14	50
1	B	452/474 (95%)	402 (89%)	47 (10%)	3 (1%)	22	59
1	C	411/474 (87%)	375 (91%)	31 (8%)	5 (1%)	13	48
All	All	1314/1422 (92%)	1182 (90%)	119 (9%)	13 (1%)	15	51

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2693	ASP
1	A	2880	GLN
1	B	2835	ASP
1	A	2717	HIS
1	C	2759	ASP

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	395/414 (95%)	349 (88%)	46 (12%)	5	27
1	B	396/414 (96%)	350 (88%)	46 (12%)	5	27
1	C	368/414 (89%)	329 (89%)	39 (11%)	6	30
All	All	1159/1242 (93%)	1028 (89%)	131 (11%)	6	28

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	2857	LEU
1	C	2923	ILE
1	C	3031	ILE
1	B	2632	THR
1	B	2605	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	2635	ASN
1	C	2654	GLN
1	C	2918	ASN
1	C	2683	GLN
1	B	2683	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 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	453/474 (95%)	-0.37	0 100 100	64, 106, 171, 212	0
1	B	454/474 (95%)	-0.35	0 100 100	68, 113, 166, 234	0
1	C	419/474 (88%)	0.14	15 (3%) 42 35	73, 147, 222, 269	0
All	All	1326/1422 (93%)	-0.20	15 (1%) 80 76	64, 117, 200, 269	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2932	GLN	3.8
1	C	2909	ILE	3.7
1	C	2923	ILE	3.5
1	C	2998	ILE	3.3
1	C	2988	ILE	3.1

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

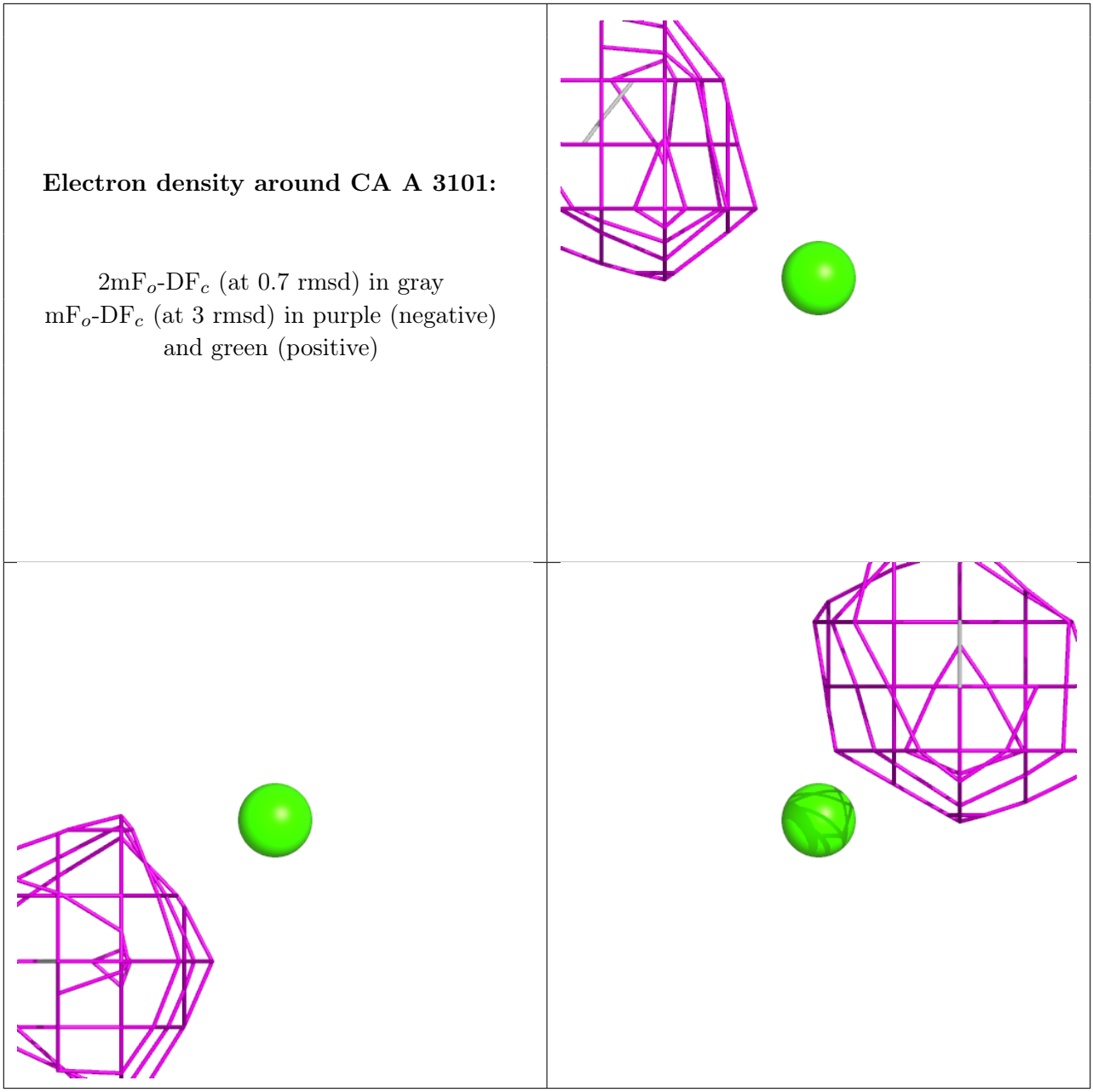
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	CA	A	3101	1/1	0.95	0.09	82,82,82,82	0
2	CA	C	3102	1/1	0.95	0.08	110,110,110,110	0
2	CA	B	3101	1/1	0.96	0.09	116,116,116,116	0
2	CA	C	3106	1/1	0.96	0.06	154,154,154,154	0
2	CA	C	3104	1/1	0.97	0.07	96,96,96,96	0
2	CA	C	3103	1/1	0.97	0.07	115,115,115,115	0
2	CA	B	3105	1/1	0.98	0.07	105,105,105,105	0
2	CA	B	3106	1/1	0.98	0.13	94,94,94,94	0
2	CA	A	3105	1/1	0.98	0.08	111,111,111,111	0
2	CA	A	3104	1/1	0.98	0.13	87,87,87,87	0
2	CA	B	3102	1/1	0.98	0.07	122,122,122,122	0
2	CA	B	3104	1/1	0.98	0.10	102,102,102,102	0
2	CA	C	3101	1/1	0.99	0.14	70,70,70,70	0
2	CA	B	3103	1/1	0.99	0.09	118,118,118,118	0
2	CA	A	3106	1/1	0.99	0.07	95,95,95,95	0
2	CA	A	3102	1/1	0.99	0.07	103,103,103,103	0
2	CA	C	3105	1/1	0.99	0.05	140,140,140,140	0
2	CA	A	3103	1/1	0.99	0.14	84,84,84,84	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.

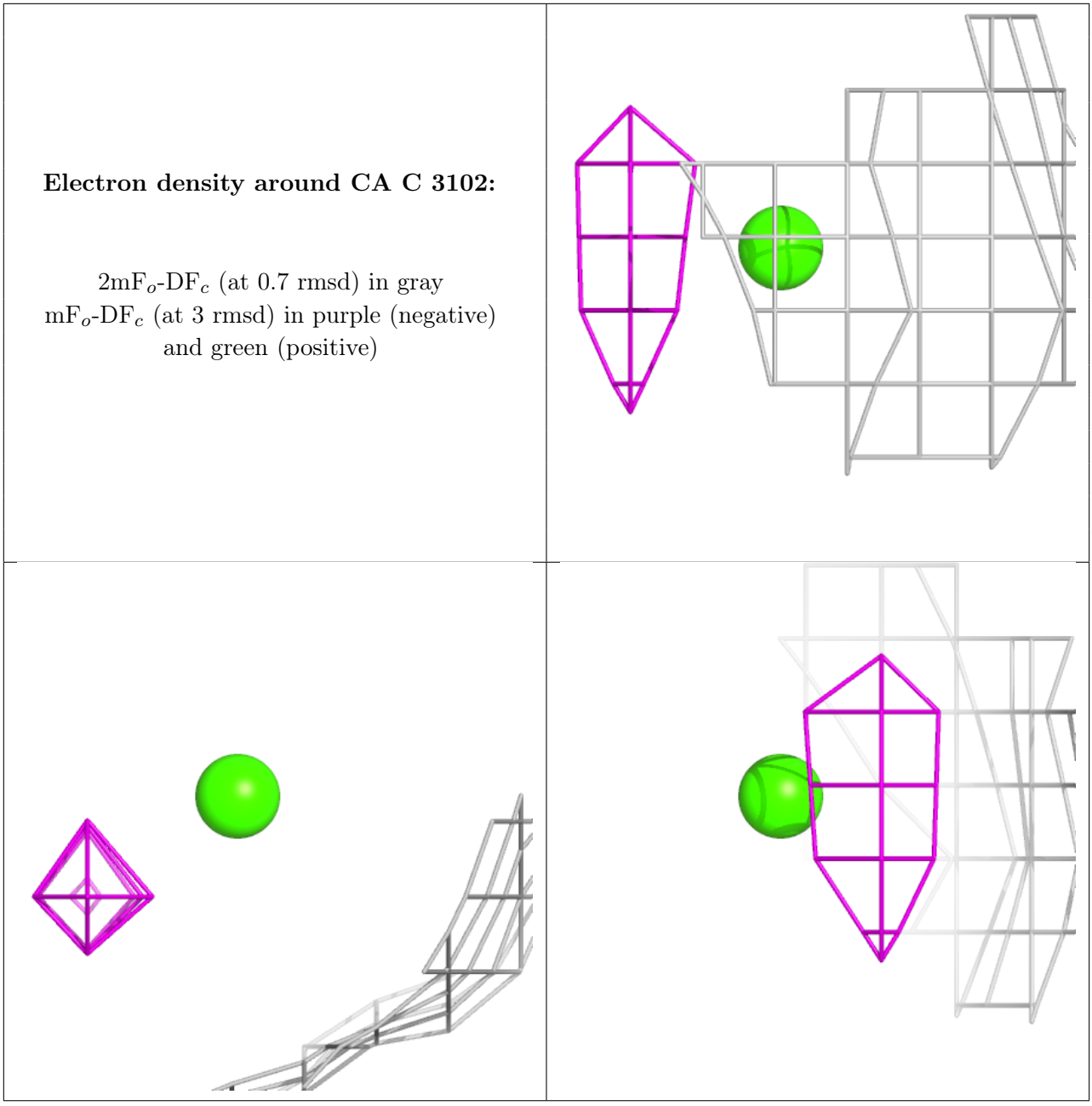
Electron density around CA A 3101:

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



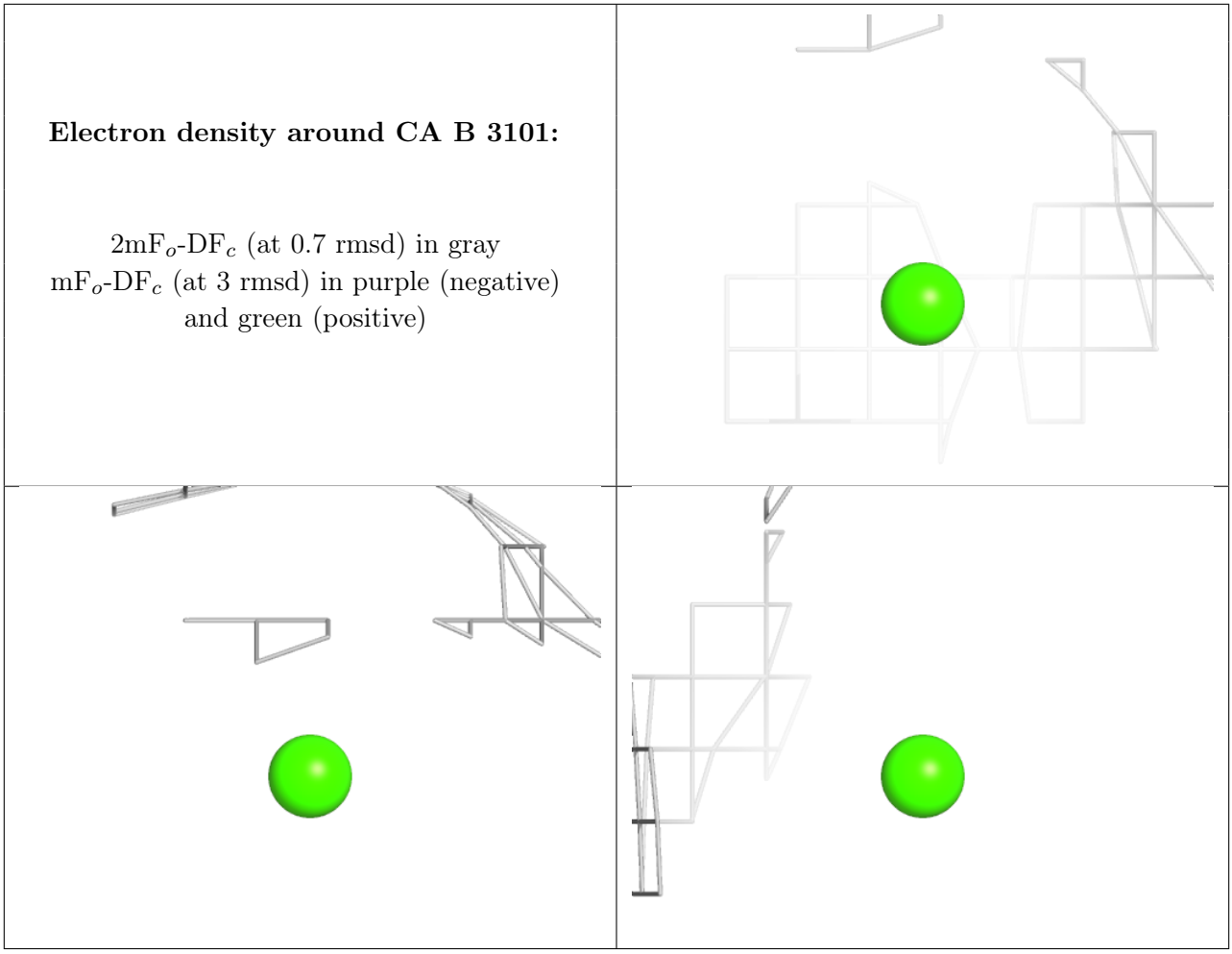
Electron density around CA C 3102:

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



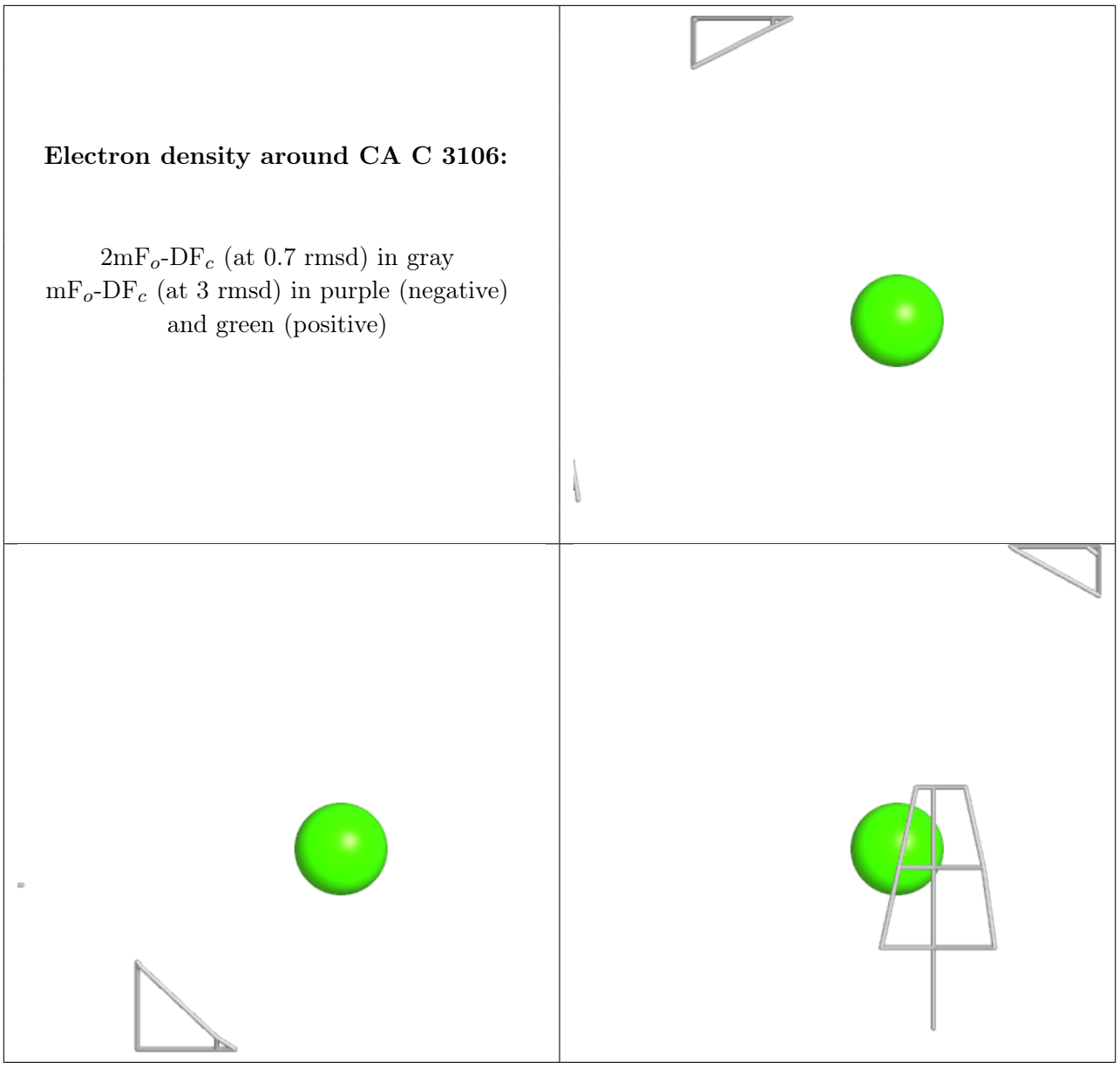
Electron density around CA B 3101:

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



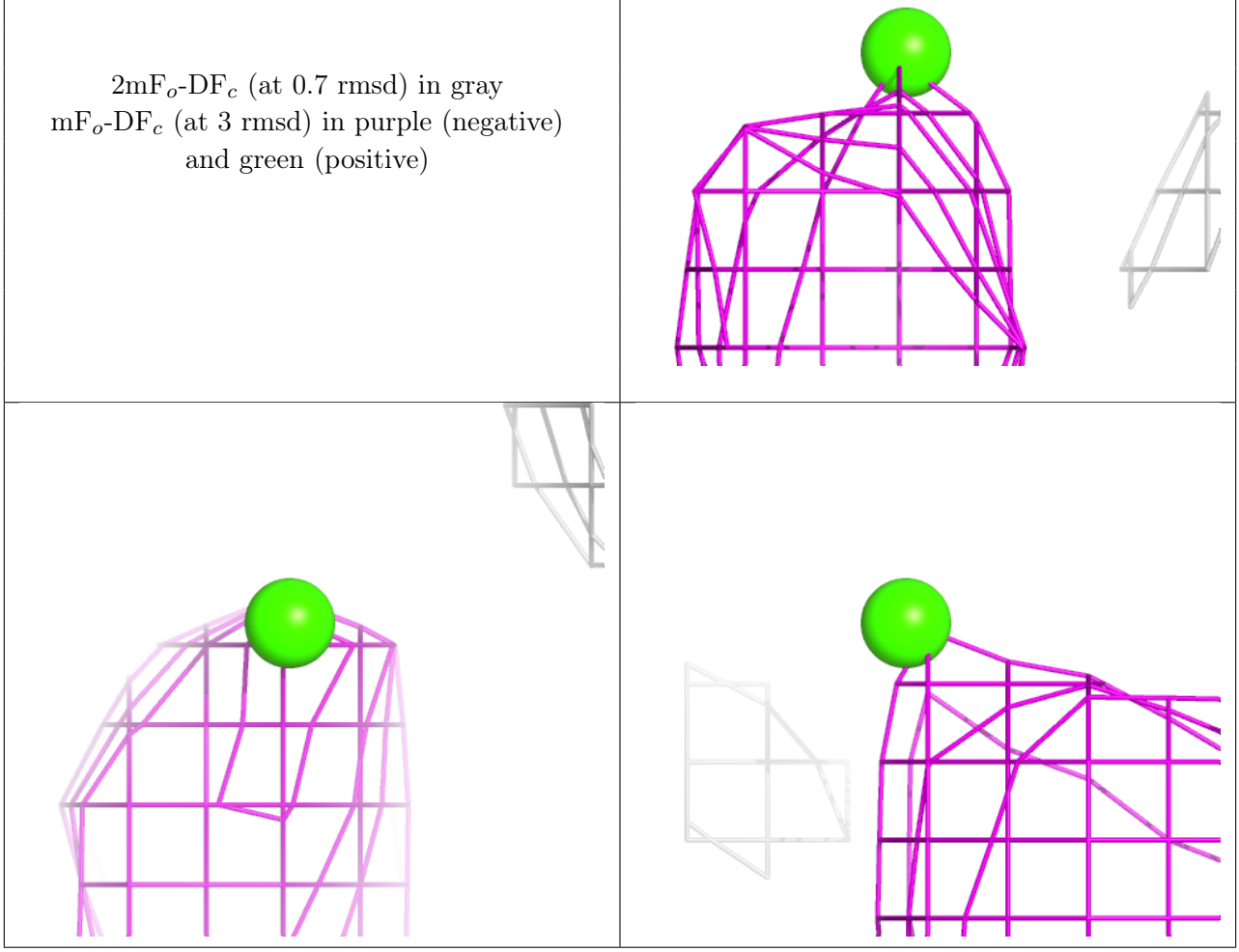
Electron density around CA C 3106:

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



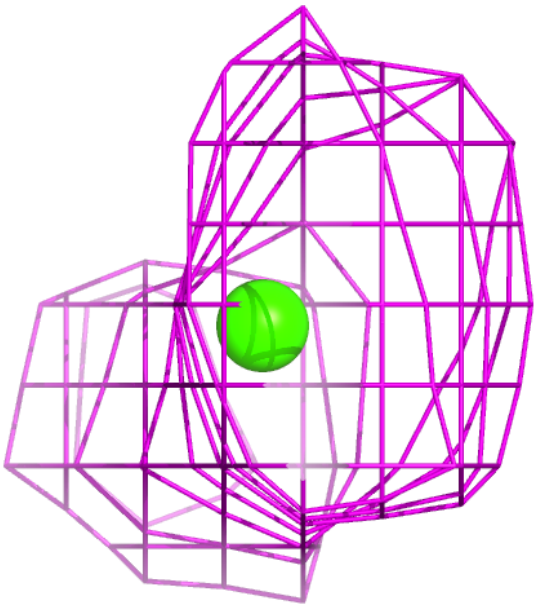
Electron density around CA C 3104:

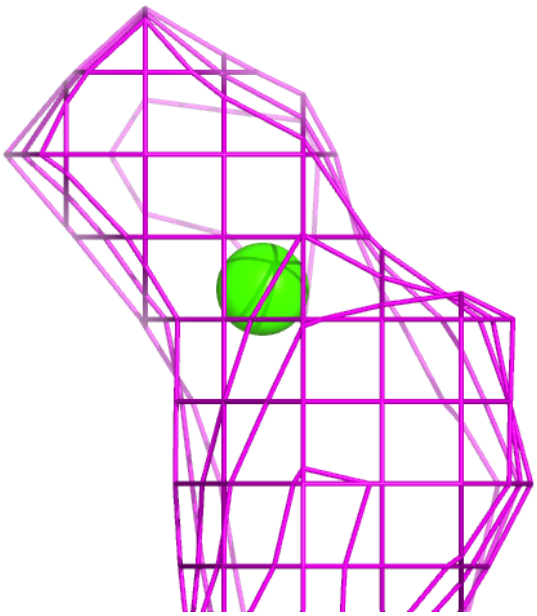
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

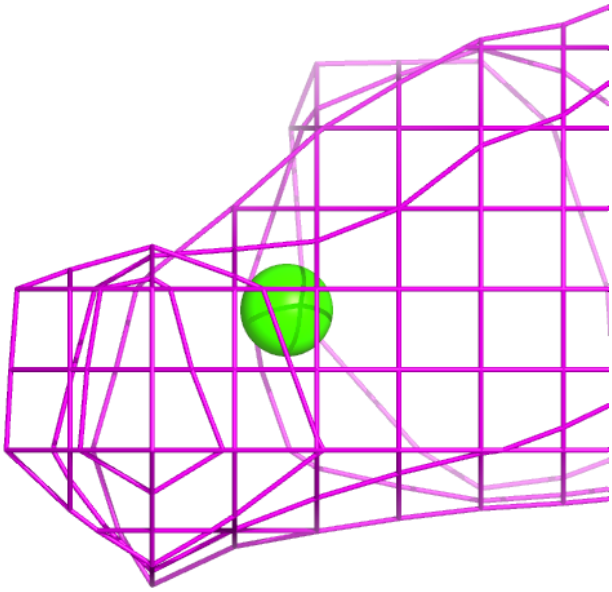


Electron density around CA C 3103:

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

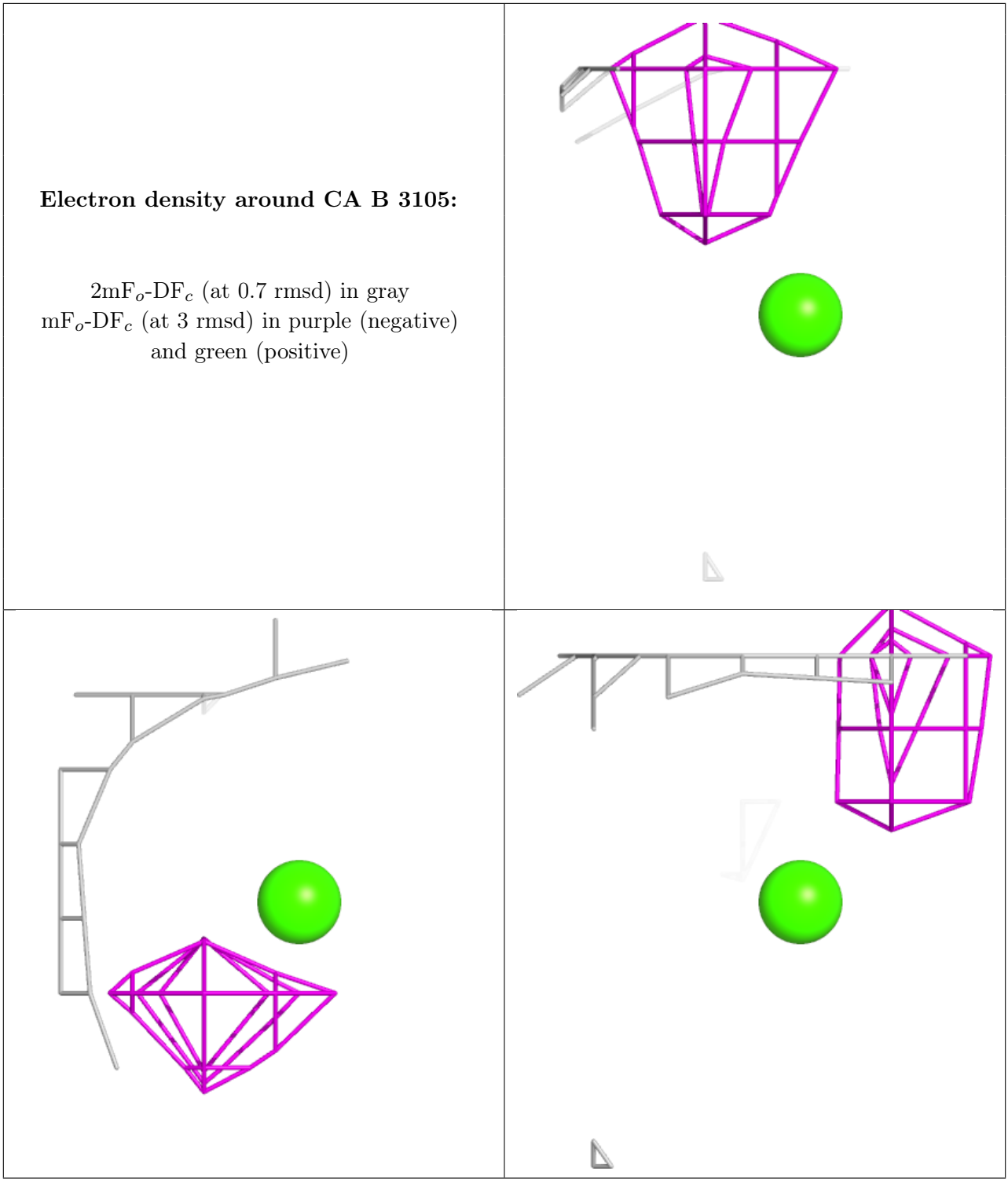






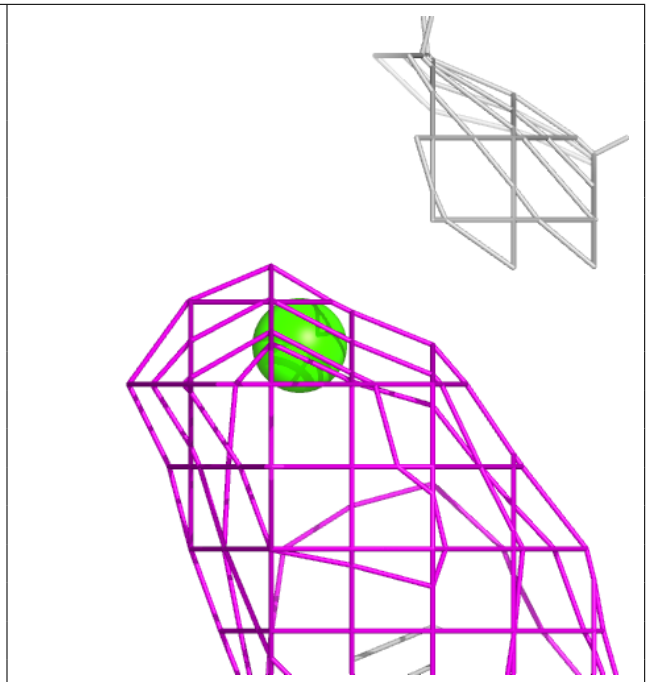
Electron density around CA B 3105:

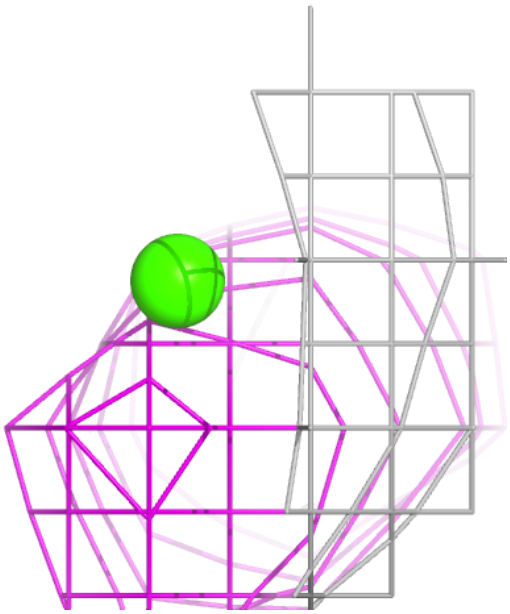
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

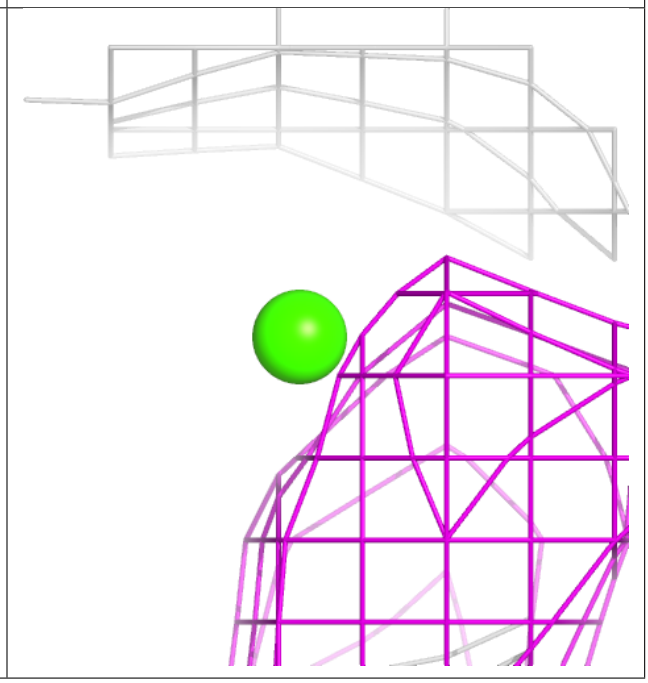


Electron density around CA B 3106:

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

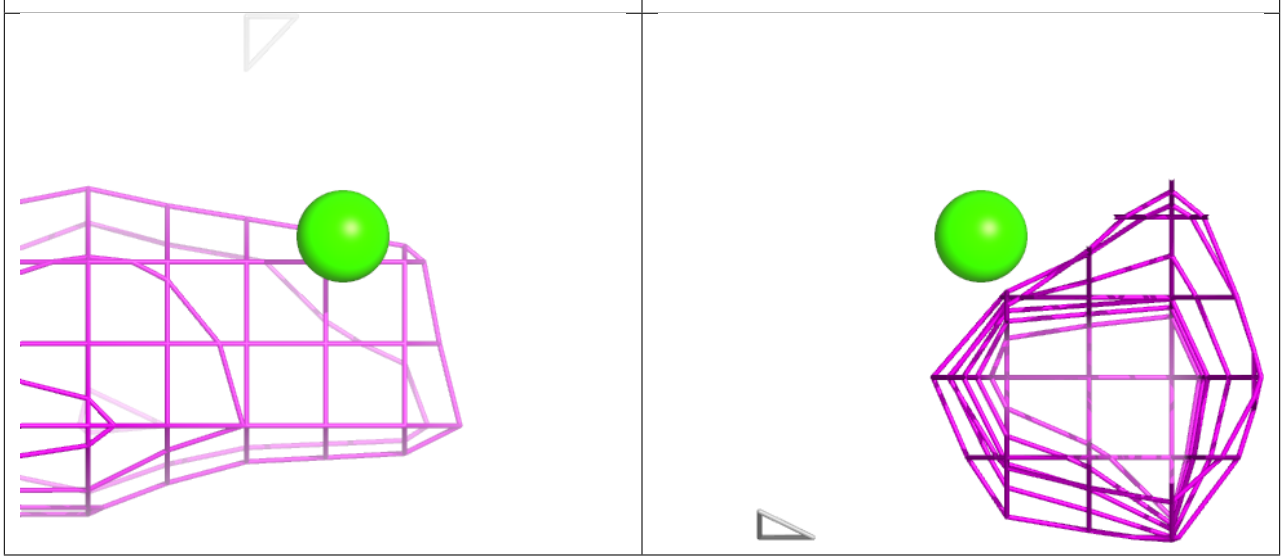






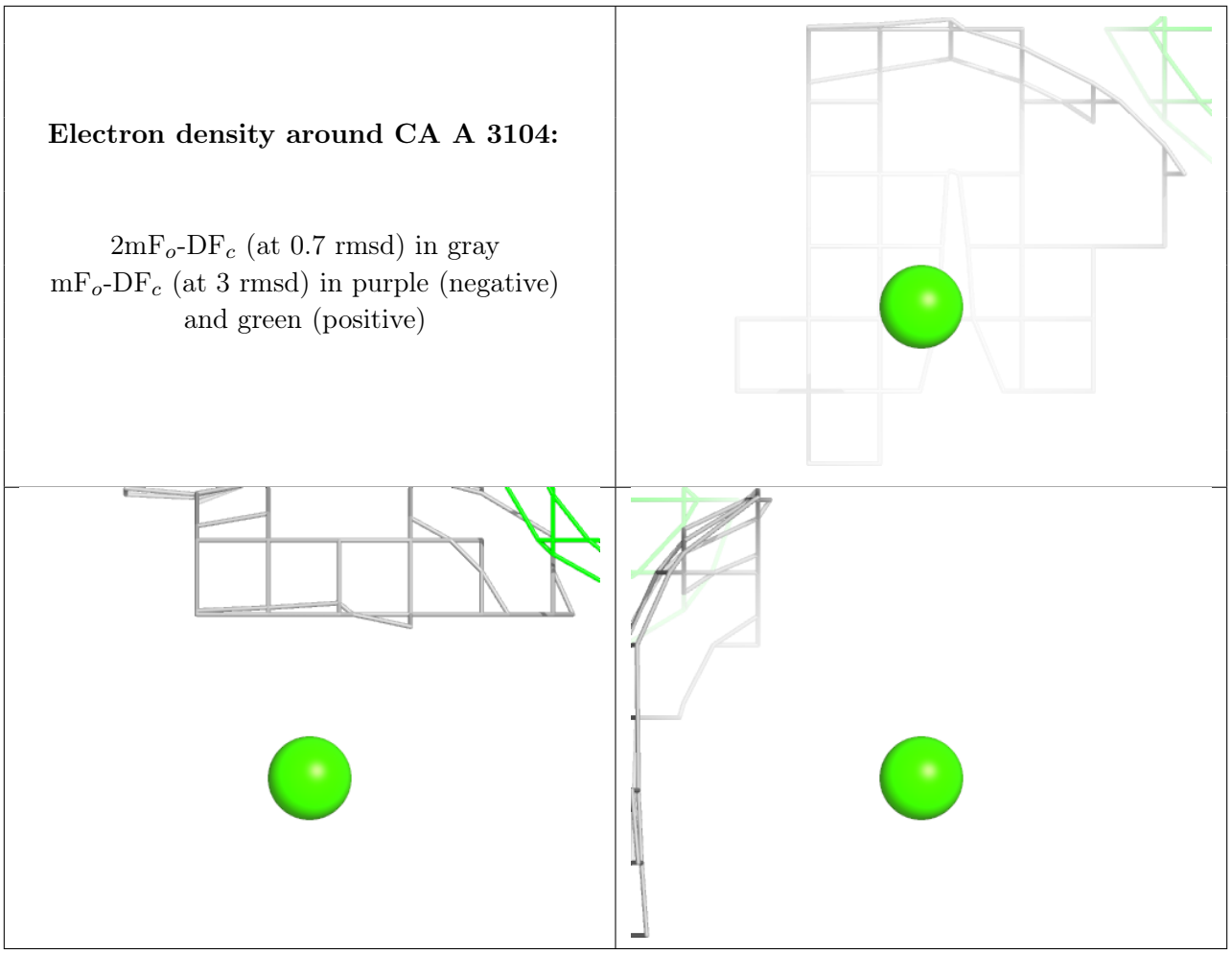
Electron density around CA A 3105:

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



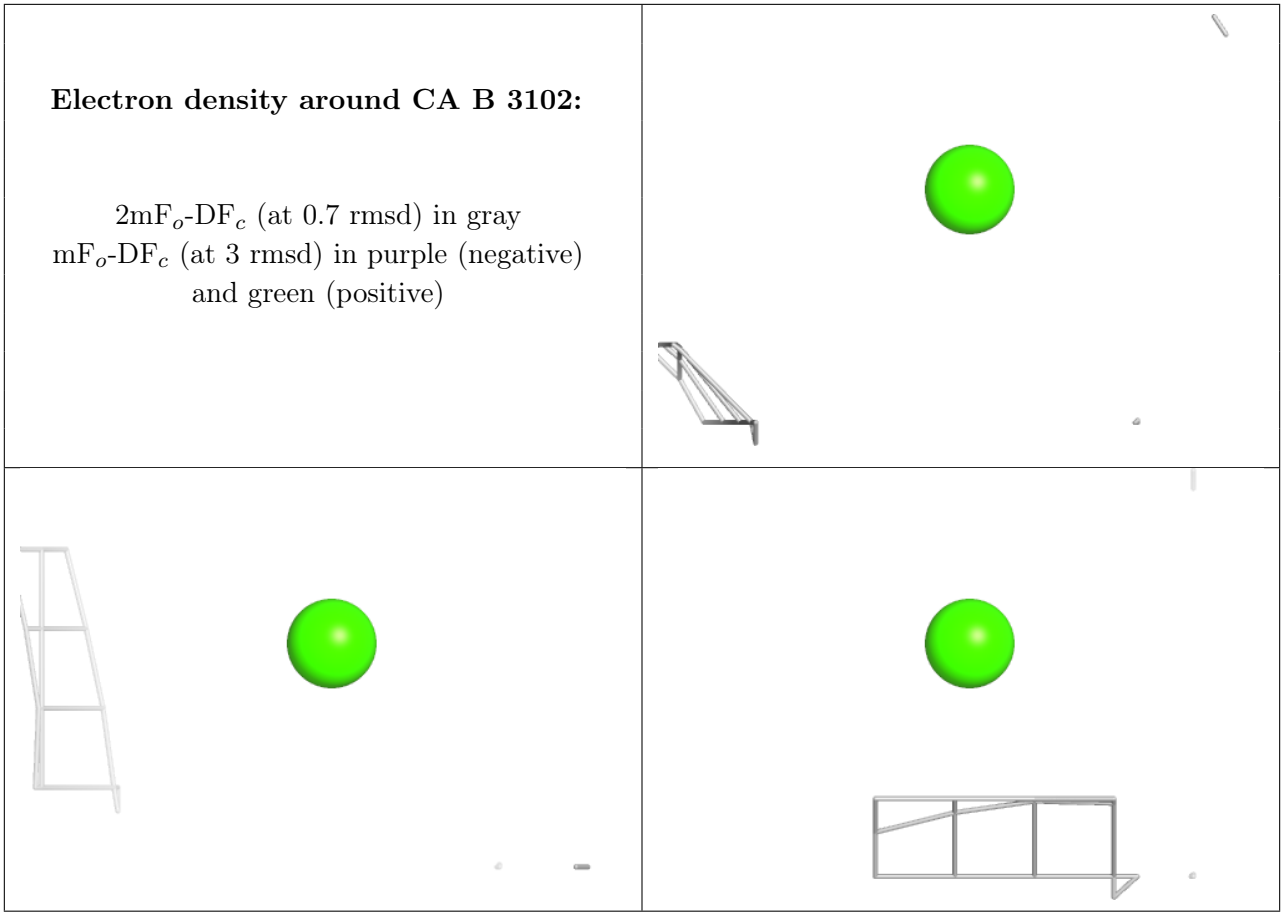
Electron density around CA A 3104:

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



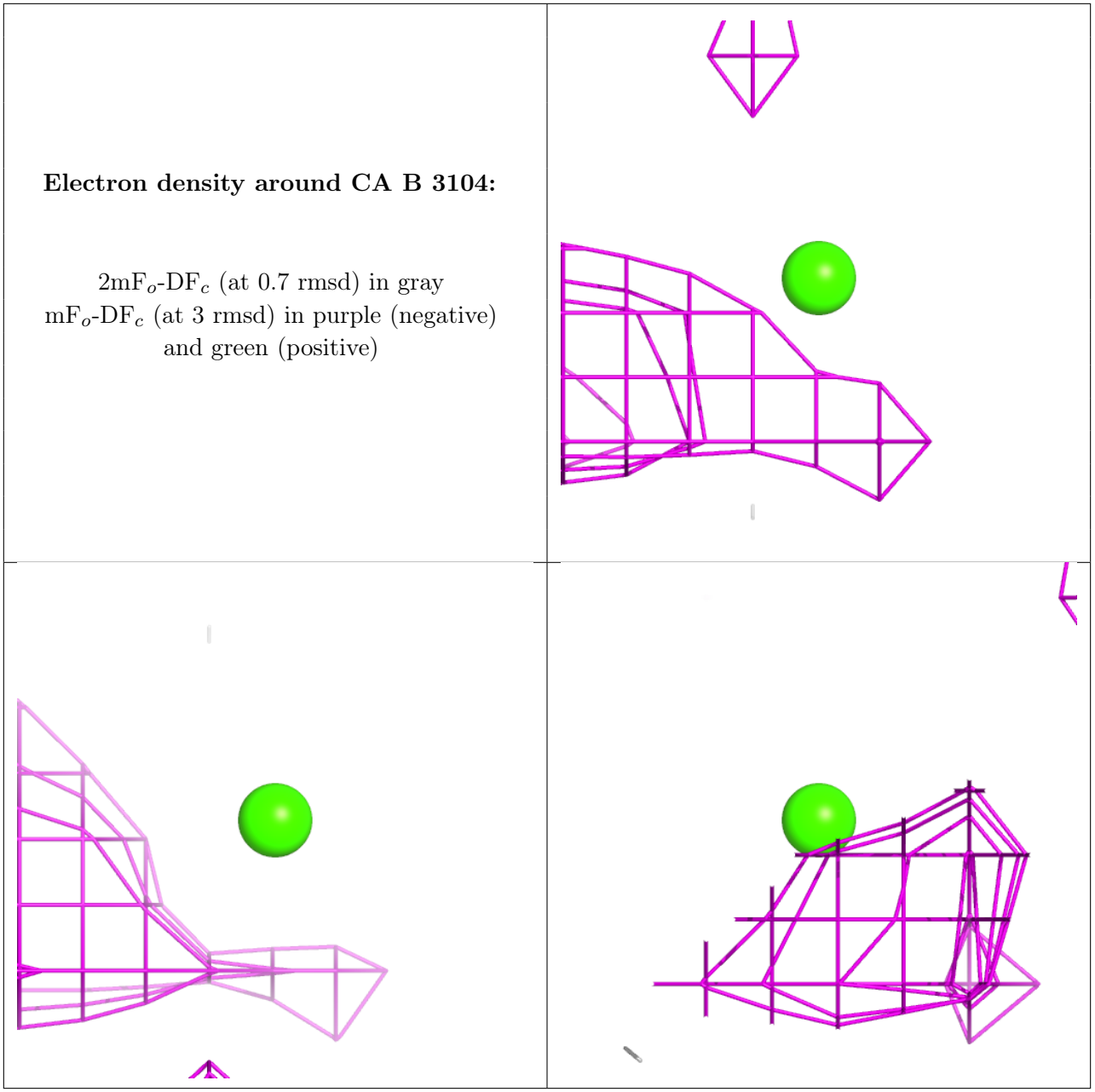
Electron density around CA B 3102:

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



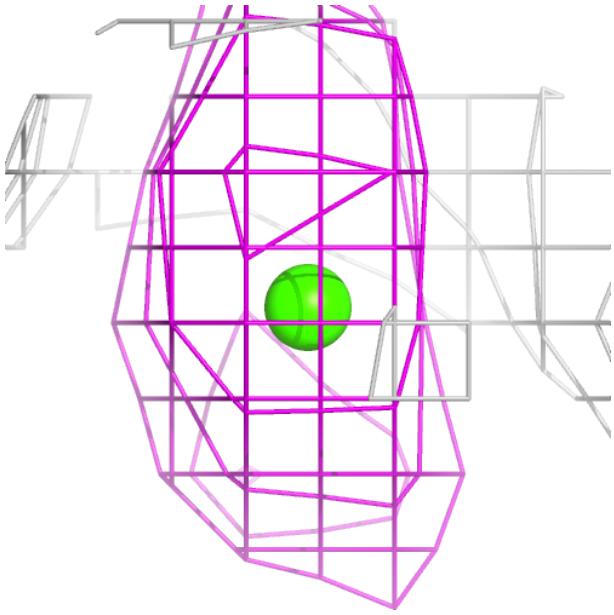
Electron density around CA B 3104:

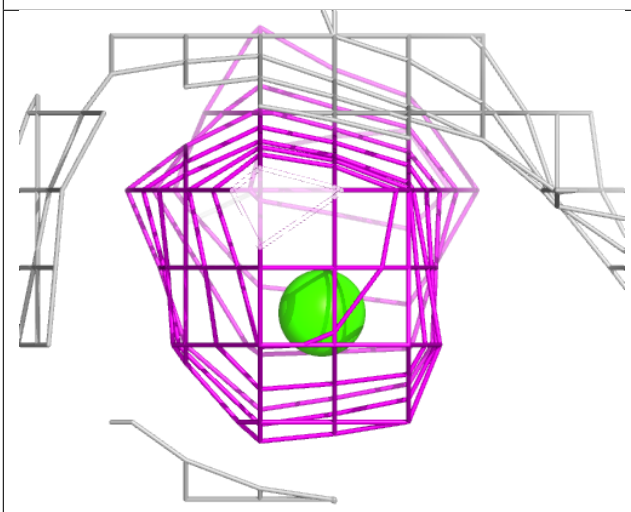
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

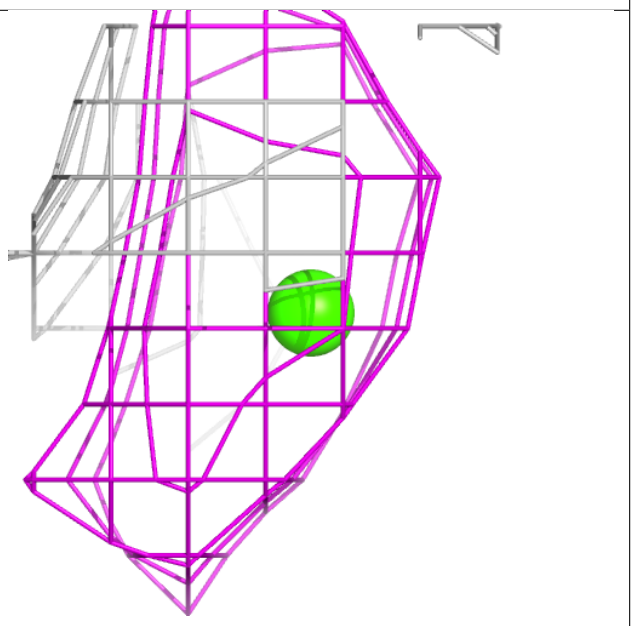


Electron density around CA C 3101:

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

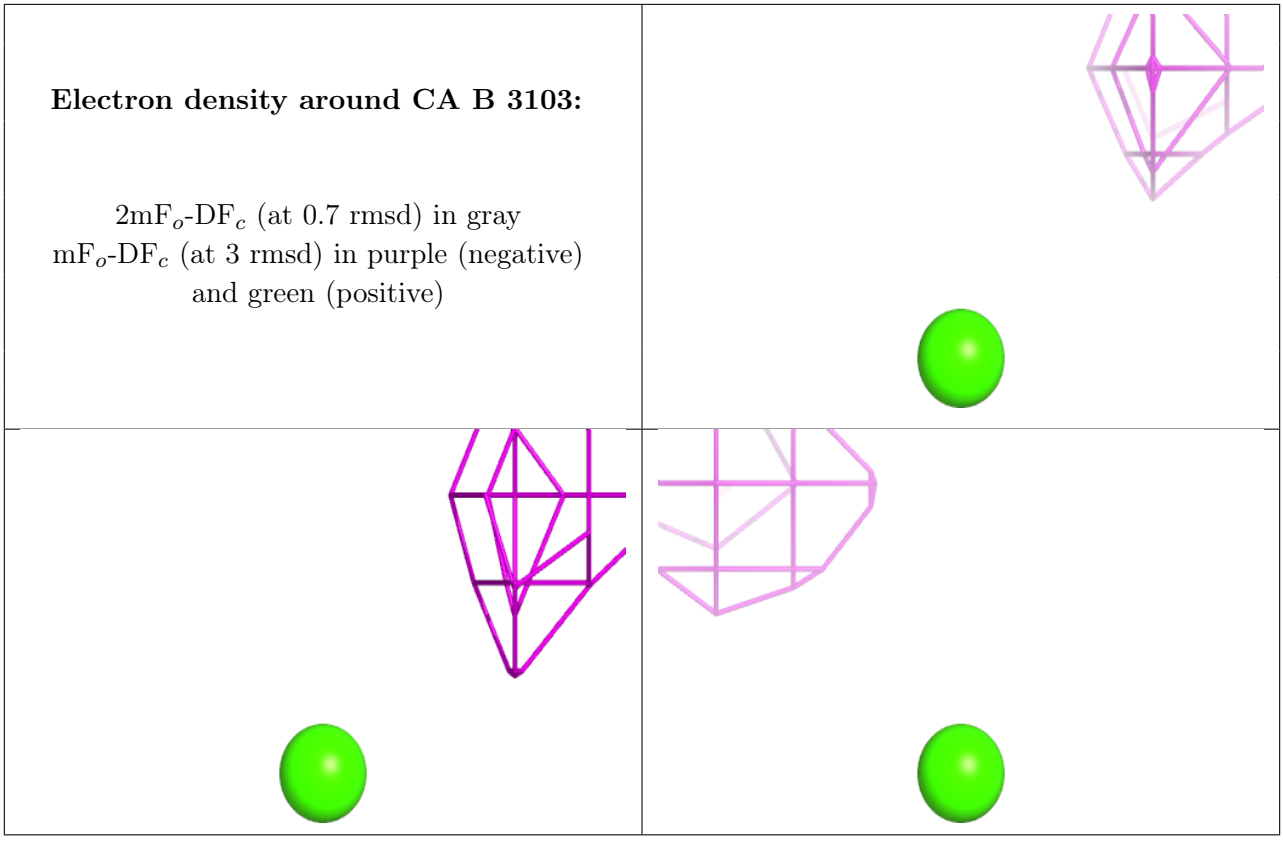






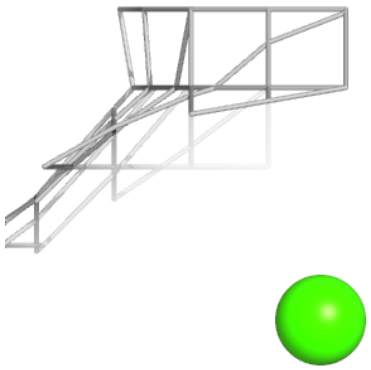
Electron density around CA B 3103:

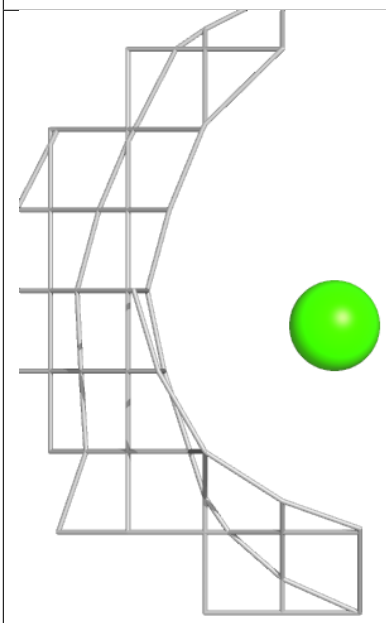
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

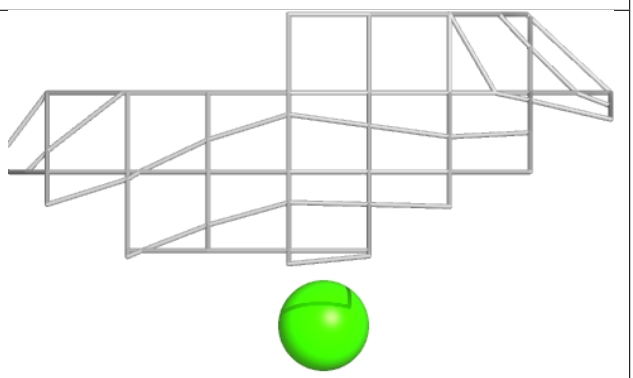


Electron density around CA A 3106:

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

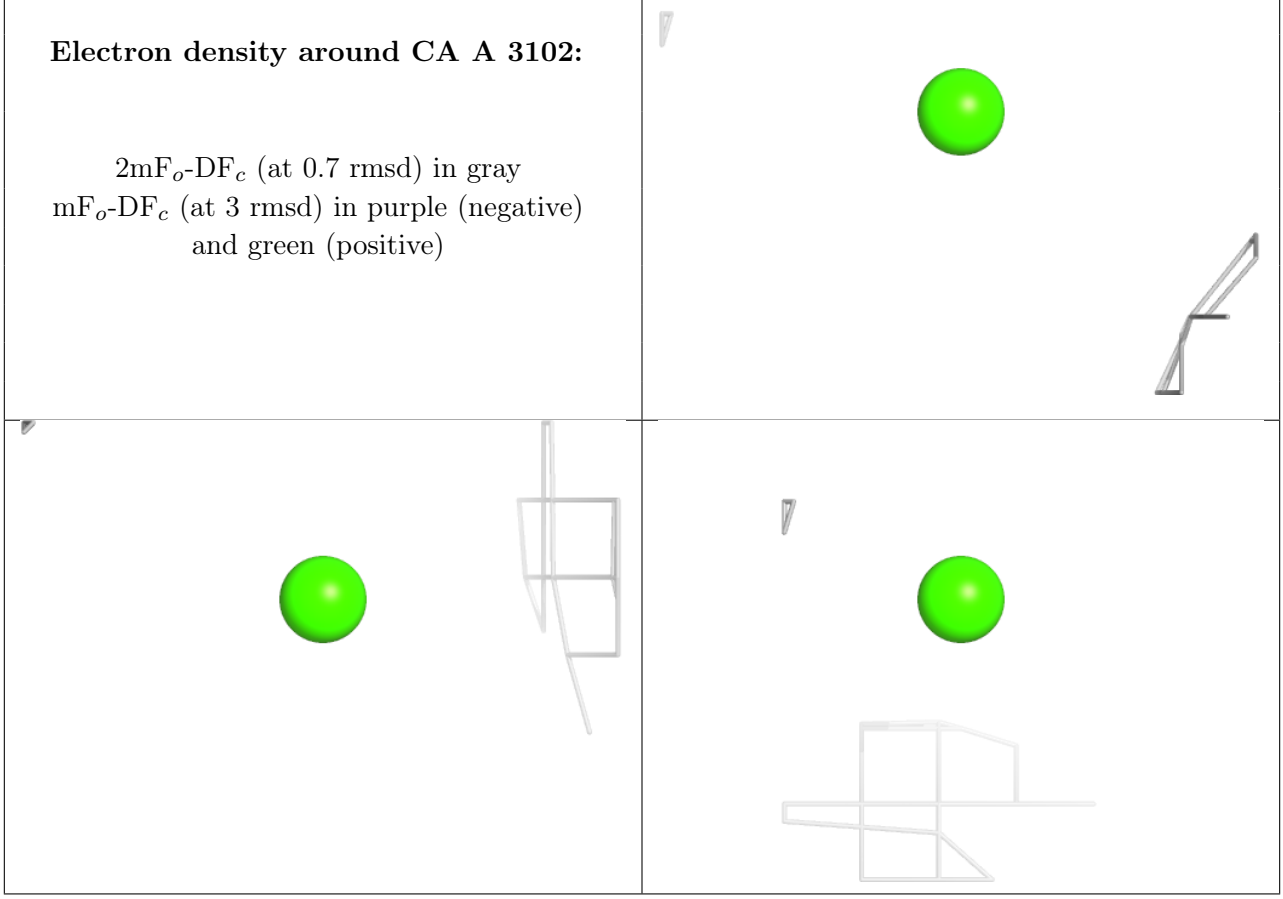






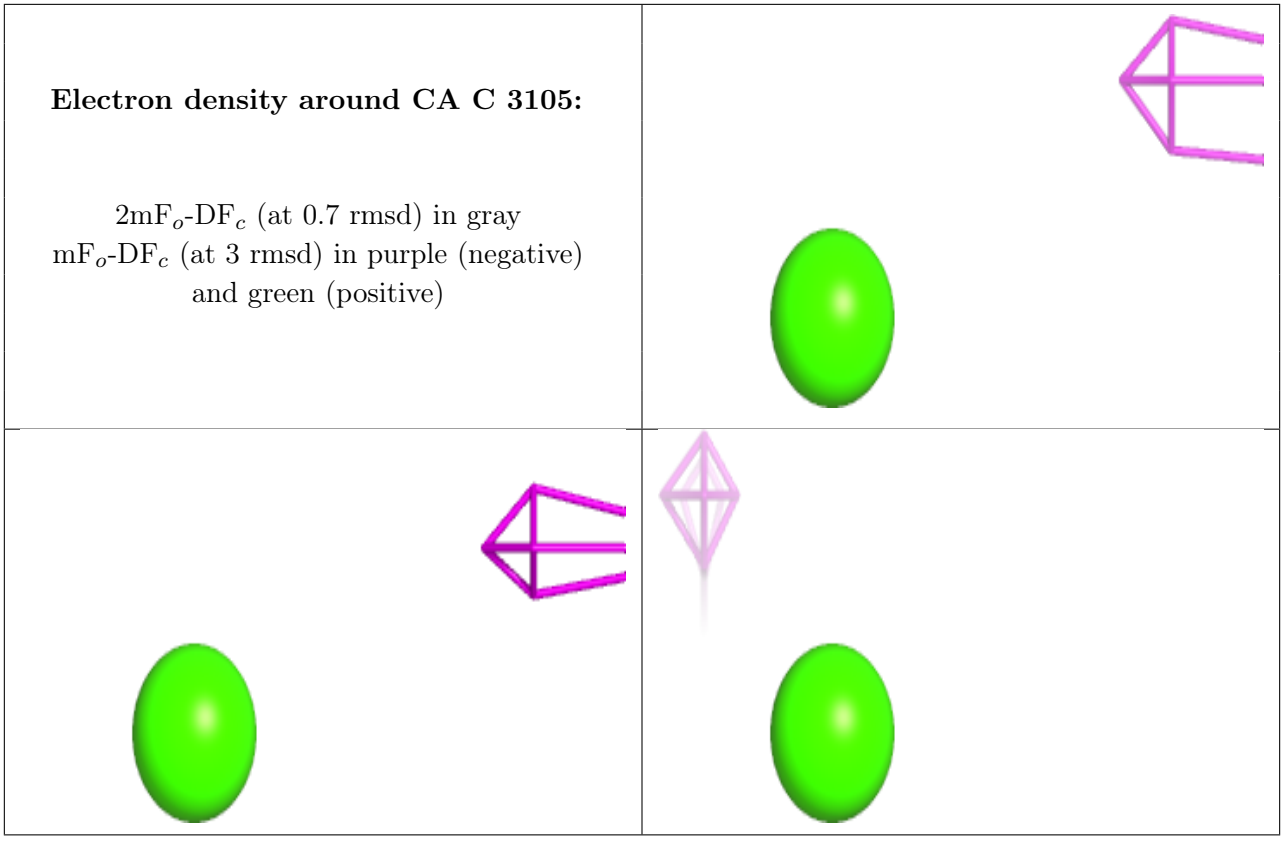
Electron density around CA A 3102:

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



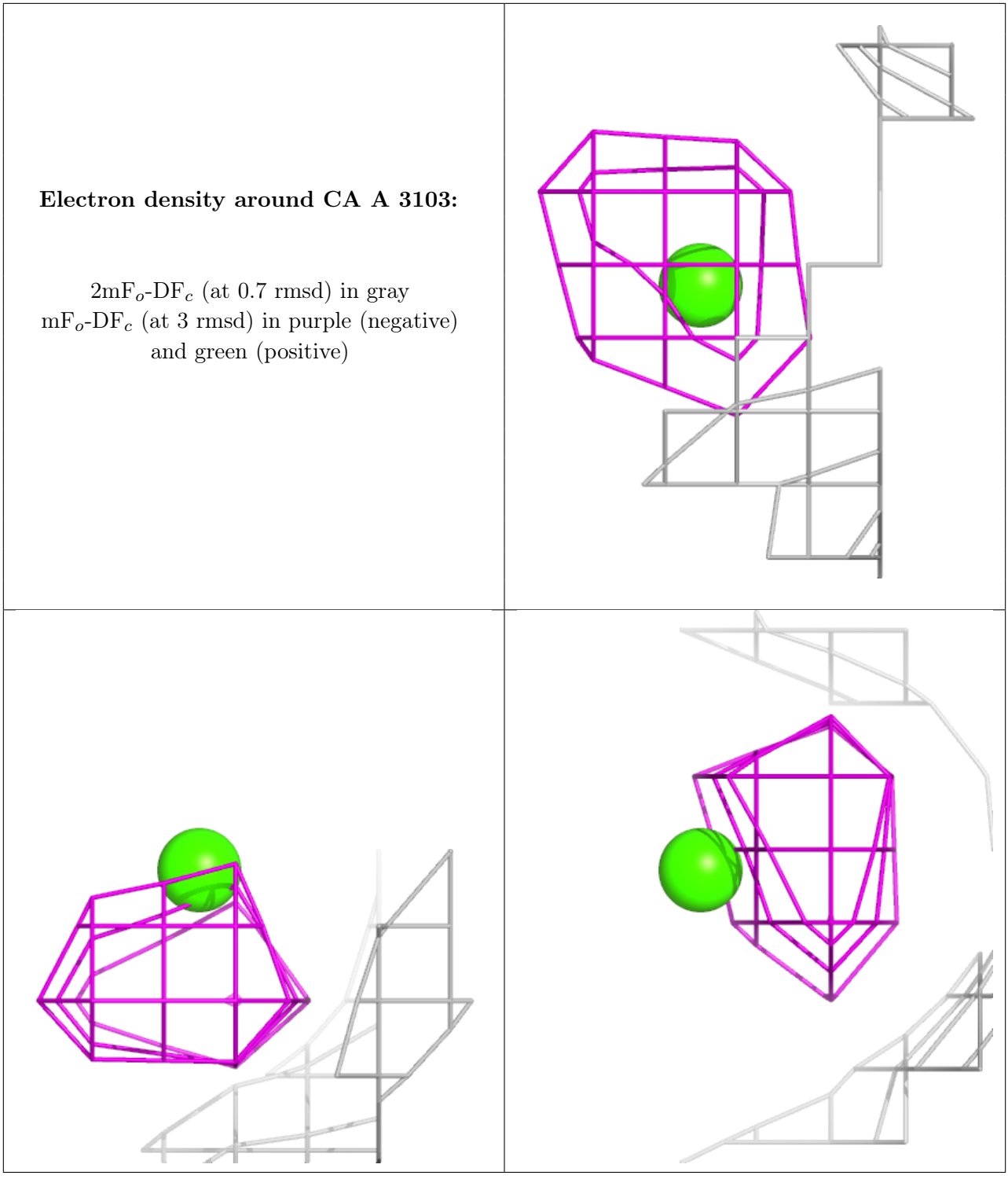
Electron density around CA C 3105:

$2mF_o$-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 3103:

$2mF_o$-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.