Full wwPDB X-ray Structure Validation Report (i)

Nov 21, 2023 – 09:24 PM JST

PDB ID	:	7WHF
Title	:	Heimdallarchaeota gelsolin (2DGel) bound to rabbit actin
Authors	:	Robinson, R.C.; Akil, C.
Deposited on	:	2021-12-30
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 (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:

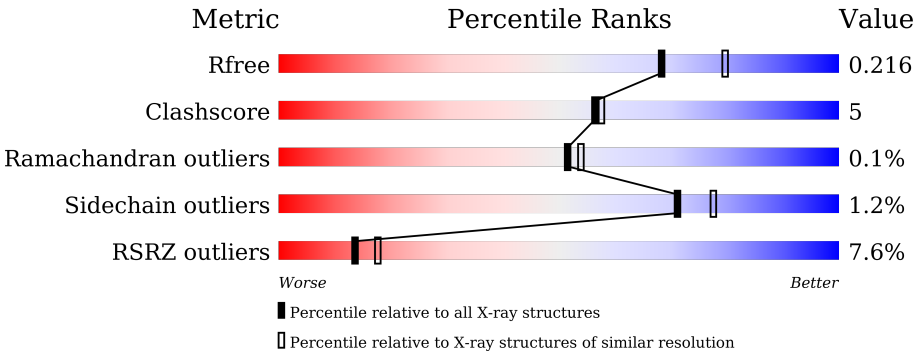
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (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	377	4% 85% 8% • 6%
1	B	377	6% 80% 12% • 7%
2	C	284	10% 70% 8% • 21%
2	G	284	7% 73% 6% • 20%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 9793 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.

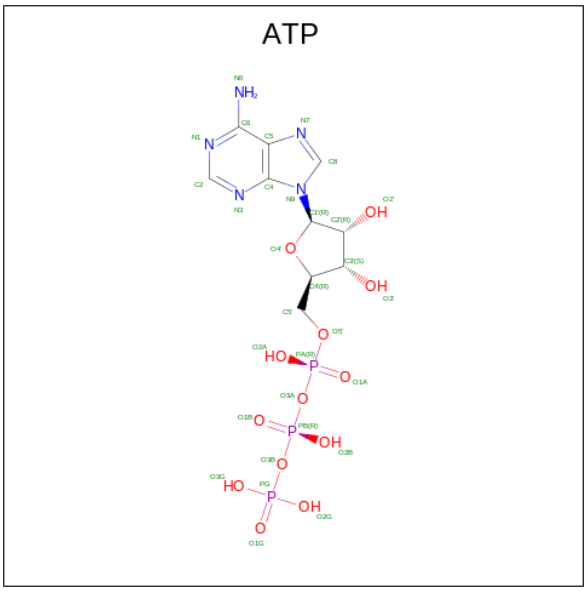
- Molecule 1 is a protein called Actin, alpha skeletal muscle.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	355	Total	C	N	O	S	13	3	0
			2793	1773	463	537	20			
1	B	350	Total	C	N	O	S	16	4	0
			2758	1754	454	529	21			

- Molecule 2 is a protein called Heimdallarchaeota Gelsolin domain-containing protein 2DGel.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	226	Total	C	N	O	S	8	1	0
			1807	1142	293	368	4			
2	C	225	Total	C	N	O	S	0	0	0
			1790	1133	289	364	4			

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).

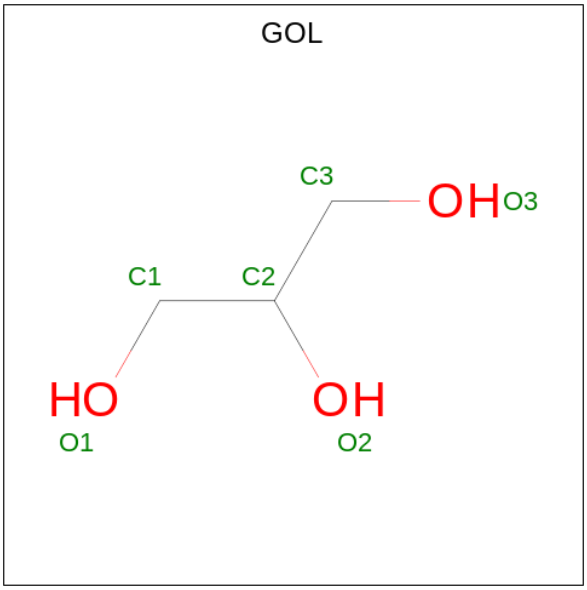


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		
4	G	8	Total	Ca	0	0
			8	8		
4	B	1	Total	Ca	0	0
			1	1		
4	C	8	Total	Ca	0	0
			8	8		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

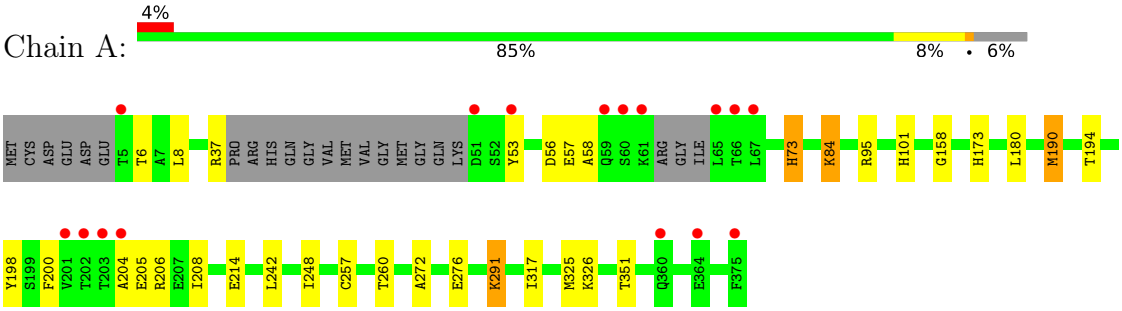
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	185	Total O 185 185	0	0
6	G	114	Total O 114 114	0	0
6	B	153	Total O 153 153	0	0
6	C	101	Total O 101 101	0	0

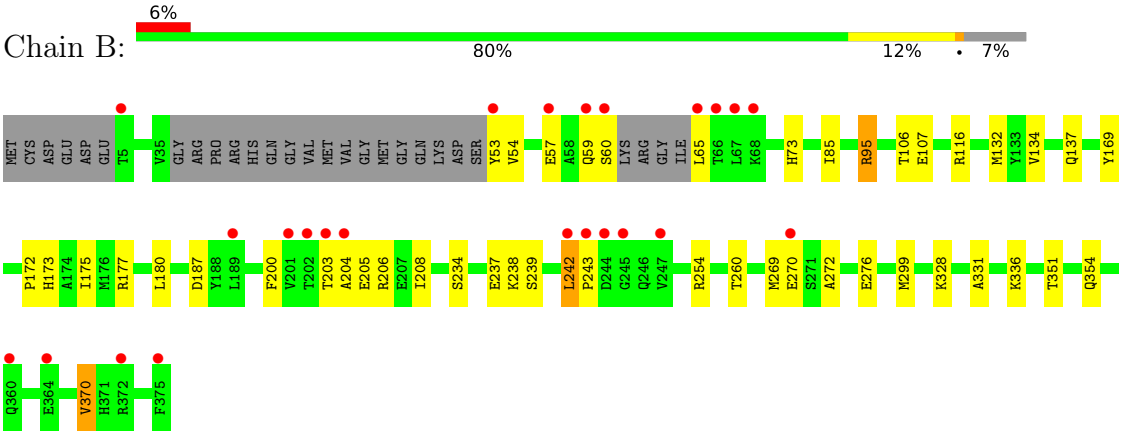
3 Residue-property plots (i)

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.

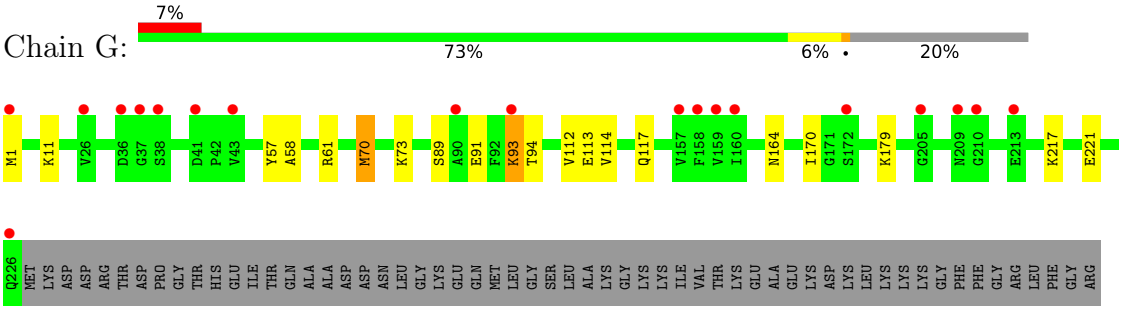
• Molecule 1: Actin, alpha skeletal muscle



• Molecule 1: Actin, alpha skeletal muscle

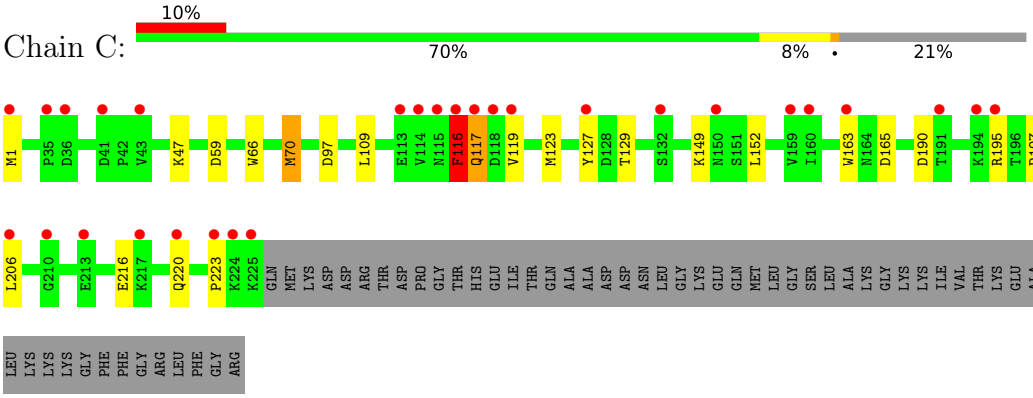


• Molecule 2: Heimdallarchaeota Gelsolin domain-containing protein 2DGel



• Molecule 2: Heimdallarchaeota Gelsolin domain-containing protein 2DGel

Chain C:


10%
70%
8%
21%
M1 P35 D36 D41 P42 V43 K47 D59 W66 M70 D97 L109 E113 V114 N115 F116 Q117 D118 V119 M123 Y127 D128 T129 S132 K149 N150 S151 L152 V159 I160 W163 N164 D165 D190 T191 K194 R195 T196 P197 M198 V199 Y200
L206 G210 E213 E216 K217 Q220 P223 K224 K225 GLN MET LYS ASP ASP ARG THR ASP PRO GLY THR HIS GLU ILE THR GLN ALA ALA ASP ASP ASN LEU GLY LYS GLU GLN MET LEU GLY SER LEU ALA LYS GLY LYS LYS ILE VAL THR LYS GLU ALA GLU LYS ASP LYS
LEU LYS LYS LYS GLY PHE PHE GLY ARG LEU PHE GLY ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.26Å 100.35Å 171.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.79 – 2.10 19.94 – 2.10	Depositor EDS
% Data completeness (in resolution range)	81.3 (19.79-2.10) 81.4 (19.94-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.56 (at 2.09Å)	Xtriage
Refinement program	PHENIX 1.17.1	Depositor
R, R_{free}	0.165 , 0.216 0.165 , 0.216	Depositor DCC
R_{free} test set	3093 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	24.1	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 61.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9793	wwPDB-VP
Average B, all atoms (Å^2)	39.0	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 34.24 % 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 7.0412e-04. 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: HIC, ATP, GOL, 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	0.67	0/2848	0.73	2/3856 (0.1%)
1	B	0.62	0/2816	0.80	6/3815 (0.2%)
2	C	0.60	0/1827	0.70	1/2470 (0.0%)
2	G	0.64	0/1844	0.76	3/2493 (0.1%)
All	All	0.63	0/9335	0.75	12/12634 (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	B	0	1
2	C	0	1
All	All	0	2

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	95	ARG	CD-NE-CZ	11.21	139.29	123.60
1	A	84	LYS	CB-CG-CD	-10.92	83.20	111.60
1	B	95	ARG	NE-CZ-NH2	-10.26	115.17	120.30
2	G	93	LYS	CD-CE-NZ	-9.97	88.77	111.70
2	G	93	LYS	CB-CG-CD	-7.92	91.02	111.60
2	G	11	LYS	CA-CB-CG	7.62	130.17	113.40
1	B	270	GLU	N-CA-CB	7.50	124.09	110.60
1	B	370	VAL	CG1-CB-CG2	-7.46	98.96	110.90
1	B	242	LEU	CB-CG-CD1	6.92	122.77	111.00
1	B	269	MET	C-N-CA	-6.91	104.42	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	6	THR	CA-CB-CG2	-5.72	104.39	112.40
2	C	206	LEU	CB-CG-CD2	-5.64	101.41	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	95	ARG	Sidechain
2	C	116	PHE	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	2793	0	2762	28	0
1	B	2758	0	2729	35	0
2	C	1790	0	1732	21	0
2	G	1807	0	1745	15	0
3	A	31	0	12	1	0
3	B	31	0	12	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	8	0	0	0	0
4	G	8	0	0	0	0
5	B	12	0	16	2	0
6	A	185	0	0	2	0
6	B	153	0	0	8	0
6	C	101	0	0	1	0
6	G	114	0	0	3	0
All	All	9793	0	9008	90	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 (90) 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:239:SER:H	2:C:198:MET:HE1	1.10	1.09
1:B:107:GLU:OE1	1:B:116:ARG:NH2	2.06	0.89
1:B:354:GLN:OE1	6:B:501:HOH:O	1.89	0.88
1:A:173:HIS:ND1	6:A:502:HOH:O	2.19	0.76
1:B:351:THR:HB	2:C:70:MET:HE2	1.70	0.74
1:B:239:SER:N	2:C:198:MET:HE1	1.95	0.73
2:G:112:VAL:HG12	2:G:114:VAL:HG13	1.73	0.71
1:B:187:ASP:OD1	1:B:206:ARG:NH2	2.24	0.71
1:B:173:HIS:ND1	6:B:506:HOH:O	2.22	0.71
1:B:336:LYS:NZ	6:B:507:HOH:O	2.24	0.71
1:B:65:LEU:N	6:B:508:HOH:O	2.25	0.69
1:B:351:THR:HB	2:C:70:MET:CE	2.23	0.69
1:B:60:SER:O	6:B:503:HOH:O	2.10	0.68
2:C:149:LYS:HA	2:C:152:LEU:HD23	1.75	0.67
1:A:37:ARG:NH2	1:A:84:LYS:HZ3	1.96	0.64
2:G:170:ILE:HD12	2:G:179:LYS:HG3	1.80	0.64
1:B:239:SER:H	2:C:198:MET:CE	2.00	0.64
1:B:204:ALA:O	1:B:208:ILE:HG12	2.00	0.62
1:B:65:LEU:O	6:B:504:HOH:O	2.16	0.62
1:B:234:SER:O	1:B:237:GLU:HG3	1.99	0.62
2:C:119:VAL:CG2	2:C:223:PRO:HG3	2.30	0.61
1:A:351:THR:HB	2:G:70:MET:CE	2.29	0.61
1:A:37:ARG:CZ	1:A:84:LYS:HE2	2.30	0.61
1:B:254:ARG:NH2	5:B:403:GOL:H32	2.16	0.60
1:A:291:LYS:HD2	1:A:325:MET:SD	2.42	0.60
1:B:169:TYR:OH	6:B:505:HOH:O	2.17	0.58
2:G:117:GLN:N	6:G:404:HOH:O	2.37	0.58
1:B:254:ARG:HH22	5:B:403:GOL:H32	1.67	0.57
1:A:351:THR:HB	2:G:70:MET:HE3	1.86	0.57
2:C:119:VAL:HG23	2:C:223:PRO:HG3	1.87	0.56
2:C:198:MET:HG2	2:C:200:TYR:CZ	2.42	0.55
1:B:328:LYS:NZ	6:B:510:HOH:O	2.35	0.54
1:A:190:MET:HE1	1:A:206:ARG:CD	2.37	0.54
1:A:95:ARG:HG2	1:A:95:ARG:HH11	1.73	0.53
1:B:208:ILE:HG21	1:B:242:LEU:HD23	1.90	0.52
1:B:116:ARG:HG2	1:B:370:VAL:HG21	1.91	0.51
2:C:117:GLN:O	2:C:119:VAL:HG13	2.11	0.51
2:G:179:LYS:NZ	6:G:405:HOH:O	2.42	0.51
1:B:272:ALA:HB1	1:B:276:GLU:HB2	1.93	0.51
1:B:299:MET:HE2	1:B:331:ALA:HB2	1.92	0.51
2:G:89:SER:O	2:G:93:LYS:HG3	2.11	0.51
1:A:198:TYR:CE1	1:A:248:ILE:HD12	2.46	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:234:SER:CA	1:B:237:GLU:HG3	2.42	0.49
2:C:116:PHE:CZ	2:C:223:PRO:HB3	2.48	0.49
1:A:194:THR:HA	1:A:198:TYR:O	2.12	0.49
2:C:127:TYR:CE1	2:C:129:THR:HG22	2.48	0.49
2:C:123:MET:HE1	2:C:152:LEU:HD22	1.93	0.49
1:A:351:THR:HB	2:G:70:MET:HE2	1.95	0.48
2:C:216:GLU:O	2:C:220:GLN:HG3	2.13	0.48
1:A:214:GLU:HG2	3:A:401:ATP:C5	2.49	0.47
1:A:53:TYR:O	1:A:58:ALA:HB2	2.15	0.46
2:G:113:GLU:H	2:G:113:GLU:CD	2.19	0.46
1:B:172:PRO:HA	1:B:175:ILE:HD12	1.98	0.46
2:C:119:VAL:HG22	2:C:223:PRO:HG3	1.96	0.46
1:A:200:PHE:HA	1:A:205:GLU:HB3	1.97	0.46
1:B:234:SER:C	1:B:237:GLU:HG3	2.36	0.46
1:B:242:LEU:HB3	1:B:243:PRO:CD	2.45	0.46
2:C:59:ASP:HB3	2:C:109:LEU:HG	1.96	0.45
2:G:217:LYS:O	2:G:221:GLU:HG3	2.17	0.45
2:C:1:MET:N	6:C:408:HOH:O	2.50	0.45
2:G:73:LYS:HB2	2:G:73:LYS:HE2	1.73	0.44
1:B:106:THR:HB	1:B:137:GLN:HG3	1.99	0.44
1:A:180:LEU:HD11	1:A:260:THR:HG22	1.98	0.44
1:A:37:ARG:CZ	1:A:84:LYS:HZ3	2.29	0.44
1:A:242:LEU:HD11	1:A:248:ILE:HD11	1.99	0.44
1:B:132:MET:HE2	1:B:134:VAL:HG23	1.99	0.44
1:A:73:HIC:HD2	1:A:158:GLY:O	2.18	0.43
1:A:291:LYS:HE2	1:A:326:LYS:H	1.83	0.43
1:B:200:PHE:HA	1:B:205:GLU:HB3	2.00	0.43
1:B:54:VAL:HG13	1:B:85:ILE:HD13	2.01	0.43
1:A:204:ALA:O	1:A:208:ILE:HG13	2.18	0.43
1:A:190:MET:HE1	1:A:206:ARG:HD3	2.01	0.43
1:B:238:LYS:NZ	2:C:165:ASP:OD1	2.52	0.42
1:A:190:MET:HE1	1:A:206:ARG:HD2	2.01	0.42
1:A:8:LEU:HG	1:A:101:HIS:HB3	2.00	0.42
2:G:1:MET:HE2	6:G:476:HOH:O	2.19	0.42
1:B:53:TYR:HD1	1:B:57:GLU:HG2	1.84	0.42
2:C:66:TRP:HE1	2:C:70:MET:HE2	1.85	0.42
1:A:53:TYR:HD2	1:A:57:GLU:CG	2.33	0.42
2:G:91:GLU:O	2:G:94:THR:HB	2.20	0.42
1:A:272:ALA:HB1	1:A:276:GLU:HB2	2.03	0.41
1:A:56:ASP:OD2	6:A:503:HOH:O	2.22	0.41
2:C:195:ARG:O	2:C:195:ARG:HG2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:59:GLN:O	1:B:203:THR:HG21	2.21	0.41
2:G:57:TYR:CG	2:G:58:ALA:N	2.90	0.40
2:C:190:ASP:OD1	2:C:197:PRO:HD2	2.20	0.40
1:A:95:ARG:HG2	1:A:95:ARG:NH1	2.37	0.40
2:G:57:TYR:O	2:G:61:ARG:HG3	2.21	0.40
1:B:180:LEU:HD11	1:B:260:THR:HG22	2.04	0.40

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	351/377 (93%)	344 (98%)	7 (2%)	0	100	100
1	B	347/377 (92%)	340 (98%)	7 (2%)	0	100	100
2	C	223/284 (78%)	216 (97%)	6 (3%)	1 (0%)	34	32
2	G	225/284 (79%)	220 (98%)	5 (2%)	0	100	100
All	All	1146/1322 (87%)	1120 (98%)	25 (2%)	1 (0%)	51	54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	163	TRP

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	304/319 (95%)	301 (99%)	3 (1%)	76 82
1	B	301/319 (94%)	300 (100%)	1 (0%)	92 95
2	C	199/247 (81%)	193 (97%)	6 (3%)	41 44
2	G	201/247 (81%)	198 (98%)	3 (2%)	65 71
All	All	1005/1132 (89%)	992 (99%)	13 (1%)	71 75

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	190	MET
1	A	257	CYS
1	A	291	LYS
2	G	70	MET
2	G	164[A]	ASN
2	G	164[B]	ASN
1	B	177	ARG
2	C	47	LYS
2	C	70	MET
2	C	97	ASP
2	C	116	PHE
2	C	117	GLN
2	C	198	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	G	117	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)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	HIC	A	73	1	8,11,12	1.93	2 (25%)	6,14,16	1.35	2 (33%)
1	HIC	B	73	1	8,11,12	1.94	2 (25%)	6,14,16	0.98	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	HIC	A	73	1	-	0/5/6/8	0/1/1/1
1	HIC	B	73	1	-	0/5/6/8	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	73	HIC	CD2-CG	4.38	1.42	1.36
1	A	73	HIC	CD2-CG	4.15	1.42	1.36
1	B	73	HIC	CD2-NE2	2.26	1.41	1.38
1	A	73	HIC	CD2-NE2	2.04	1.41	1.38

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	73	HIC	CB-CA-C	-2.28	107.20	111.47
1	A	73	HIC	CG-CD2-NE2	-2.23	105.36	107.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	73	HIC	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 22 ligands modelled in this entry, 18 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	ATP	B	401	4	26,33,33	0.78	1 (3%)	31,52,52	1.17	4 (12%)
5	GOL	B	404	-	5,5,5	0.90	0	5,5,5	1.15	0
5	GOL	B	403	-	5,5,5	1.63	1 (20%)	5,5,5	0.81	0
3	ATP	A	401	4	26,33,33	0.87	1 (3%)	31,52,52	1.38	3 (9%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ATP	B	401	4	-	4/18/38/38	0/3/3/3
5	GOL	B	404	-	-	4/4/4/4	-
5	GOL	B	403	-	-	2/4/4/4	-
3	ATP	A	401	4	-	2/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	403	GOL	C3-C2	2.33	1.61	1.51
3	B	401	ATP	C5-C4	2.15	1.46	1.40
3	A	401	ATP	C5-C4	2.07	1.46	1.40

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	401	ATP	C2-N1-C6	3.71	125.10	118.75
3	A	401	ATP	N3-C2-N1	-3.69	122.91	128.68
3	B	401	ATP	N3-C2-N1	-3.04	123.93	128.68
3	A	401	ATP	O4'-C1'-C2'	-2.25	103.64	106.93
3	B	401	ATP	O4'-C1'-C2'	-2.07	103.90	106.93
3	B	401	ATP	C2-N1-C6	2.03	122.23	118.75
3	B	401	ATP	O2B-PB-O1B	2.01	122.18	112.24

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	404	GOL	O1-C1-C2-C3
5	B	404	GOL	C1-C2-C3-O3
5	B	404	GOL	O1-C1-C2-O2
5	B	404	GOL	O2-C2-C3-O3
3	B	401	ATP	C5'-O5'-PA-O3A
5	B	403	GOL	O2-C2-C3-O3
3	A	401	ATP	PG-O3B-PB-O2B
3	B	401	ATP	PG-O3B-PB-O1B
3	B	401	ATP	PB-O3B-PG-O2G
5	B	403	GOL	O1-C1-C2-O2
3	A	401	ATP	PG-O3B-PB-O1B
3	B	401	ATP	PG-O3B-PB-O2B

There are no ring outliers.

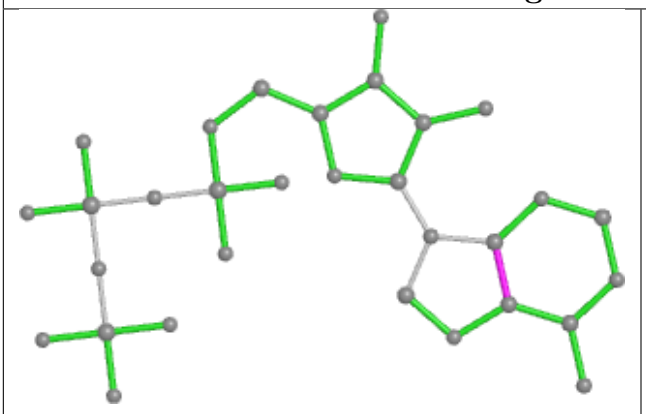
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	403	GOL	2	0
3	A	401	ATP	1	0

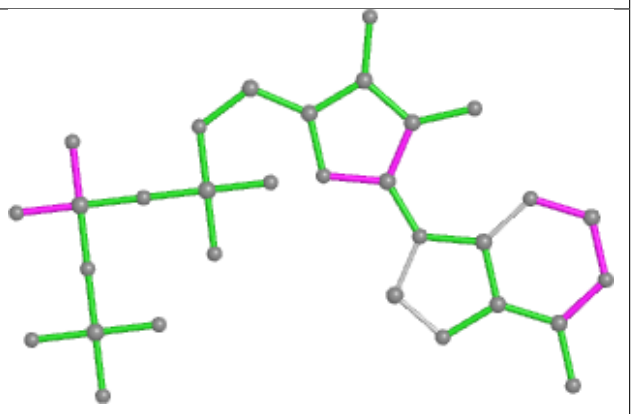
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient

equivalents in the CSD to analyse the geometry.

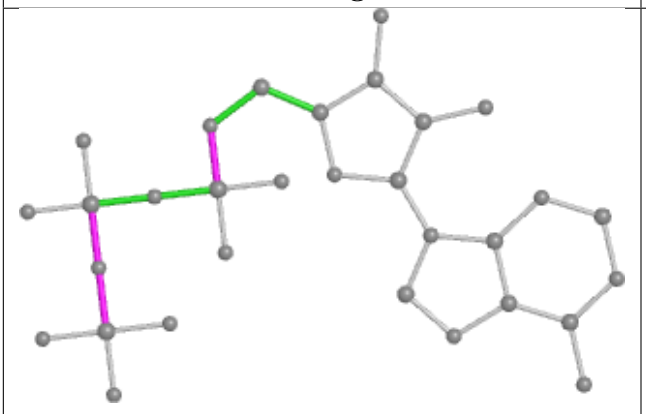
Ligand ATP B 401



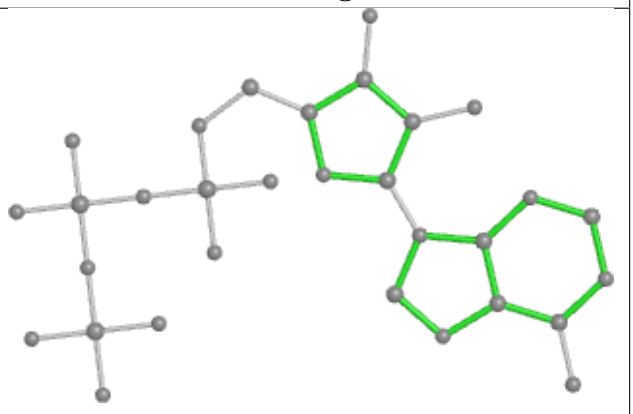
Bond lengths



Bond angles

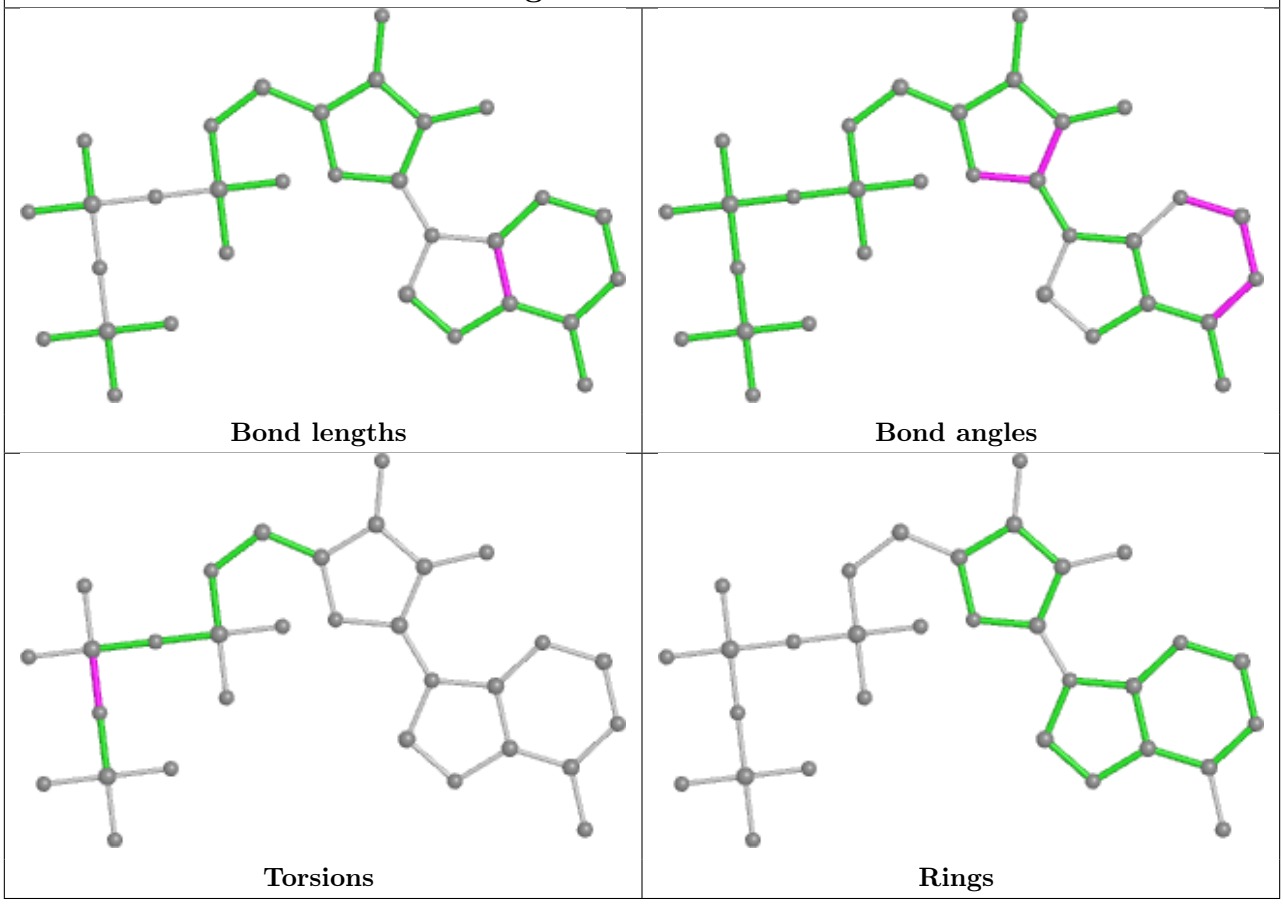


Torsions



Rings

Ligand ATP A 401



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

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	354/377 (93%)	-0.02	16 (4%)	33	38	10, 27, 73, 120	0
1	B	349/377 (92%)	0.18	24 (6%)	16	21	12, 34, 81, 125	0
2	C	225/284 (79%)	0.51	29 (12%)	3	4	13, 45, 91, 119	0
2	G	226/284 (79%)	0.32	19 (8%)	11	14	16, 39, 74, 100	0
All	All	1154/1322 (87%)	0.21	88 (7%)	13	18	10, 35, 82, 125	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	245	GLY	10.7
1	A	203	THR	8.3
1	B	53	TYR	6.9
1	A	201	VAL	6.7
2	C	116	PHE	6.7
1	B	202	THR	6.6
1	B	66	THR	6.2
2	C	115	ASN	6.2
1	B	203	THR	6.2
2	C	119	VAL	6.2
2	C	114	VAL	6.0
1	B	244	ASP	5.7
2	C	118	ASP	5.4
2	G	210	GLY	5.4
1	B	65	LEU	5.3
1	B	243	PRO	4.8
2	C	160	ILE	4.7
2	C	117	GLN	4.6
1	B	201	VAL	4.6
2	G	43	VAL	4.6
1	B	59	GLN	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	202	THR	4.5
1	A	66	THR	4.3
1	B	375	PHE	4.2
2	G	209	ASN	4.2
2	G	37	GLY	4.1
2	G	1	MET	4.0
2	C	223	PRO	4.0
2	G	226	GLN	4.0
1	B	247	VAL	3.8
2	C	224	LYS	3.7
1	B	372	ARG	3.7
2	C	1	MET	3.6
2	G	41	ASP	3.6
1	B	270	GLU	3.5
2	G	36	ASP	3.5
1	B	204	ALA	3.3
2	C	163	TRP	3.3
2	C	206	LEU	3.3
1	A	375	PHE	3.2
2	C	191	THR	3.2
1	A	5	THR	3.2
1	B	67	LEU	3.1
2	G	160	ILE	3.1
2	C	225	LYS	3.1
2	C	35	PRO	3.1
2	C	194	LYS	3.1
1	A	204	ALA	3.0
1	B	5	THR	3.0
1	A	61	LYS	2.9
1	A	360	GLN	2.9
1	A	65	LEU	2.8
1	B	360	GLN	2.8
1	B	68	LYS	2.8
2	C	43	VAL	2.7
2	C	127	TYR	2.7
1	A	53	TYR	2.7
2	G	205	GLY	2.7
1	B	57	GLU	2.7
2	C	195	ARG	2.7
2	G	213	GLU	2.7
1	A	364	GLU	2.6
1	A	60	SER	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	36	ASP	2.6
2	C	210	GLY	2.6
1	A	59	GLN	2.5
2	G	38	SER	2.4
2	C	113	GLU	2.4
1	B	364	GLU	2.4
1	B	60	SER	2.4
2	G	90	ALA	2.4
1	A	51	ASP	2.4
2	G	159	VAL	2.4
1	B	242	LEU	2.3
2	C	41	ASP	2.3
2	C	217	LYS	2.2
1	A	67	LEU	2.2
2	G	158	PHE	2.2
2	G	157	VAL	2.2
2	C	132	SER	2.2
2	G	172	SER	2.2
2	C	159	VAL	2.2
1	B	189	LEU	2.2
2	C	220	GLN	2.1
2	C	213	GLU	2.1
2	C	150	ASN	2.0
2	G	26	VAL	2.0
2	G	93	LYS	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	HIC	A	73	11/12	0.92	0.13	29,40,59,64	0
1	HIC	B	73	11/12	0.93	0.16	41,45,62,64	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

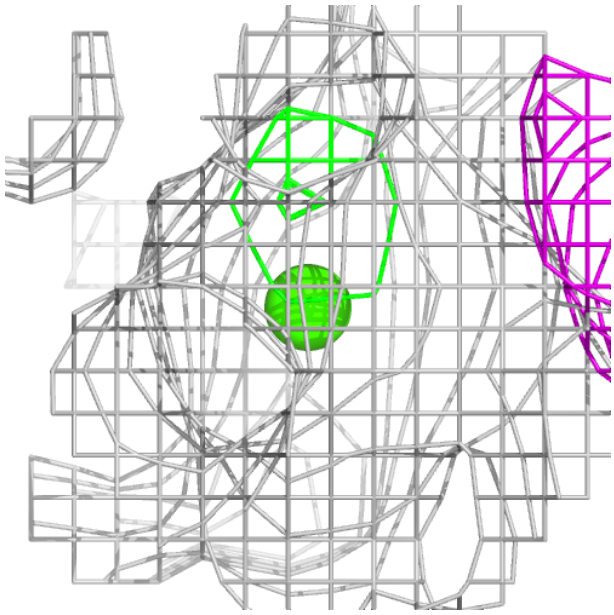
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

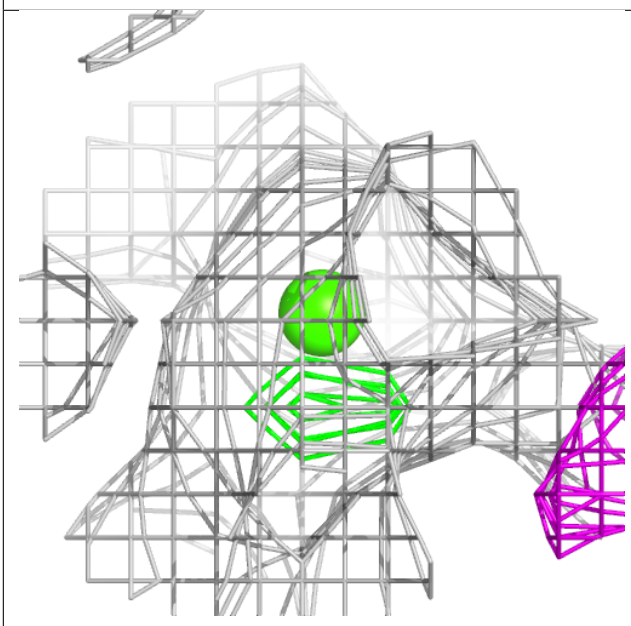
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CA	G	308	1/1	0.65	0.12	74,74,74,74	0
4	CA	C	308	1/1	0.76	0.07	73,73,73,73	0
5	GOL	B	403	6/6	0.76	0.17	49,53,56,56	0
5	GOL	B	404	6/6	0.81	0.16	62,64,64,68	0
4	CA	C	304	1/1	0.84	0.09	56,56,56,56	0
4	CA	C	302	1/1	0.96	0.17	79,79,79,79	0
4	CA	C	306	1/1	0.97	0.04	38,38,38,38	0
4	CA	C	305	1/1	0.97	0.04	38,38,38,38	0
4	CA	C	307	1/1	0.98	0.03	40,40,40,40	0
4	CA	G	306	1/1	0.98	0.04	37,37,37,37	0
3	ATP	B	401	31/31	0.98	0.07	16,22,26,33	0
4	CA	G	302	1/1	0.98	0.04	30,30,30,30	0
4	CA	C	303	1/1	0.99	0.03	29,29,29,29	0
4	CA	G	304	1/1	0.99	0.04	36,36,36,36	0
4	CA	G	305	1/1	0.99	0.03	33,33,33,33	0
4	CA	G	301	1/1	0.99	0.09	21,21,21,21	0
4	CA	G	307	1/1	0.99	0.09	29,29,29,29	0
3	ATP	A	401	31/31	0.99	0.07	13,16,22,22	0
4	CA	C	301	1/1	0.99	0.09	15,15,15,15	0
4	CA	G	303	1/1	0.99	0.02	31,31,31,31	0
4	CA	A	402	1/1	1.00	0.10	19,19,19,19	0
4	CA	B	402	1/1	1.00	0.07	22,22,22,22	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 G 308:

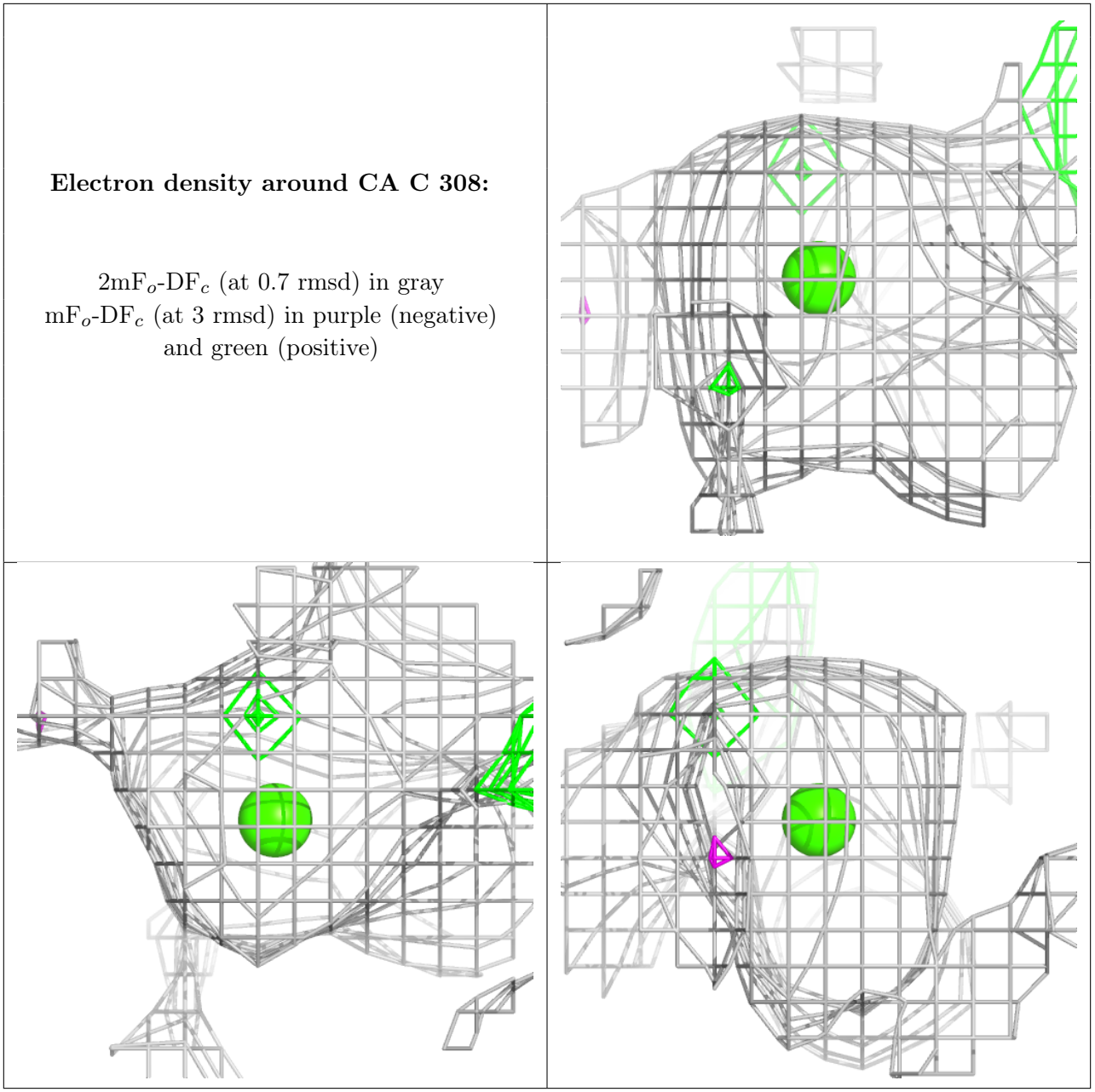
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





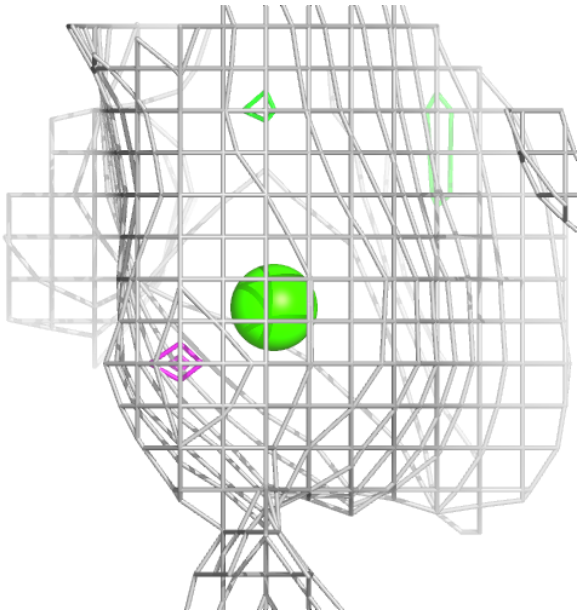
Electron density around CA C 308:

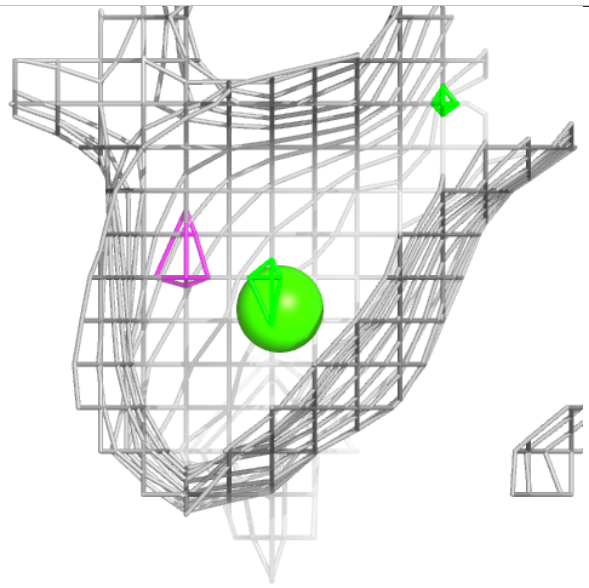
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

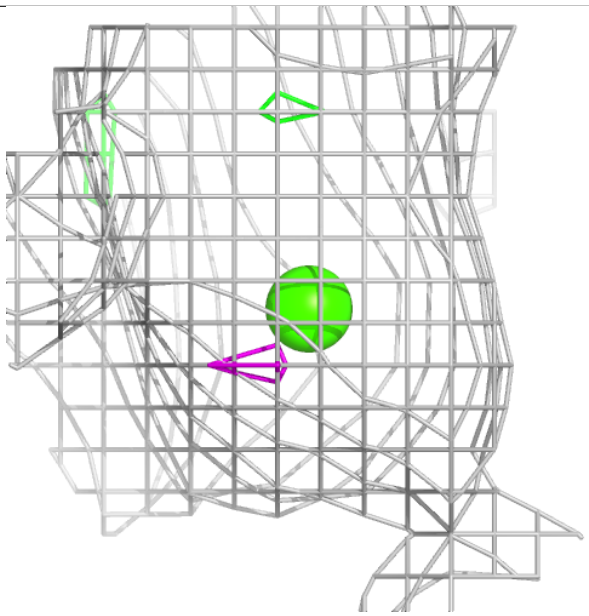


Electron density around CA C 304:

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

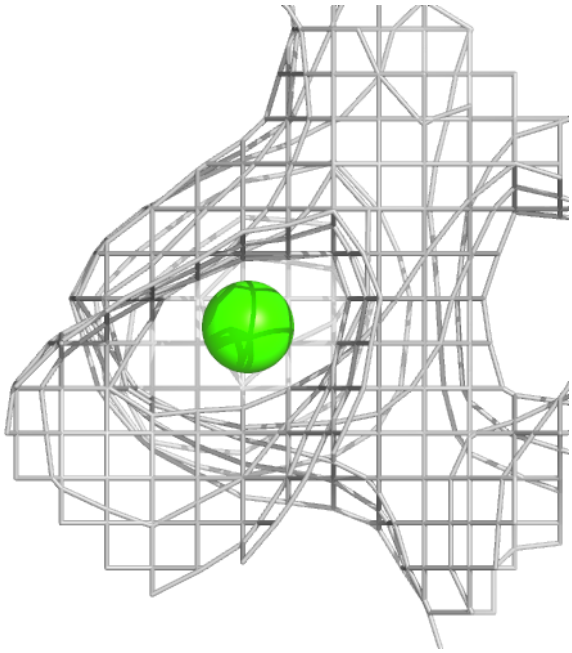


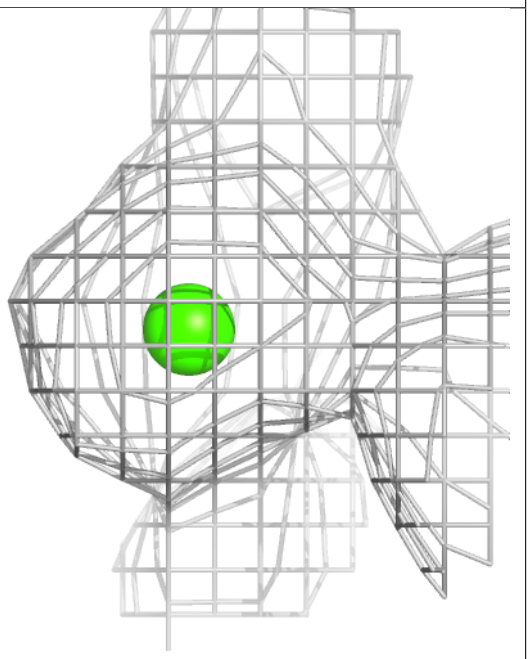




Electron density around CA C 302:

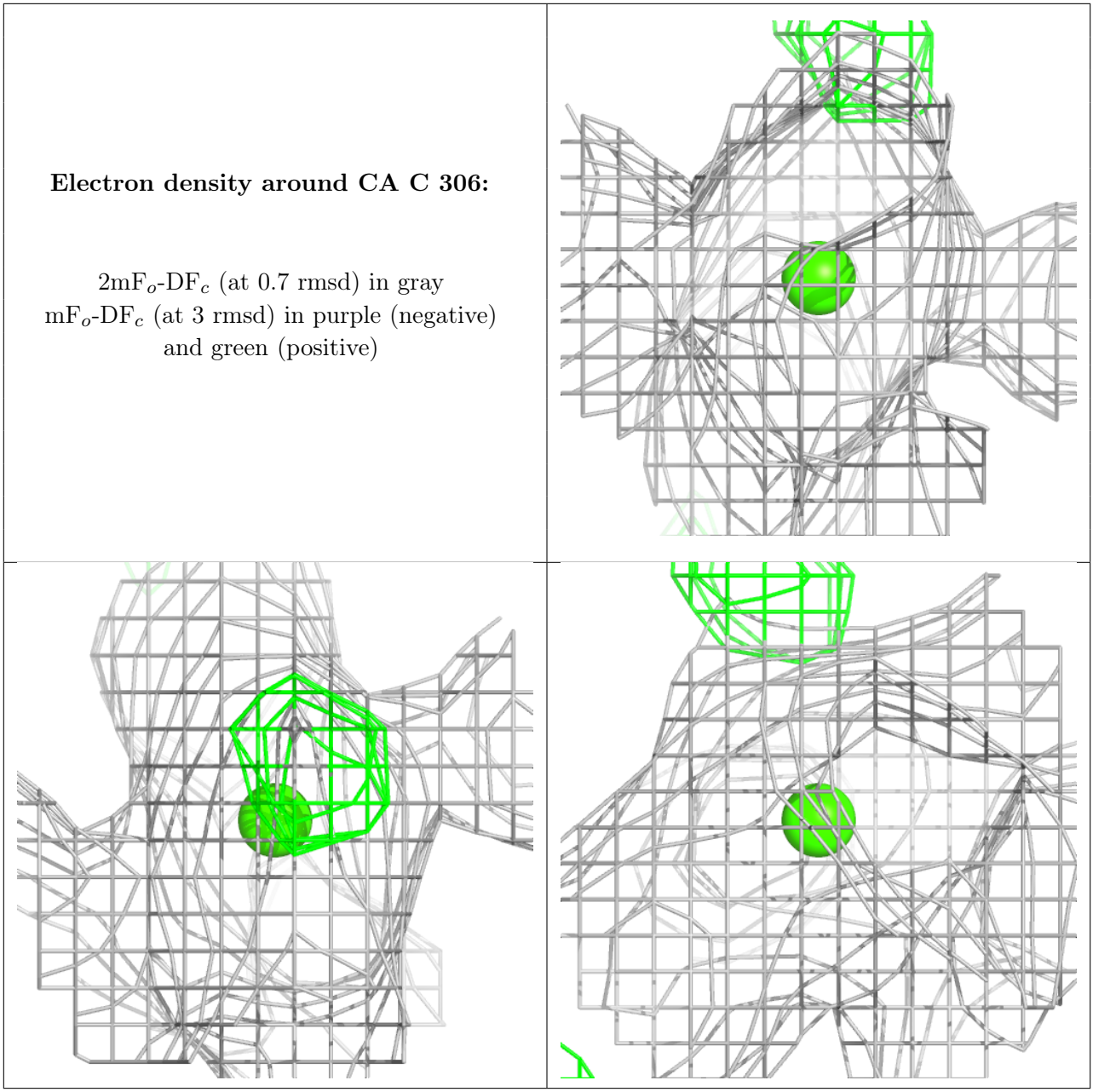
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





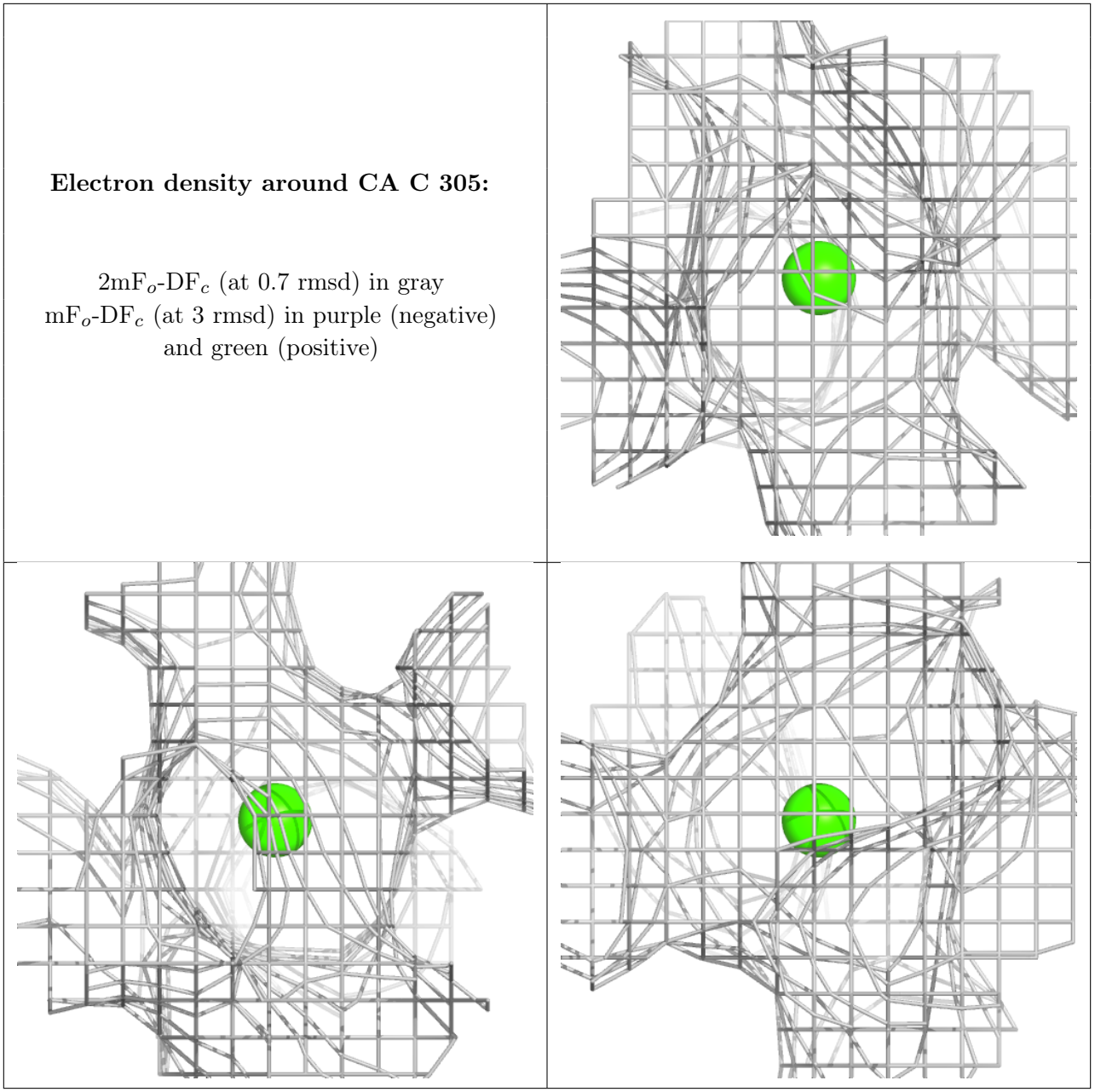
Electron density around CA C 306:

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



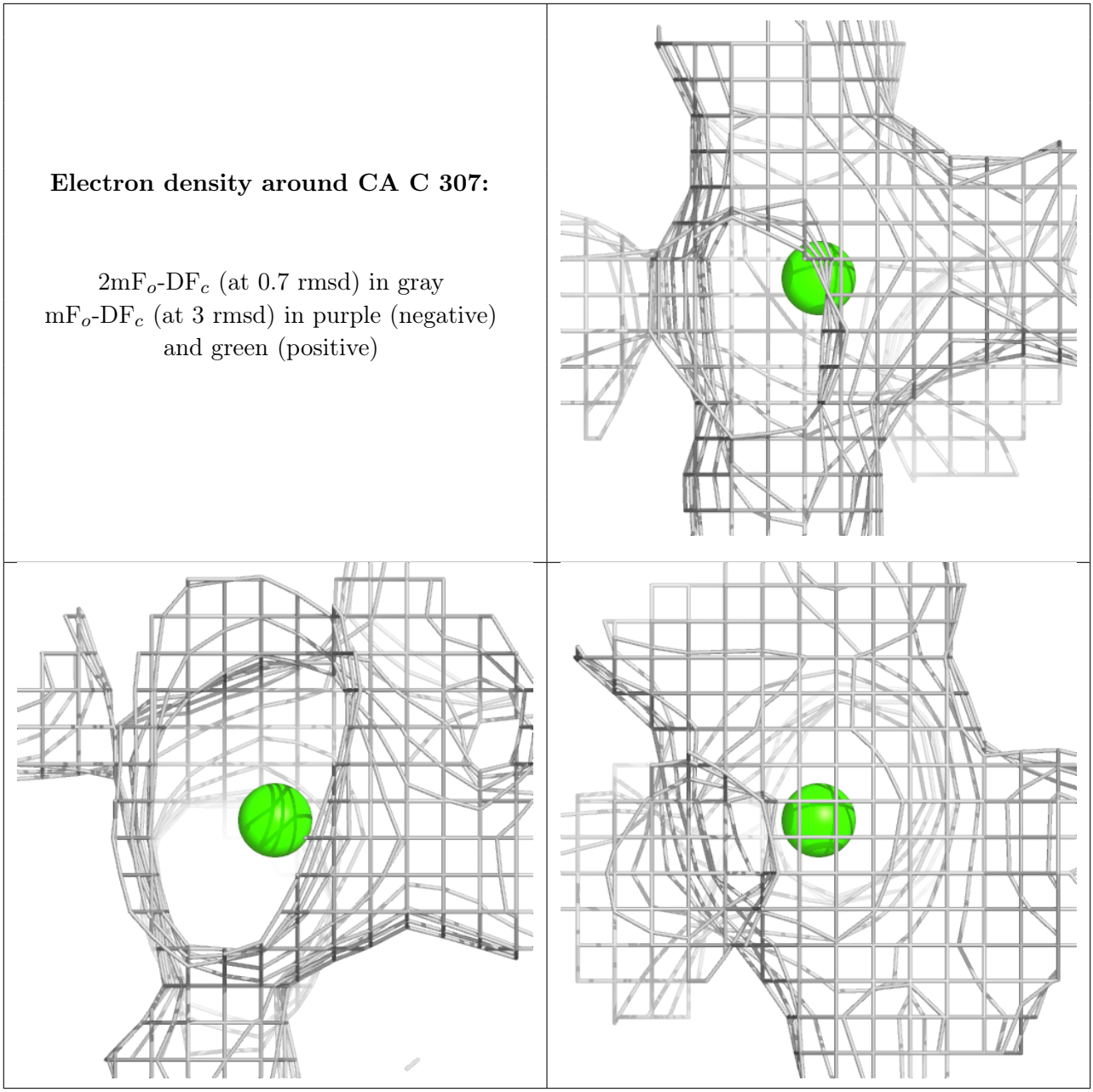
Electron density around CA C 305:

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



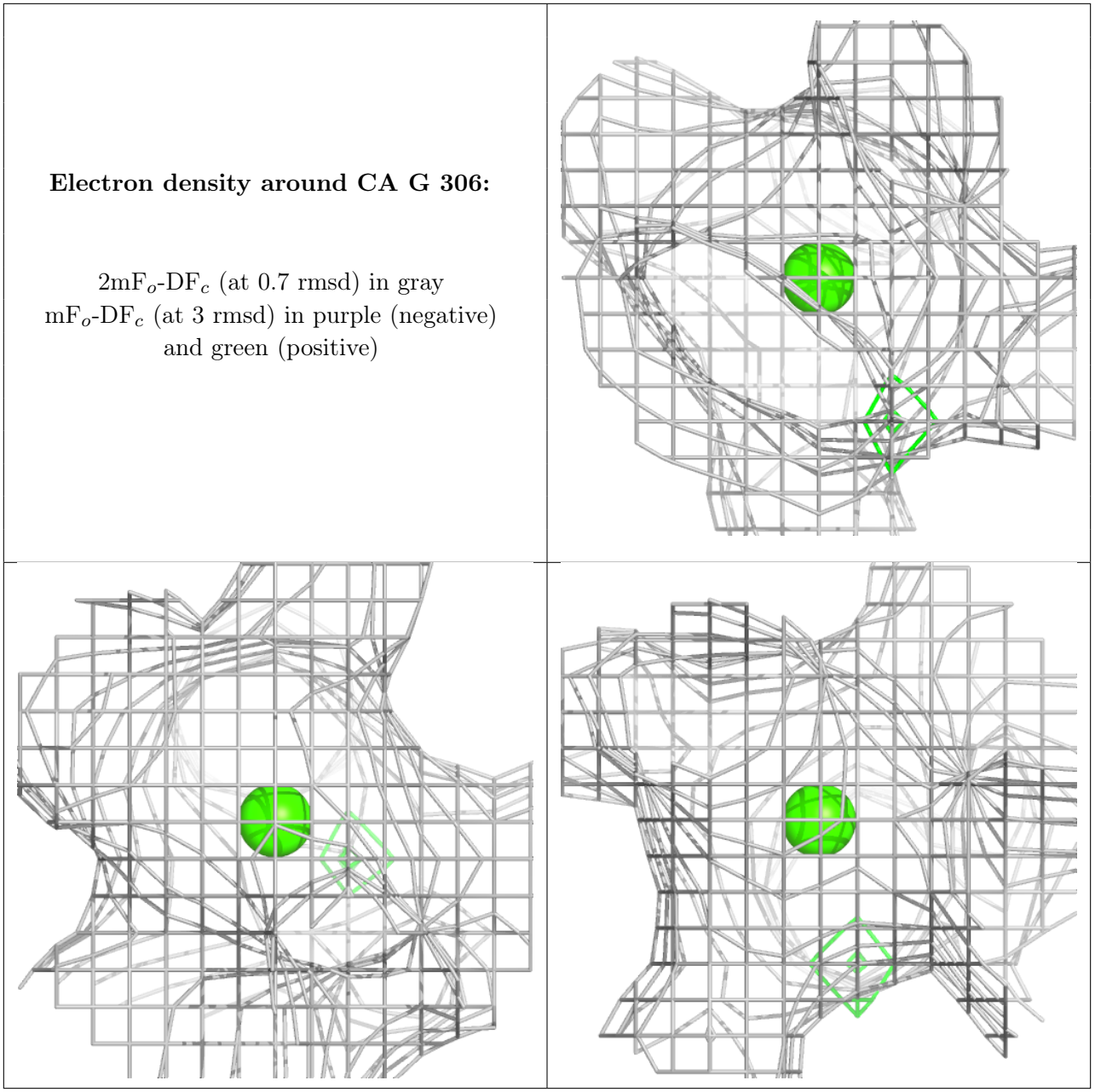
Electron density around CA C 307:

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



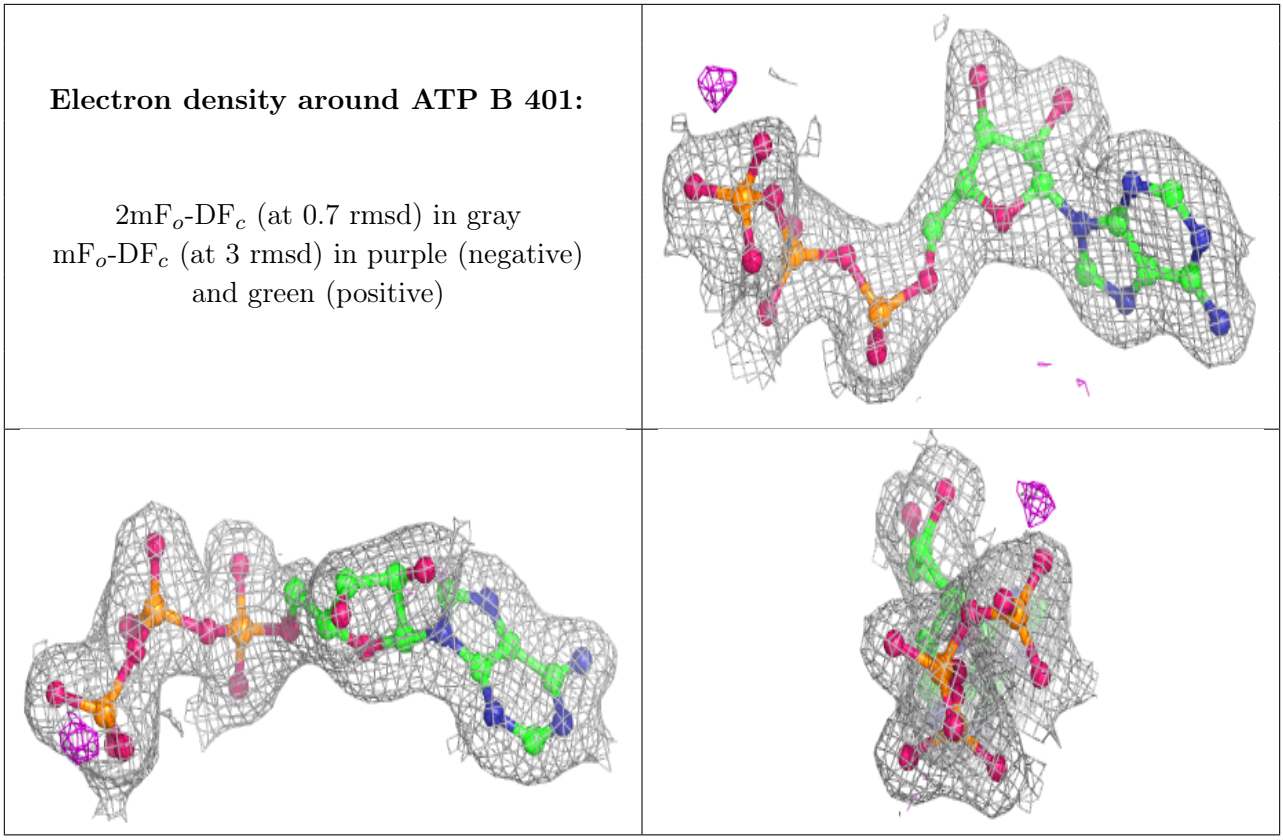
Electron density around CA G 306:

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



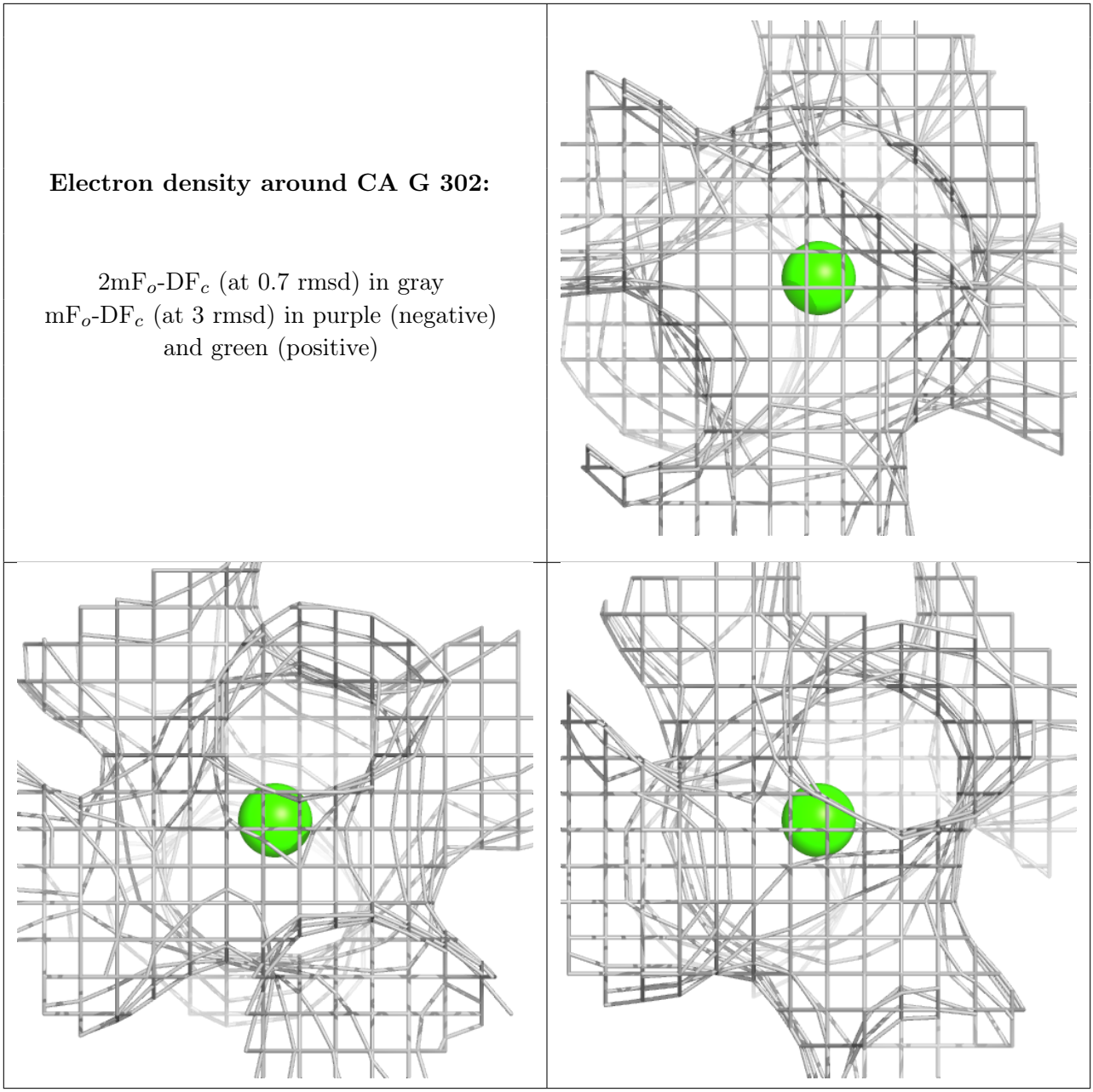
Electron density around ATP B 401:

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



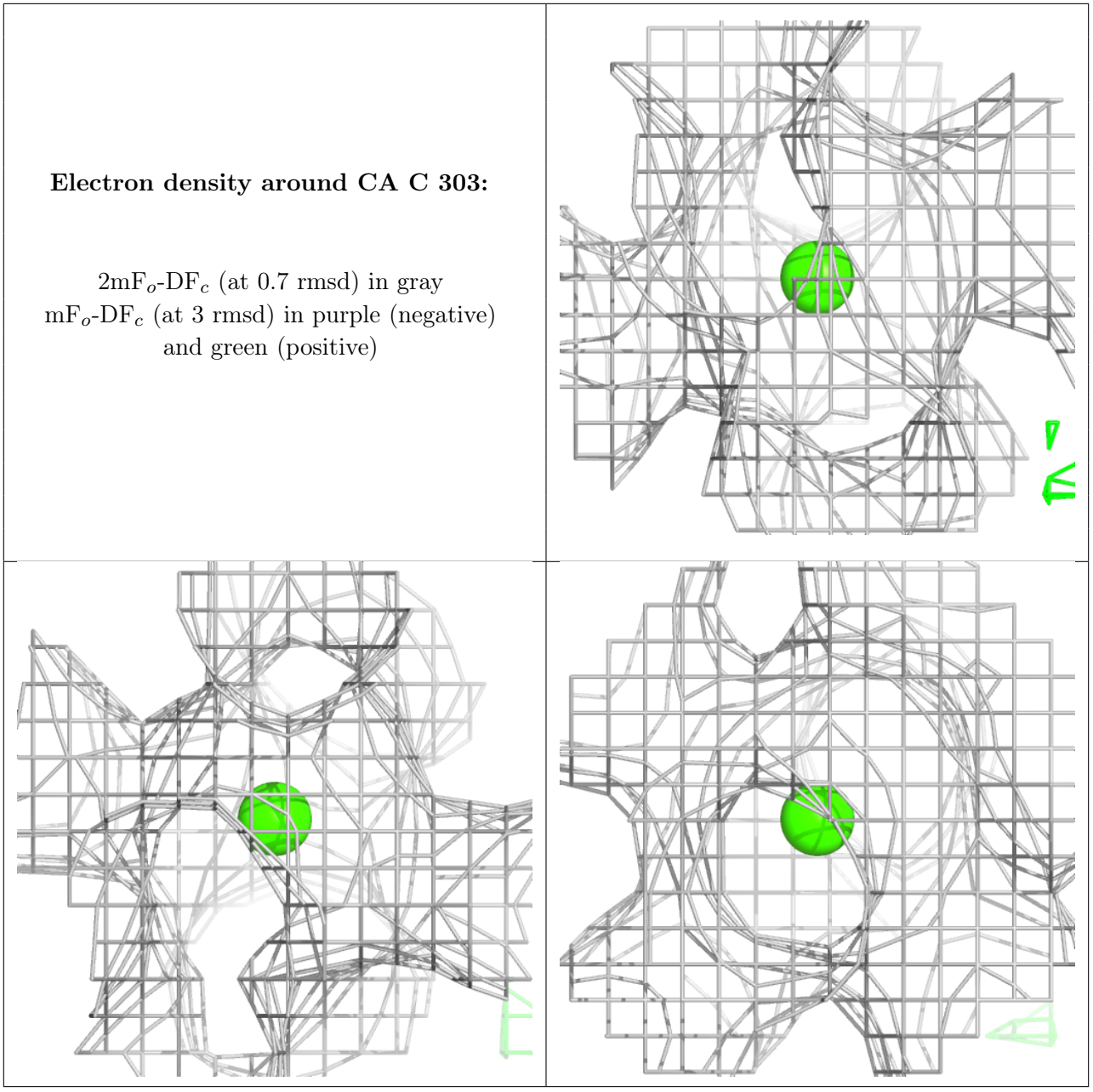
Electron density around CA G 302:

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



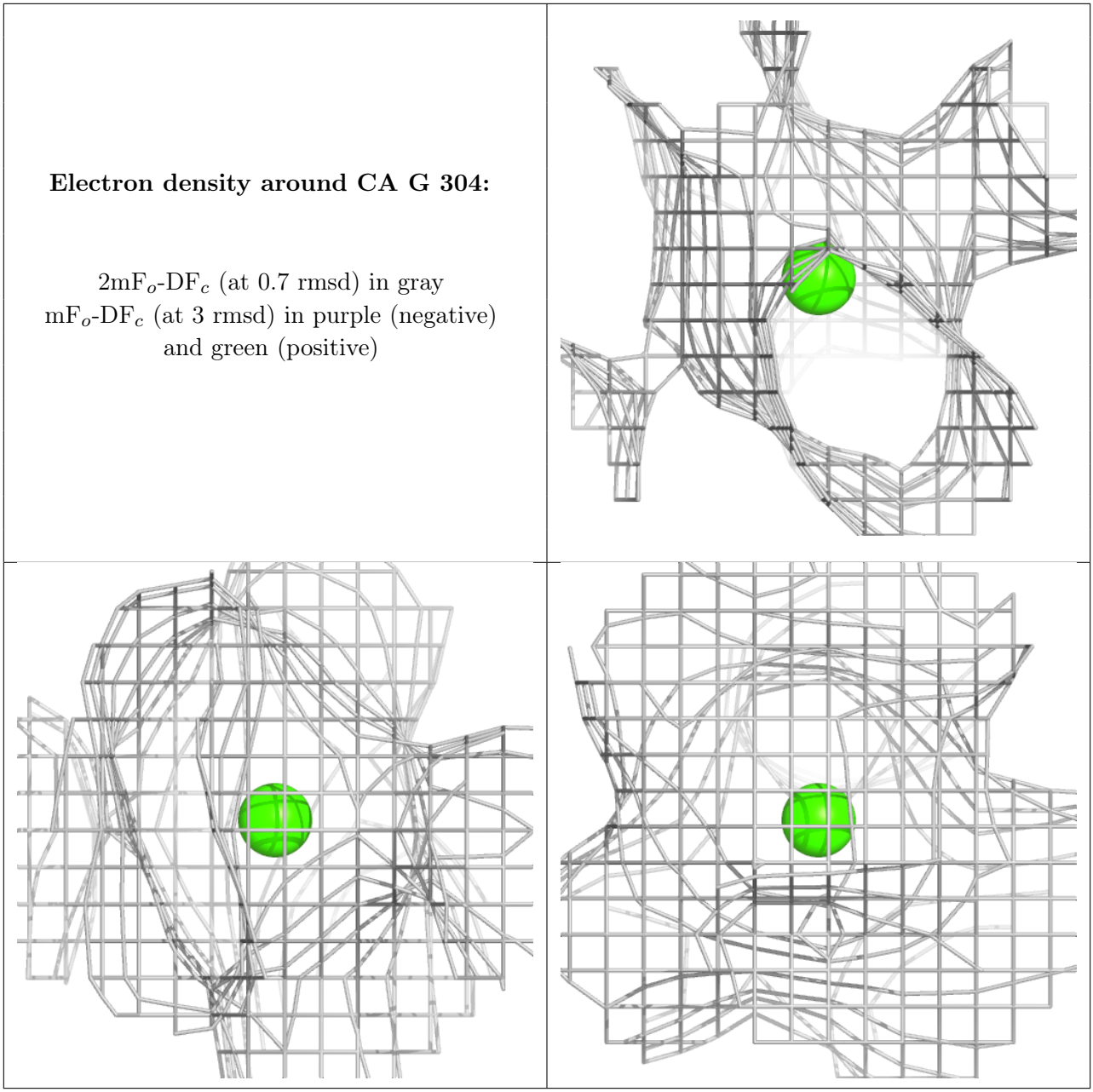
Electron density around CA C 303:

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



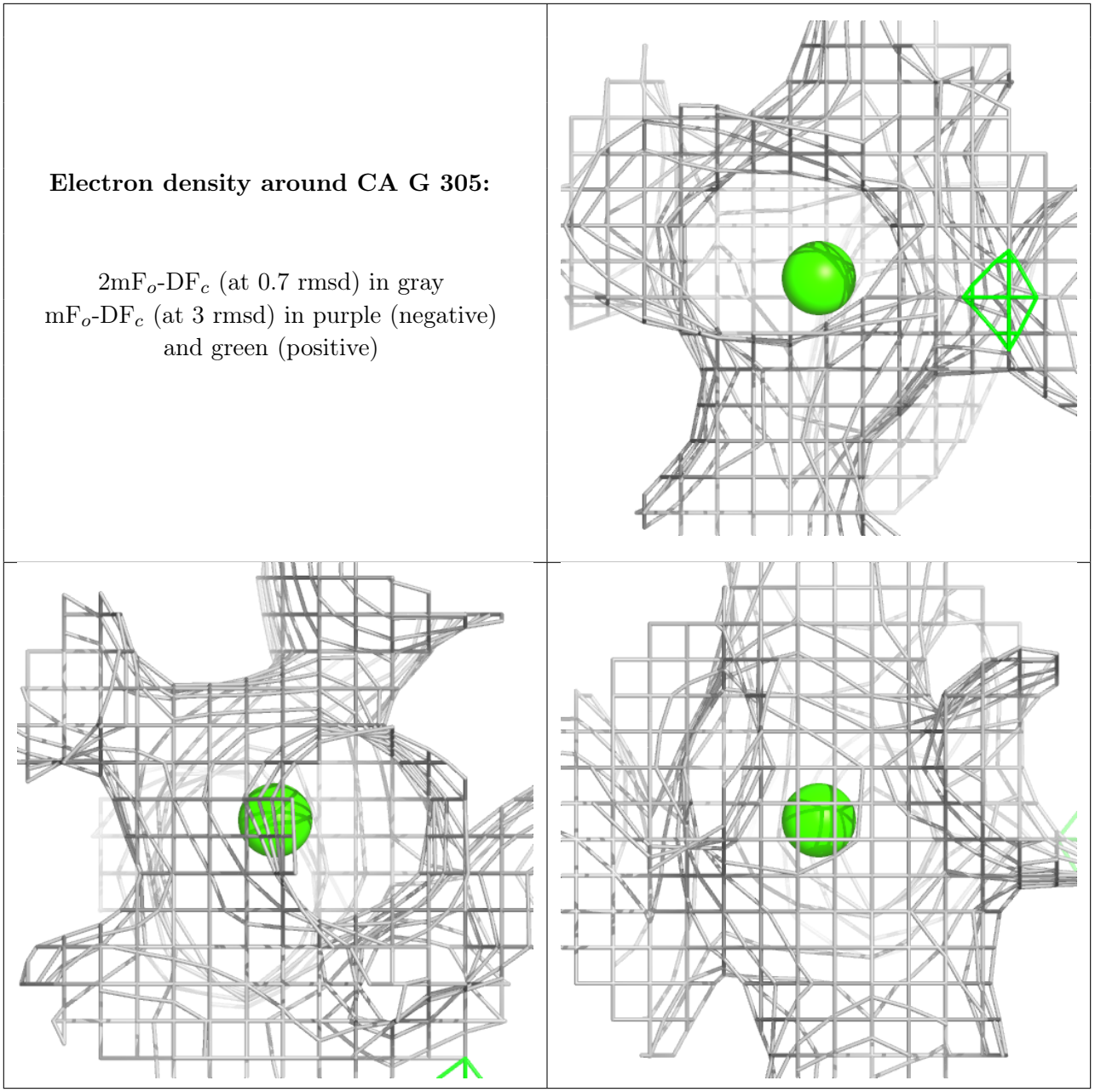
Electron density around CA G 304:

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



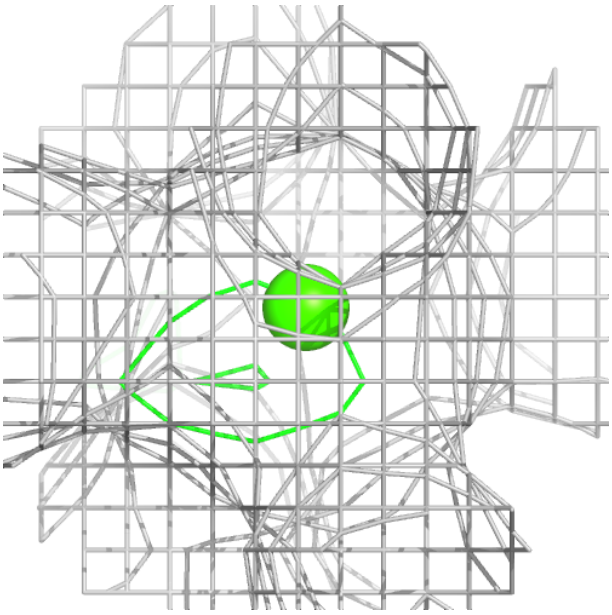
Electron density around CA G 305:

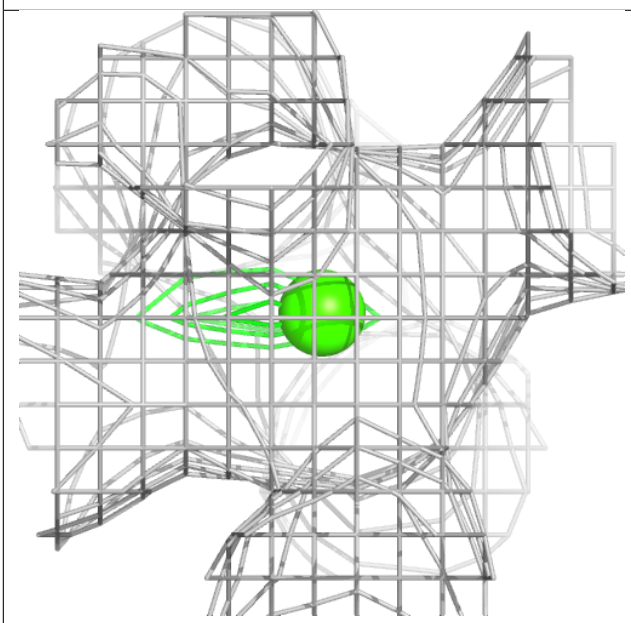
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

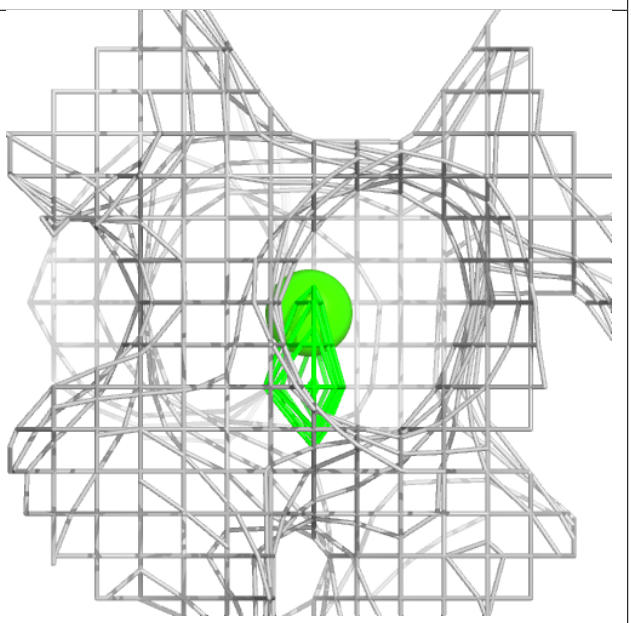


Electron density around CA G 301:

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

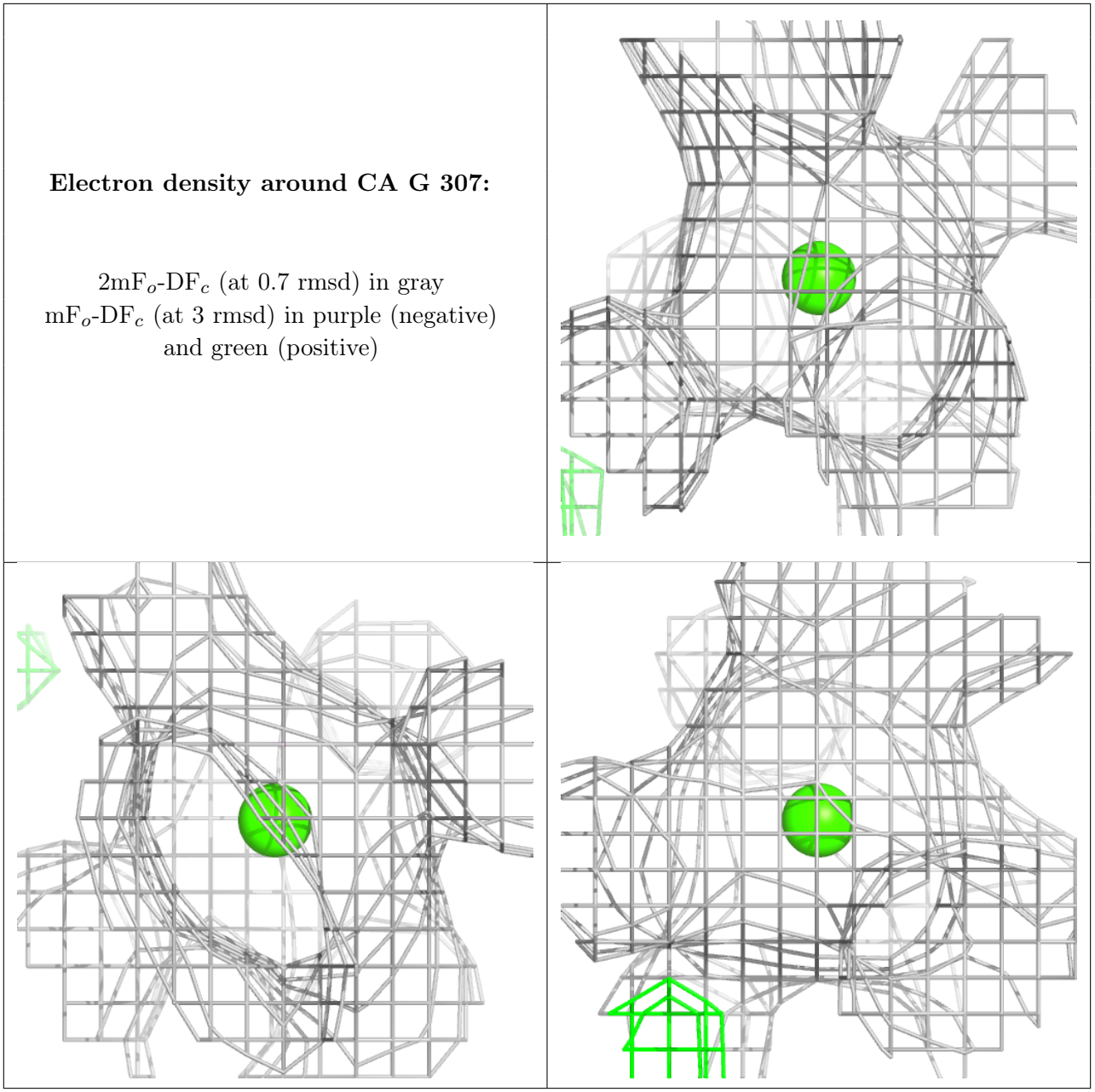






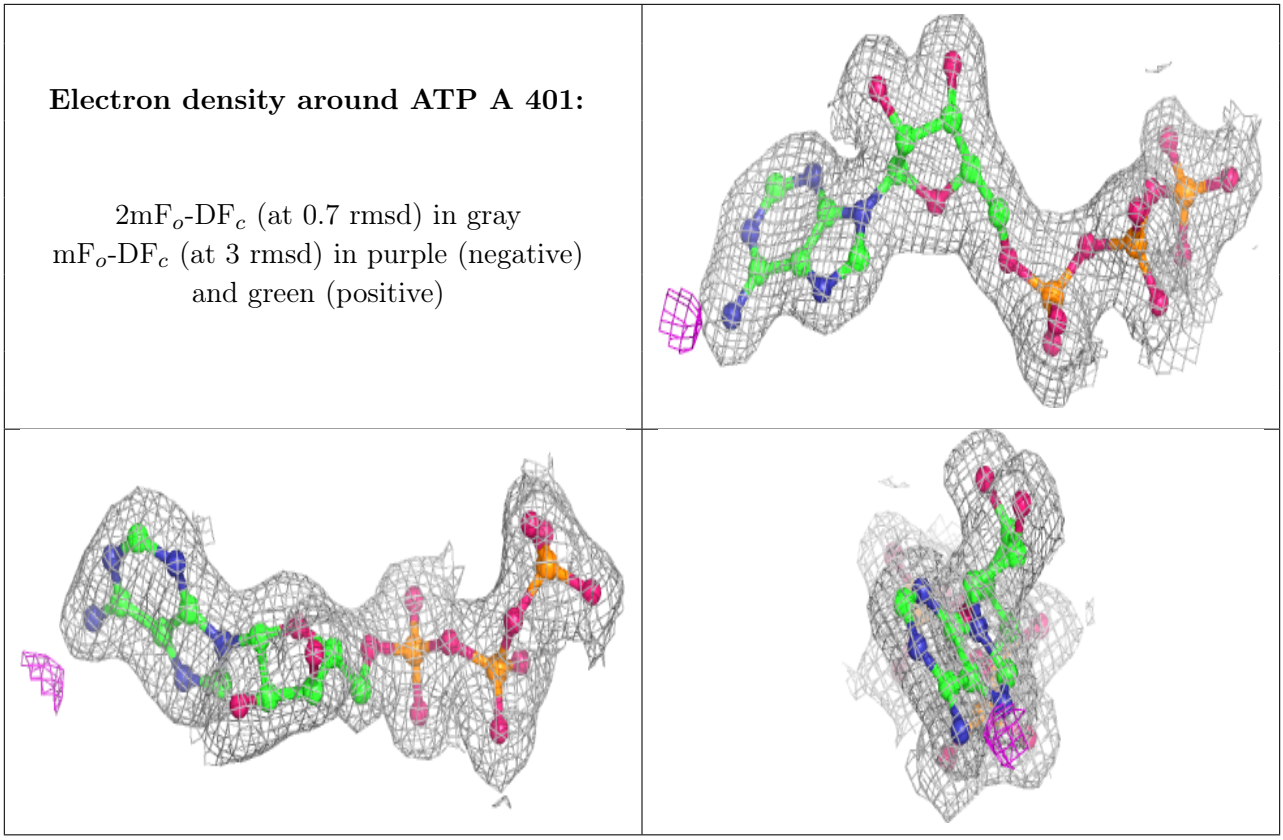
Electron density around CA G 307:

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



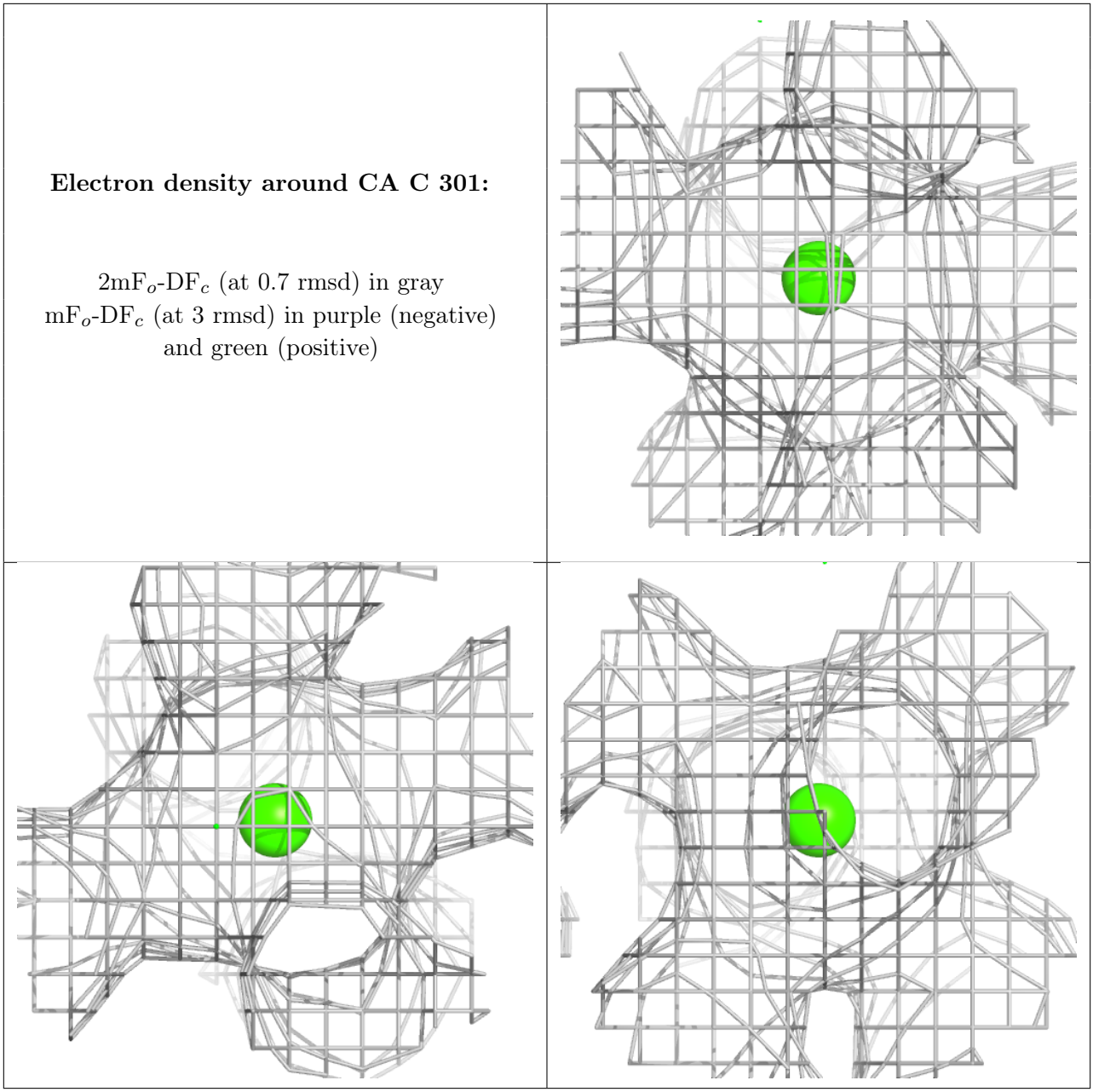
Electron density around ATP A 401:

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



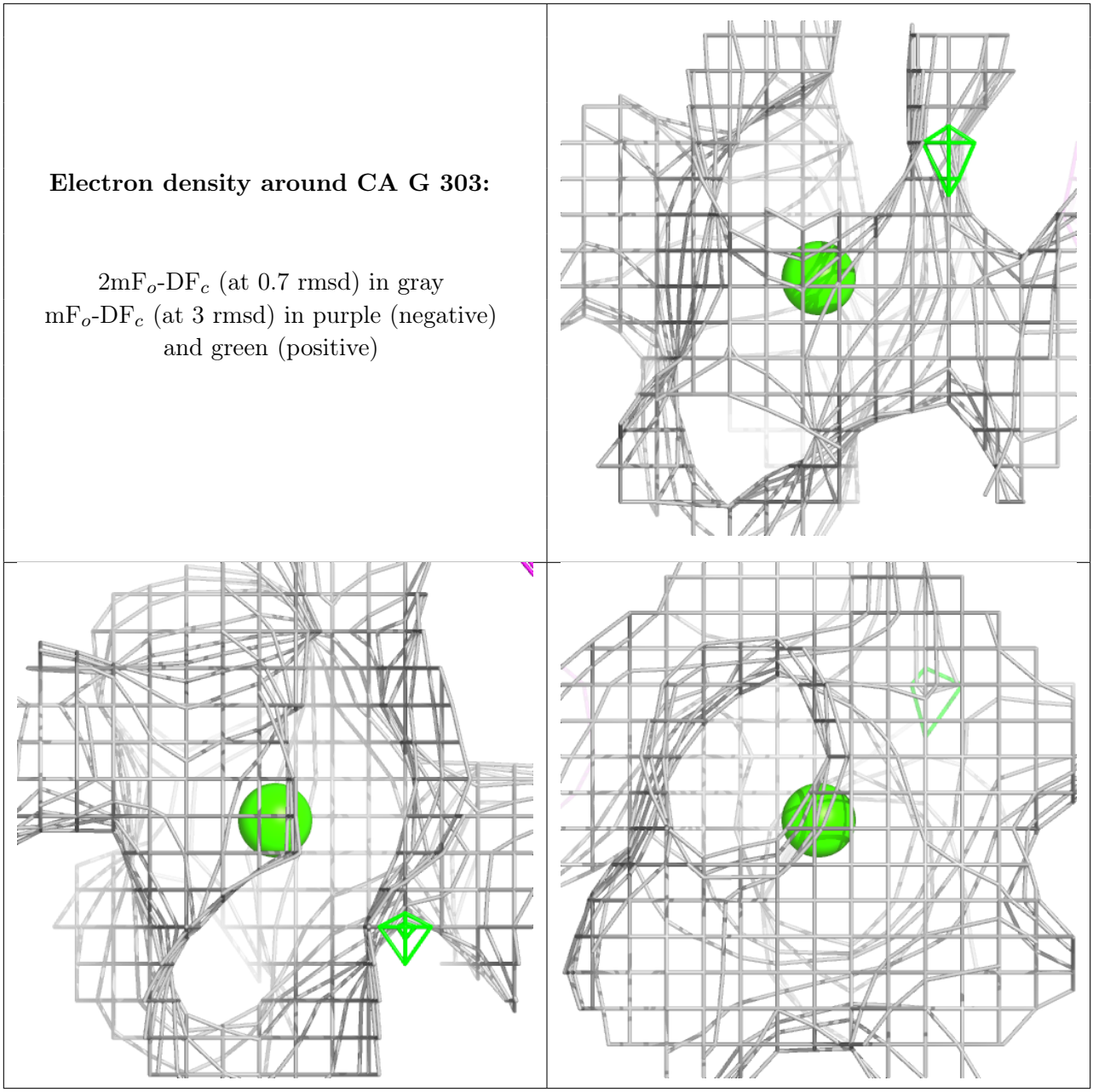
Electron density around CA C 301:

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



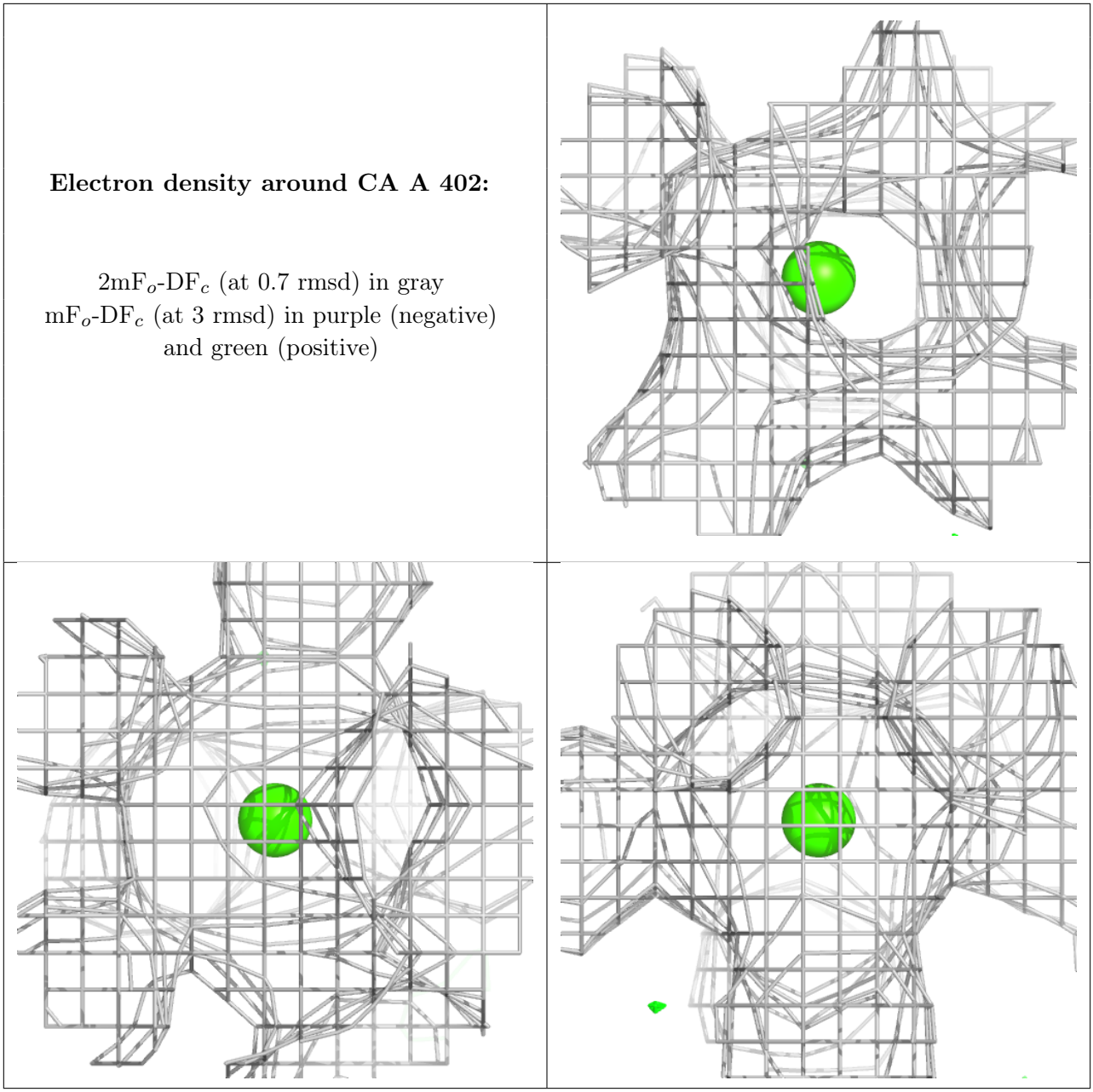
Electron density around CA G 303:

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



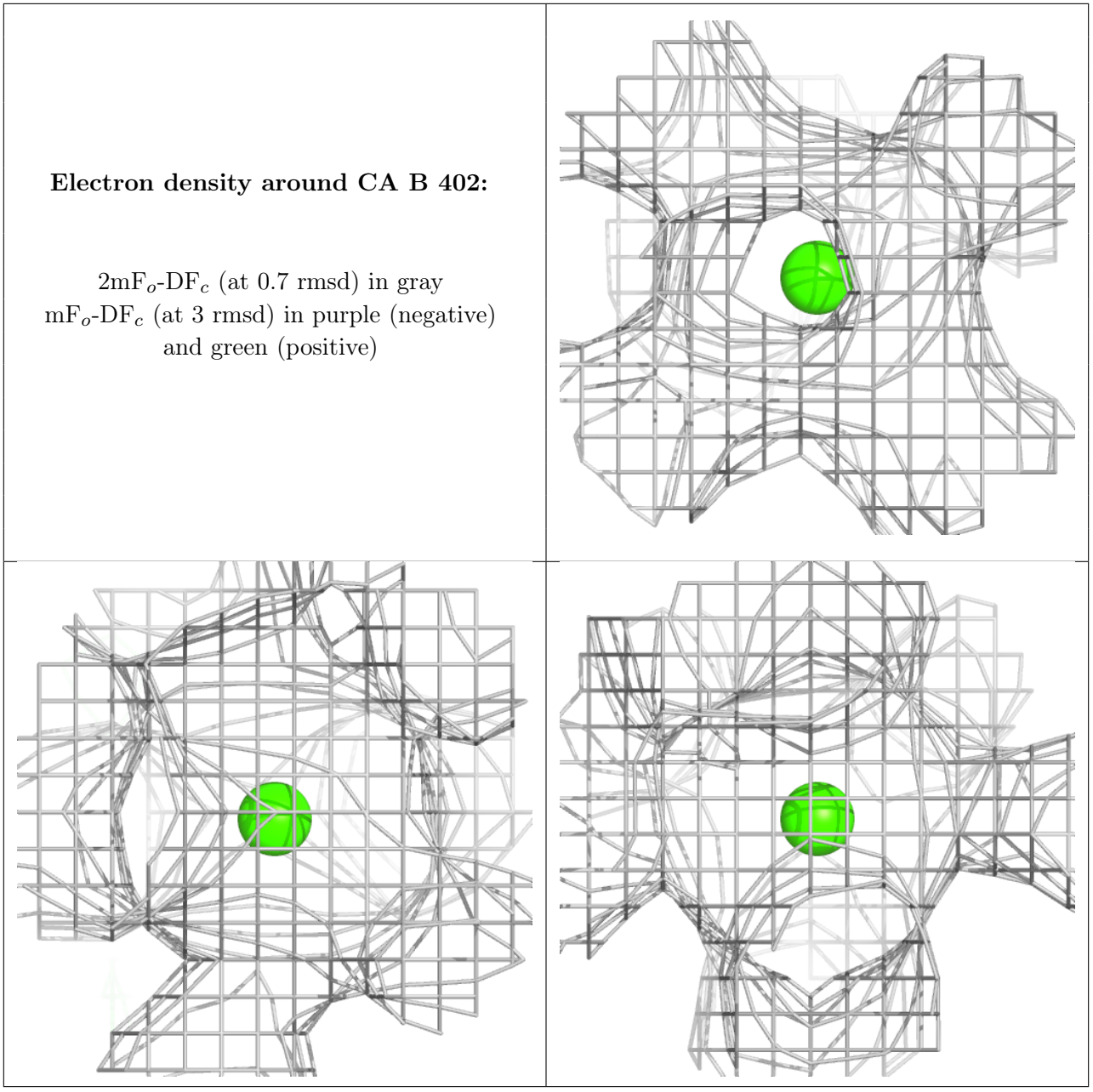
Electron density around CA A 402:

$2mF_o$-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 402:

$2mF_o$-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.