wwPDB X-ray Structure Validation Summary Report (i)

Mar 10, 2026 – 09:54 AM UTC

PDB ID	:	8Y0M / pdb_00008y0m
Title	:	beta-glucosidase mutant M279V_T308S_K361R_D433N_N514C
Authors	:	Matsuzaki, C.; Katayama, T.
Deposited on	:	2024-01-22
Resolution	:	2.30 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:

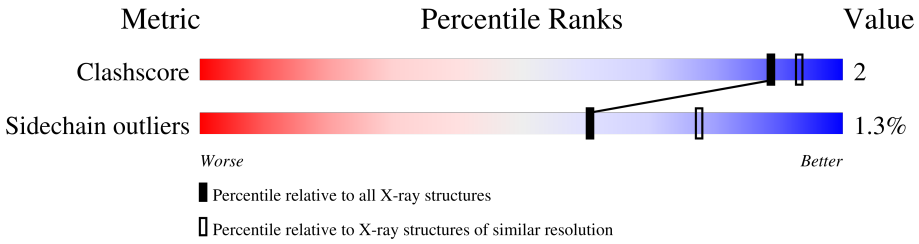
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	NOT EXECUTED
Xtriage (Phenix)	:	2.0
EDS	:	NOT EXECUTED
Buster-report	:	NOT EXECUTED
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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))
Clashscore	190562	6919 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	861	86% 5% 9%
1	B	861	86% • 9%
1	C	861	85% 5% 9%
1	D	861	86% 5% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 26960 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 beta-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	781	Total	C	N	O	S	0	0	0
			5976	3783	1017	1152	24			
1	B	781	Total	C	N	O	S	0	0	0
			5976	3783	1017	1152	24			
1	C	781	Total	C	N	O	S	0	0	0
			5976	3783	1017	1152	24			
1	D	781	Total	C	N	O	S	0	0	0
			5976	3783	1017	1152	24			

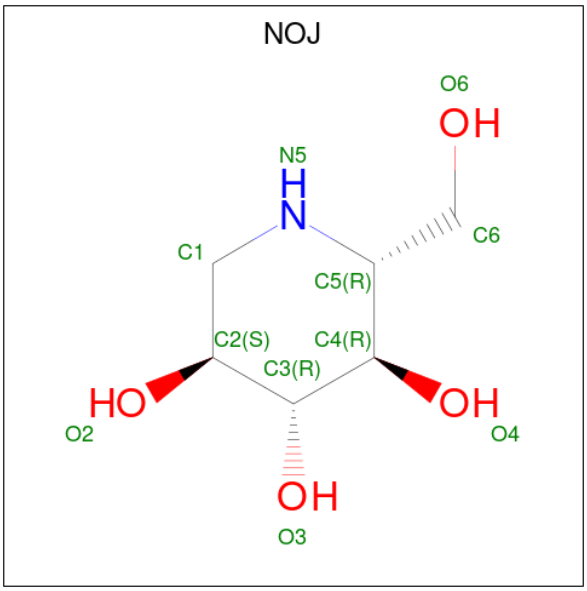
There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	279	VAL	MET	engineered mutation	UNP Q0ZUL0
A	308	SER	THR	engineered mutation	UNP Q0ZUL0
A	361	ARG	LYS	engineered mutation	UNP Q0ZUL0
A	433	ASN	ASP	engineered mutation	UNP Q0ZUL0
A	514	CYS	ASN	engineered mutation	UNP Q0ZUL0
B	279	VAL	MET	engineered mutation	UNP Q0ZUL0
B	308	SER	THR	engineered mutation	UNP Q0ZUL0
B	361	ARG	LYS	engineered mutation	UNP Q0ZUL0
B	433	ASN	ASP	engineered mutation	UNP Q0ZUL0
B	514	CYS	ASN	engineered mutation	UNP Q0ZUL0
C	279	VAL	MET	engineered mutation	UNP Q0ZUL0
C	308	SER	THR	engineered mutation	UNP Q0ZUL0
C	361	ARG	LYS	engineered mutation	UNP Q0ZUL0
C	433	ASN	ASP	engineered mutation	UNP Q0ZUL0
C	514	CYS	ASN	engineered mutation	UNP Q0ZUL0
D	279	VAL	MET	engineered mutation	UNP Q0ZUL0
D	308	SER	THR	engineered mutation	UNP Q0ZUL0
D	361	ARG	LYS	engineered mutation	UNP Q0ZUL0
D	433	ASN	ASP	engineered mutation	UNP Q0ZUL0
D	514	CYS	ASN	engineered mutation	UNP Q0ZUL0

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Mg 2 2	0	0
2	B	3	Total Mg 3 3	0	0
2	C	2	Total Mg 2 2	0	0
2	D	3	Total Mg 3 3	0	0

- Molecule 3 is 1-DEOXYNOJIRIMYCIN (CCD ID: NOJ) (formula: $C_6H_{13}NO_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 11 6 1 4	0	0
3	B	1	Total C N O 11 6 1 4	0	0
3	C	1	Total C N O 11 6 1 4	0	0
3	D	1	Total C N O 11 6 1 4	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	773	Total O 773 773	0	0
4	B	801	Total O 801 801	0	0
4	C	715	Total O 715 715	0	0
4	D	713	Total O 713 713	0	0

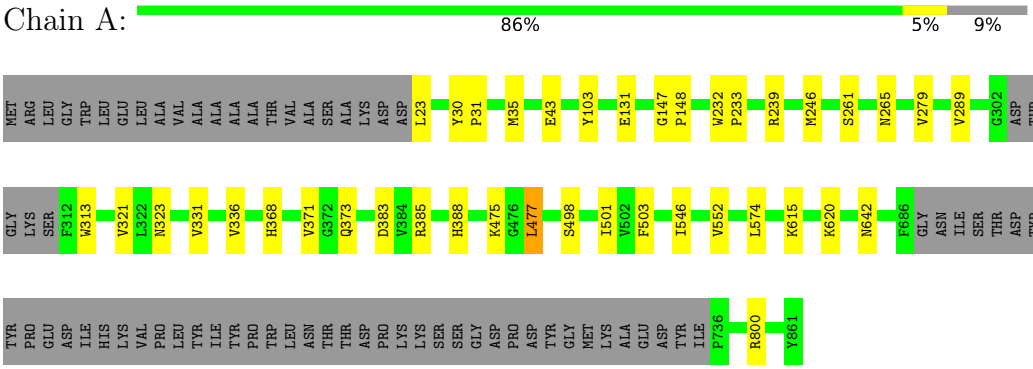
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

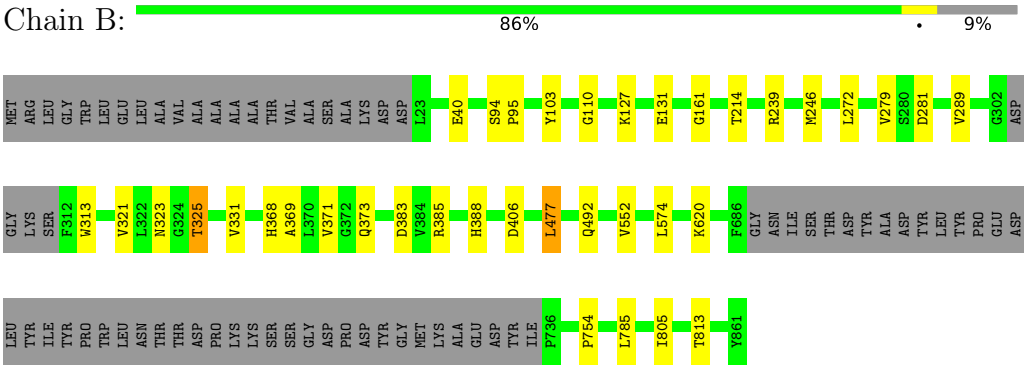
• Molecule 1: beta-glucosidase

Chain A:



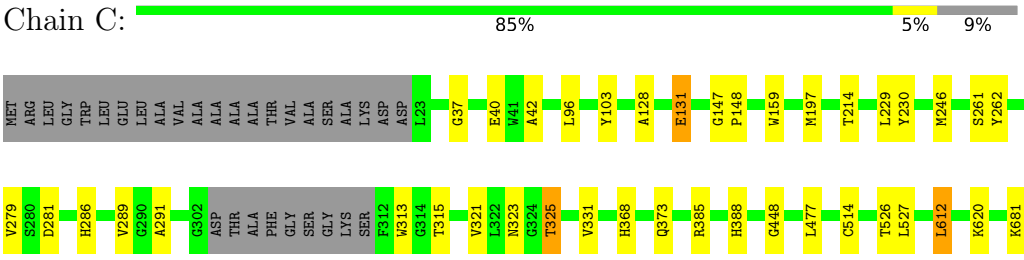
• Molecule 1: beta-glucosidase

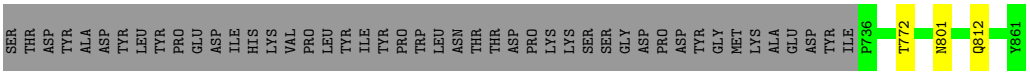
Chain B:



• Molecule 1: beta-glucosidase

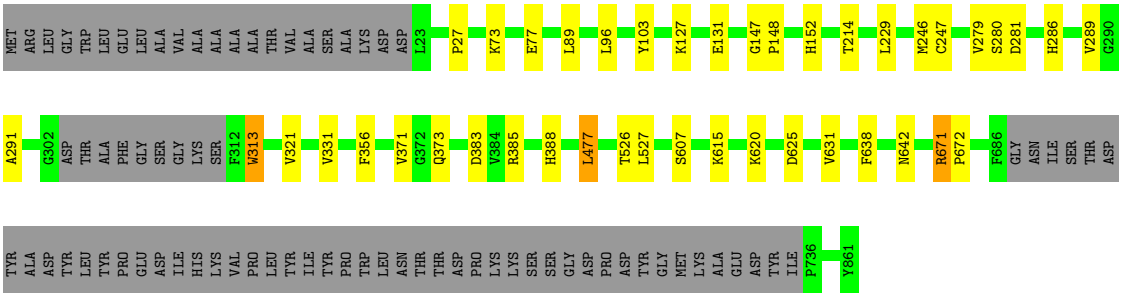
Chain C:





- Molecule 1: beta-glucosidase

Chain D:



4 Data and refinement statistics (i)

EDS was not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	138.19Å 150.16Å 174.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.20 – 2.30	Depositor
% Data completeness (in resolution range)	99.9 (40.20-2.30)	Depositor
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.18 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.162 , 0.180	Depositor
Wilson B-factor (Å^2)	21.7	Xtriage
Anisotropy	0.061	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	26960	wwPDB-VP
Average B, all atoms (Å^2)	23.0	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 38.83 % 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 3.4744e-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: NOJ, MG

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	0.46	0/6133	0.85	0/8357
1	B	0.46	0/6133	0.84	0/8357
1	C	0.46	0/6133	0.85	1/8357 (0.0%)
1	D	0.47	0/6133	0.84	1/8357 (0.0%)
All	All	0.46	0/24532	0.84	2/33428 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	131	GLU	N-CA-CB	5.96	118.89	110.12
1	D	27	PRO	N-CA-CB	5.50	106.27	103.19

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	5976	0	5718	22	0
1	B	5976	0	5718	25	0
1	C	5976	0	5718	29	0
1	D	5976	0	5718	26	0
2	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	3	0	0	1	0
3	A	11	0	13	0	0
3	B	11	0	13	1	0
3	C	11	0	13	1	0
3	D	11	0	13	1	0
4	A	773	0	0	1	0
4	B	801	0	0	1	0
4	C	715	0	0	1	0
4	D	713	0	0	3	0
All	All	26960	0	22924	89	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 2.

The worst 5 of 89 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:323:ASN:OD1	1:B:325:THR:HG23	1.79	0.82
2:D:903:MG:MG	4:D:1001:HOH:O	1.24	0.80
1:A:620:LYS:H	1:D:373:GLN:HE22	1.37	0.70
1:D:625:ASP:OD2	4:D:1001:HOH:O	2.10	0.69
1:D:152:HIS:NE2	4:D:1001:HOH:O	2.27	0.67

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

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	631/695 (91%)	621 (98%)	10 (2%)	55	73
1	B	631/695 (91%)	624 (99%)	7 (1%)	65	81
1	C	631/695 (91%)	623 (99%)	8 (1%)	61	77
1	D	631/695 (91%)	624 (99%)	7 (1%)	65	81
All	All	2524/2780 (91%)	2492 (99%)	32 (1%)	61	77

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	477	LEU
1	D	615	LYS
1	B	313	TRP
1	B	289	VAL
1	D	642	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	254	ASN
1	D	642	ASN
1	D	368	HIS
1	D	492	GLN
1	B	642	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)

Mogul was not executed - this section is therefore empty.

5.5 Carbohydrates (i)

Mogul was not executed - this section is therefore empty.

5.6 Ligand geometry (i)

Mogul was not executed - this section is therefore empty.

5.7 Other polymers (i)

Mogul was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

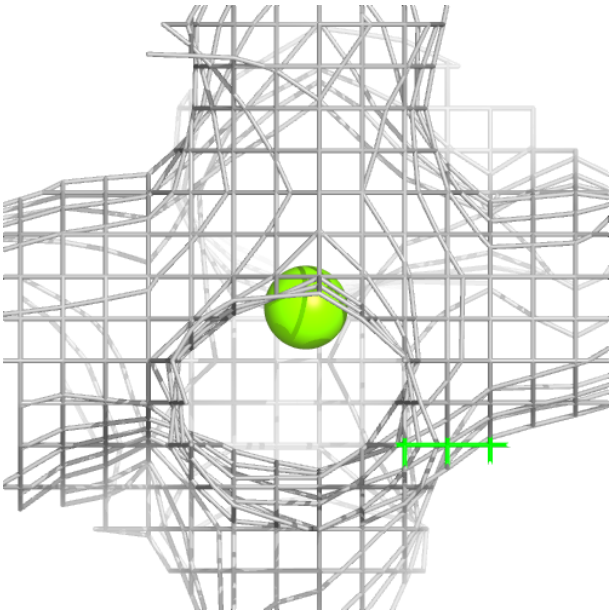
6.4 Ligands

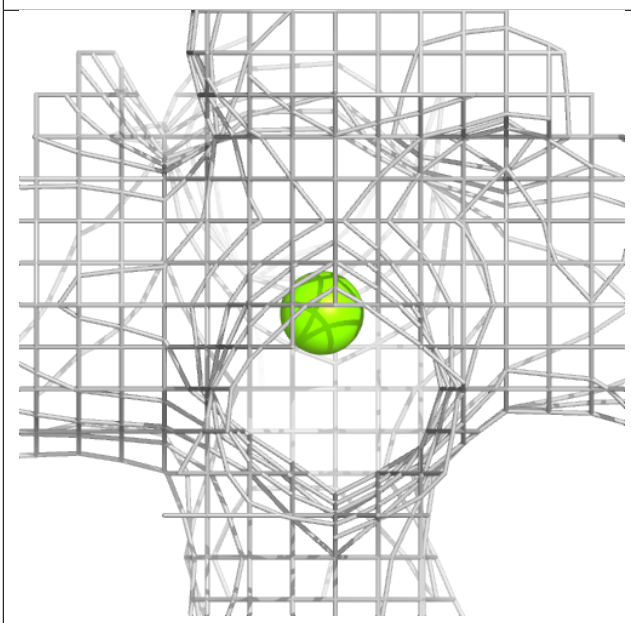
EDS was not executed - this section is therefore empty.

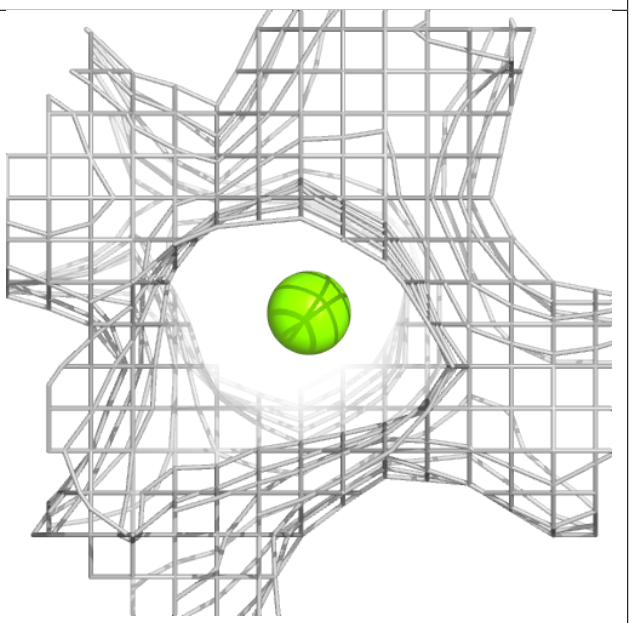
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around MG A 901:

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

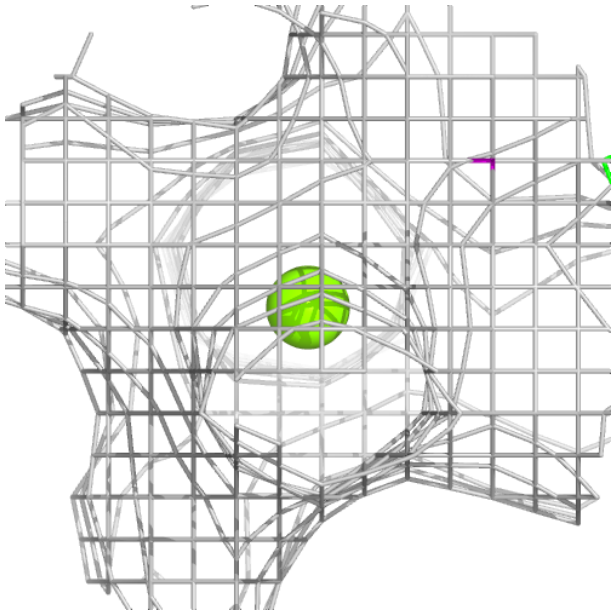


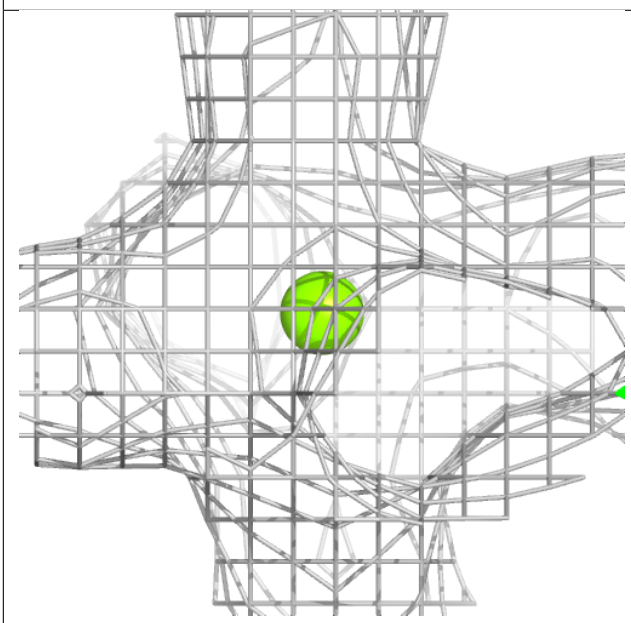




Electron density around MG A 902:

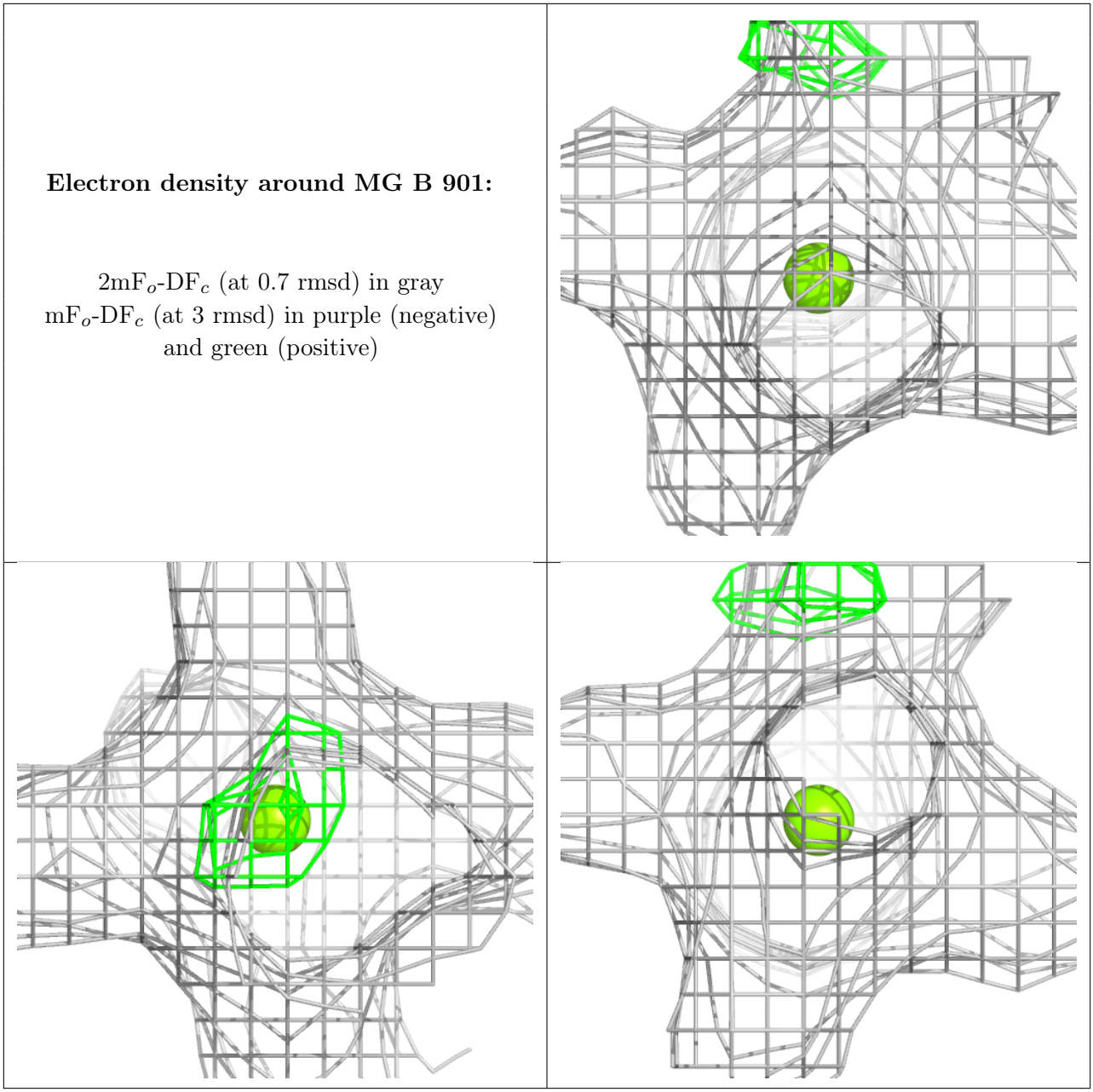
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





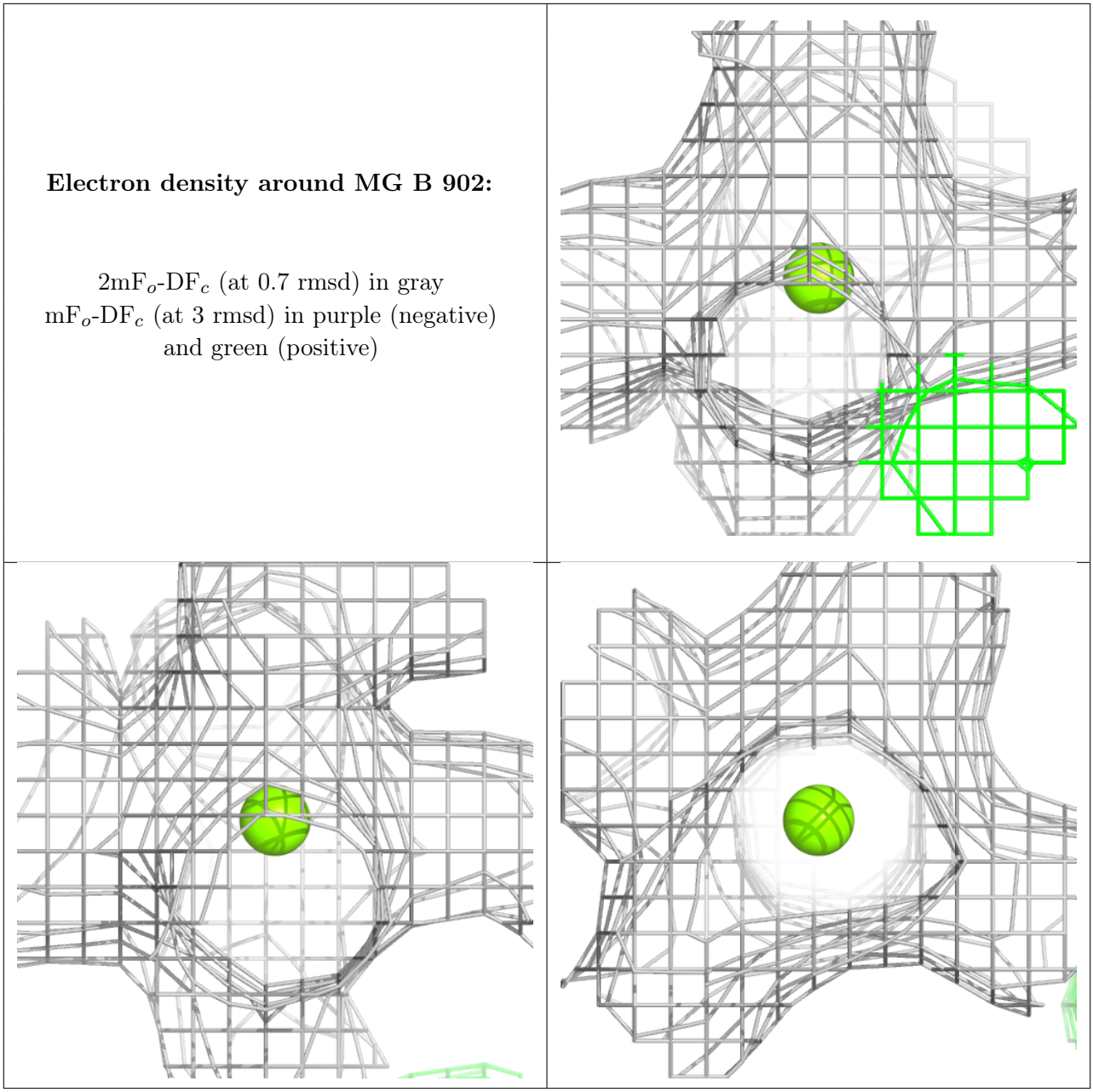
Electron density around MG B 901:

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



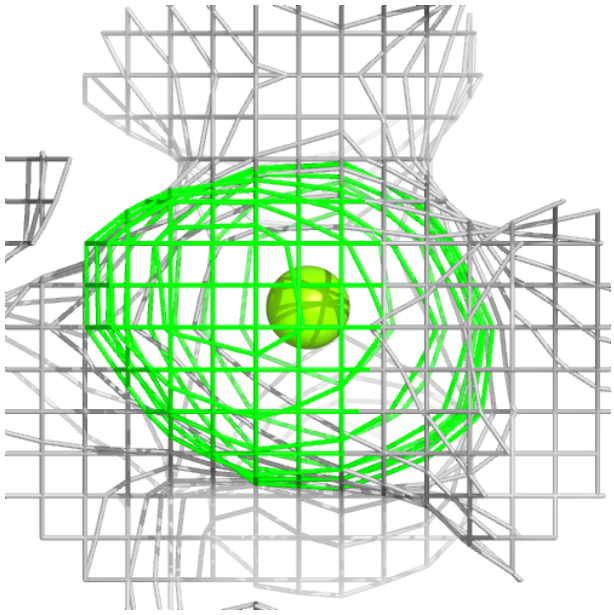
Electron density around MG B 902:

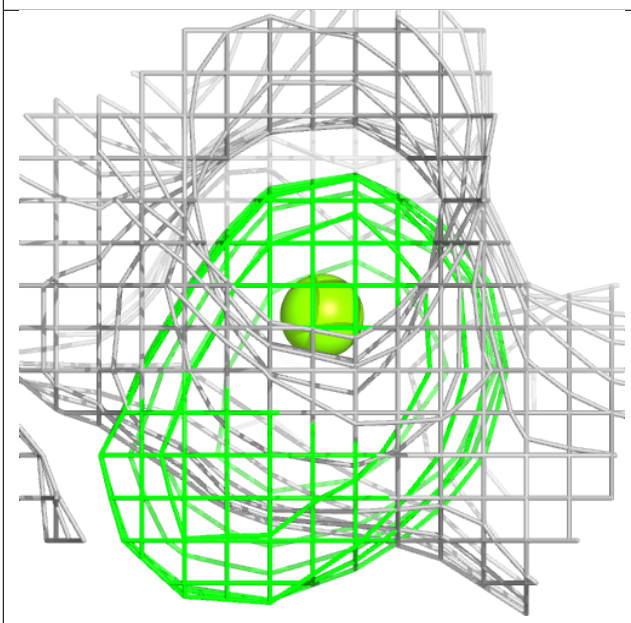
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

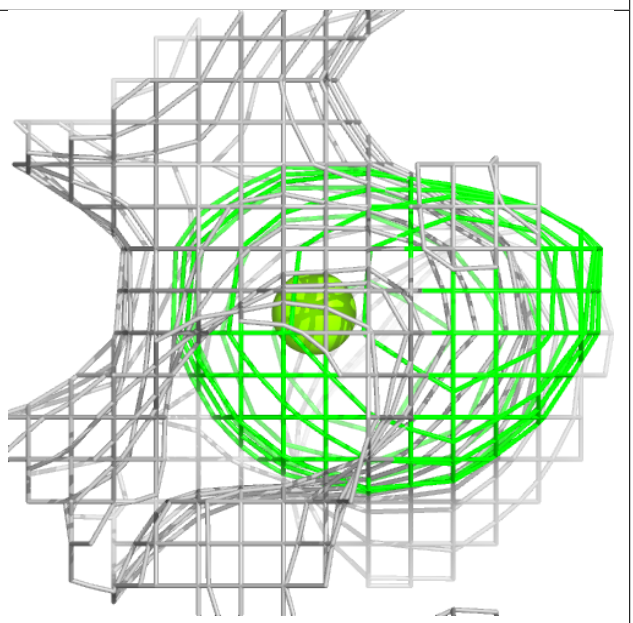


Electron density around MG B 903:

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

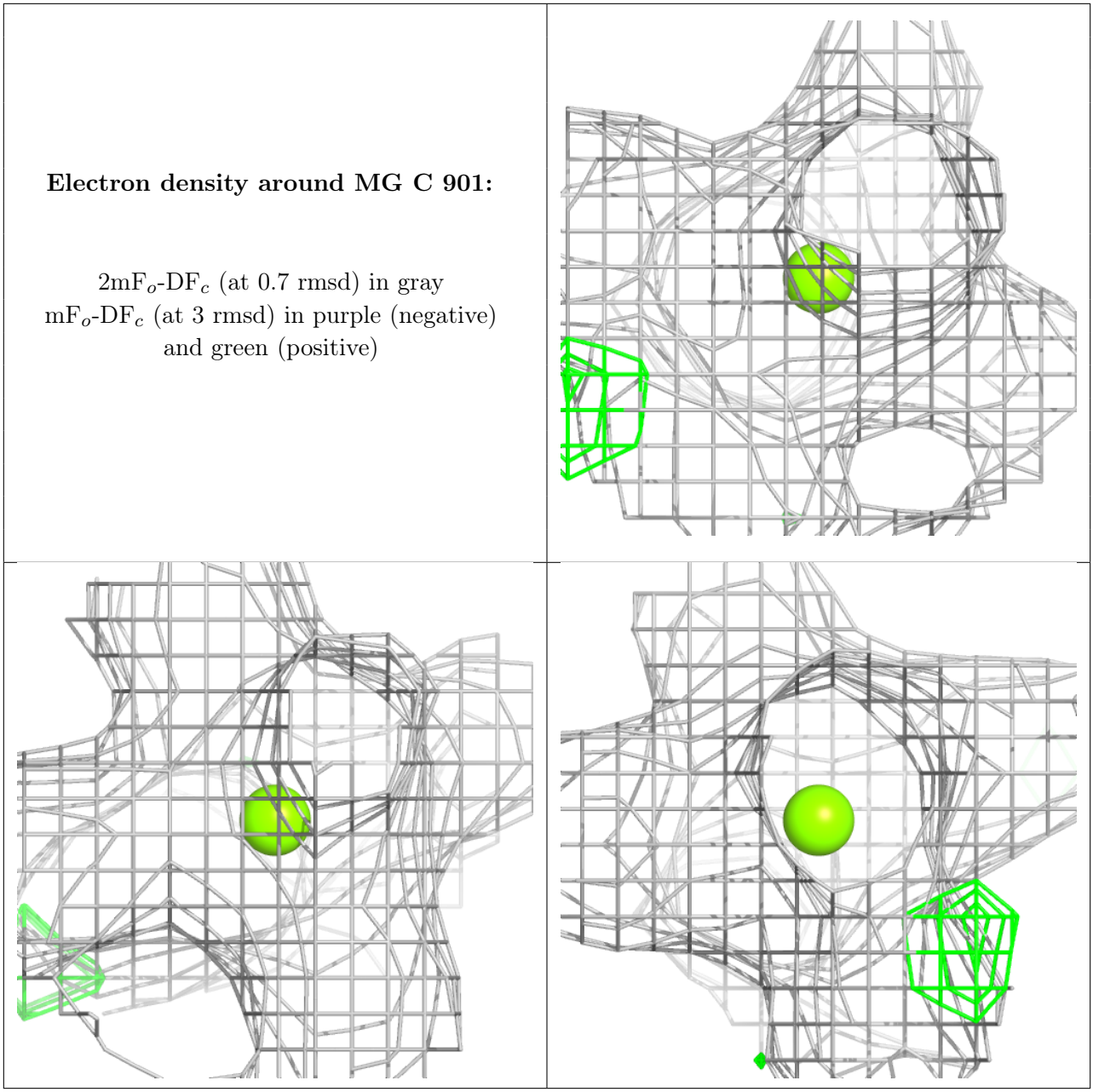






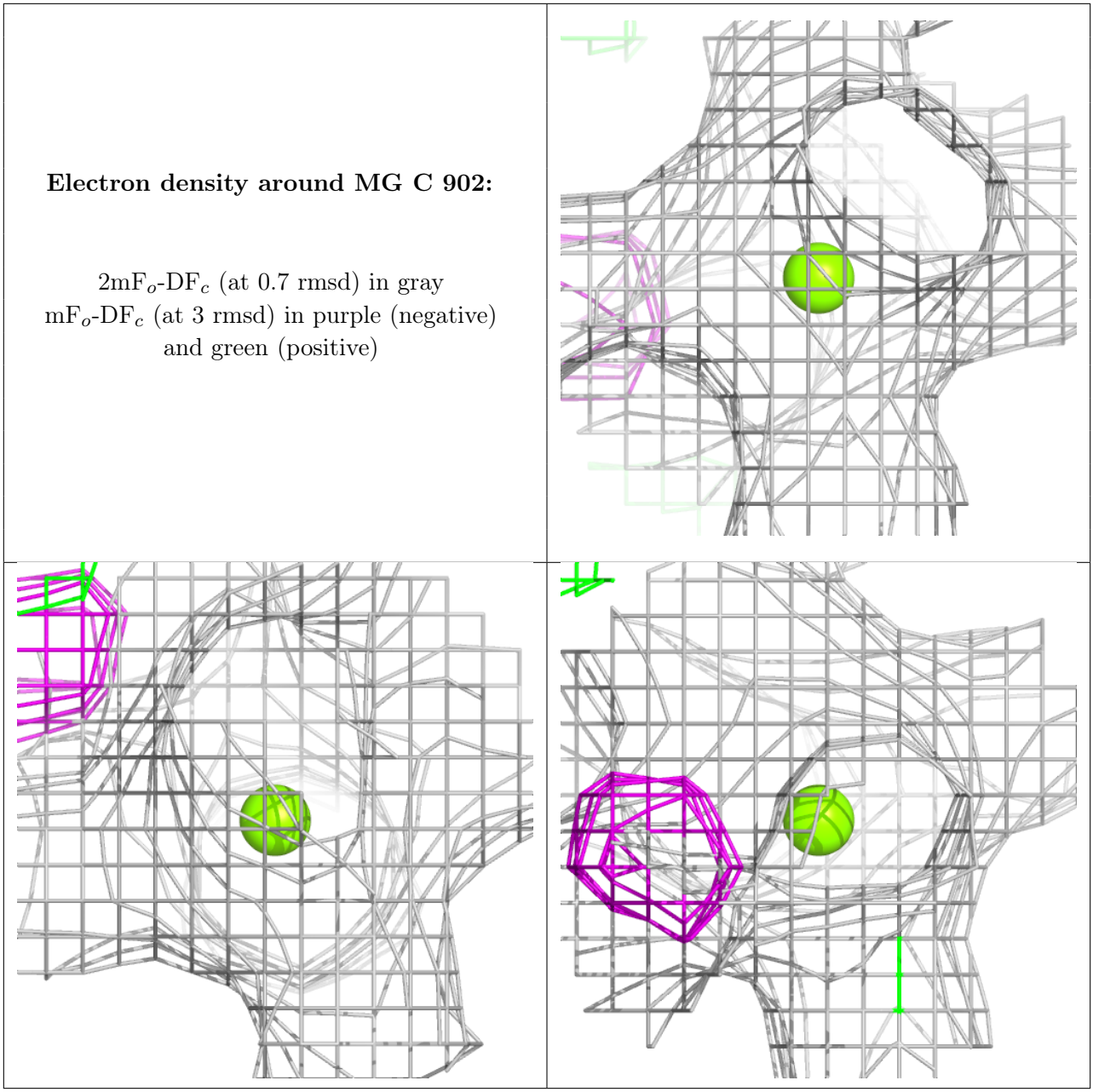
Electron density around MG C 901:

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



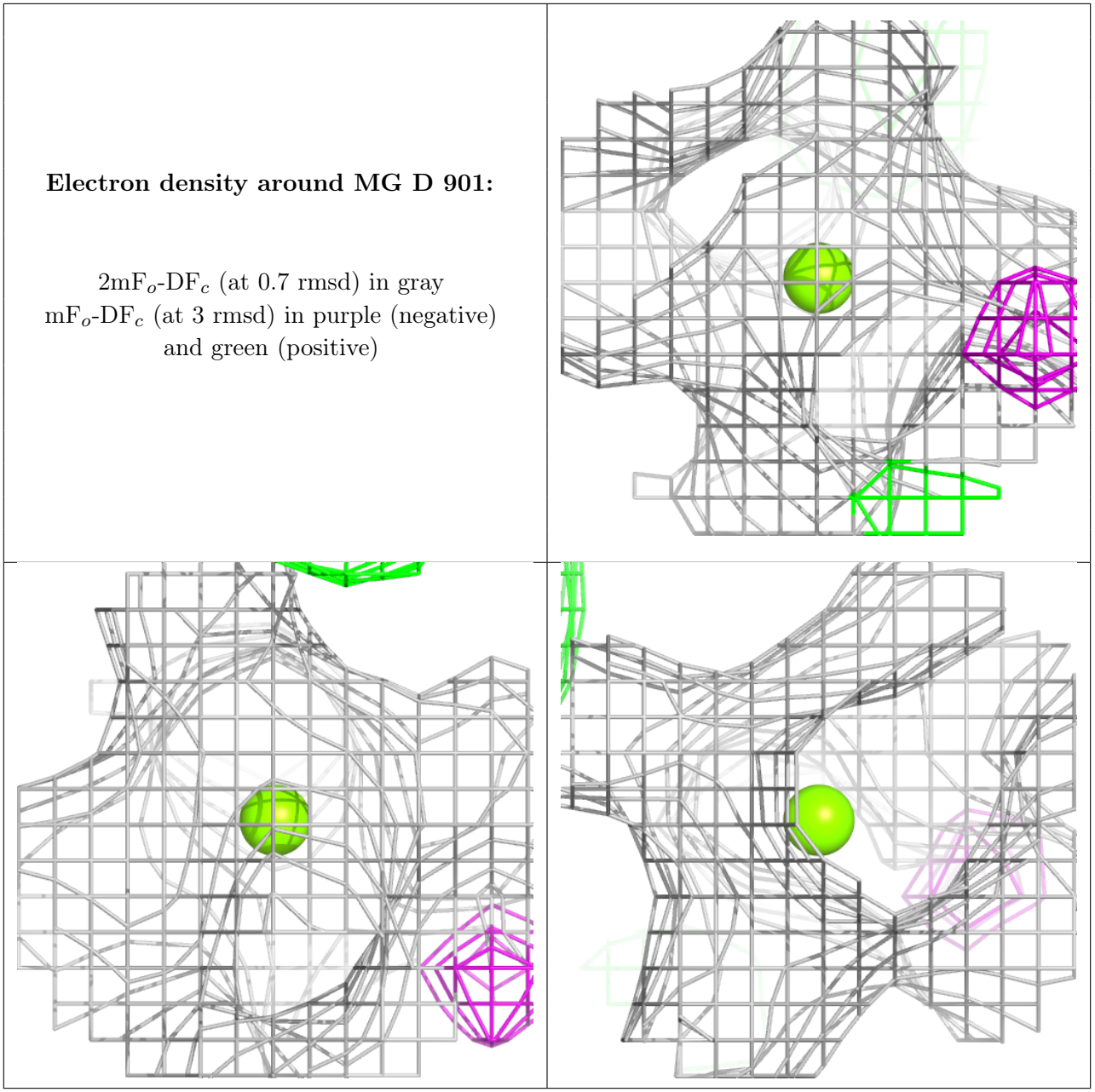
Electron density around MG C 902:

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



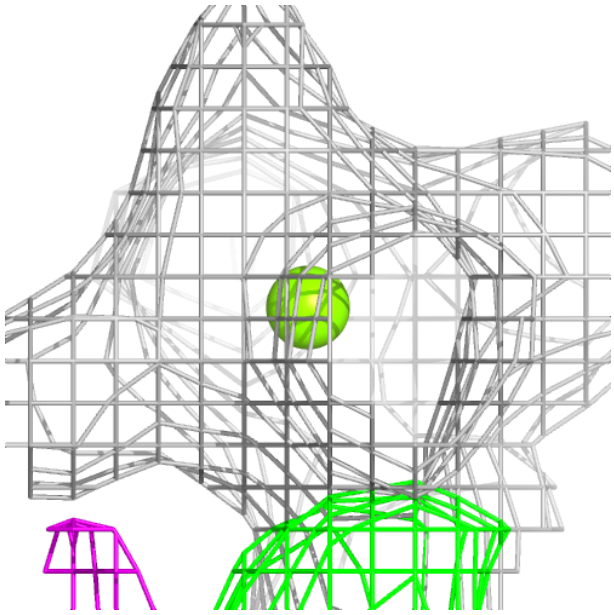
Electron density around MG D 901:

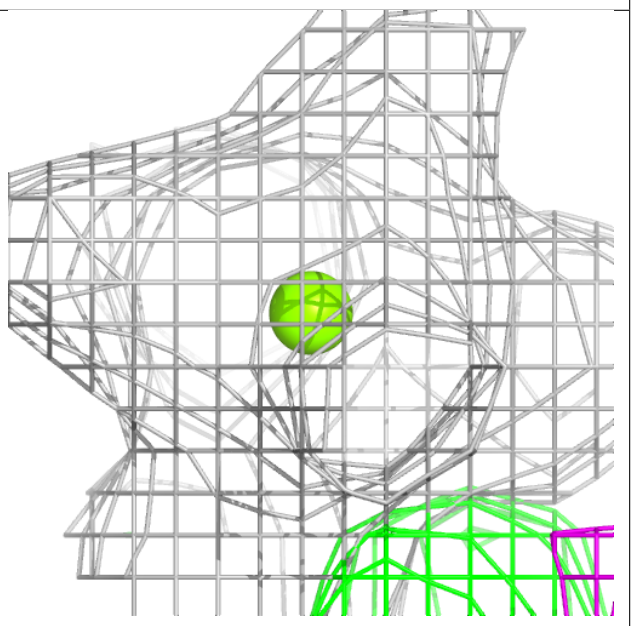
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around MG D 902:

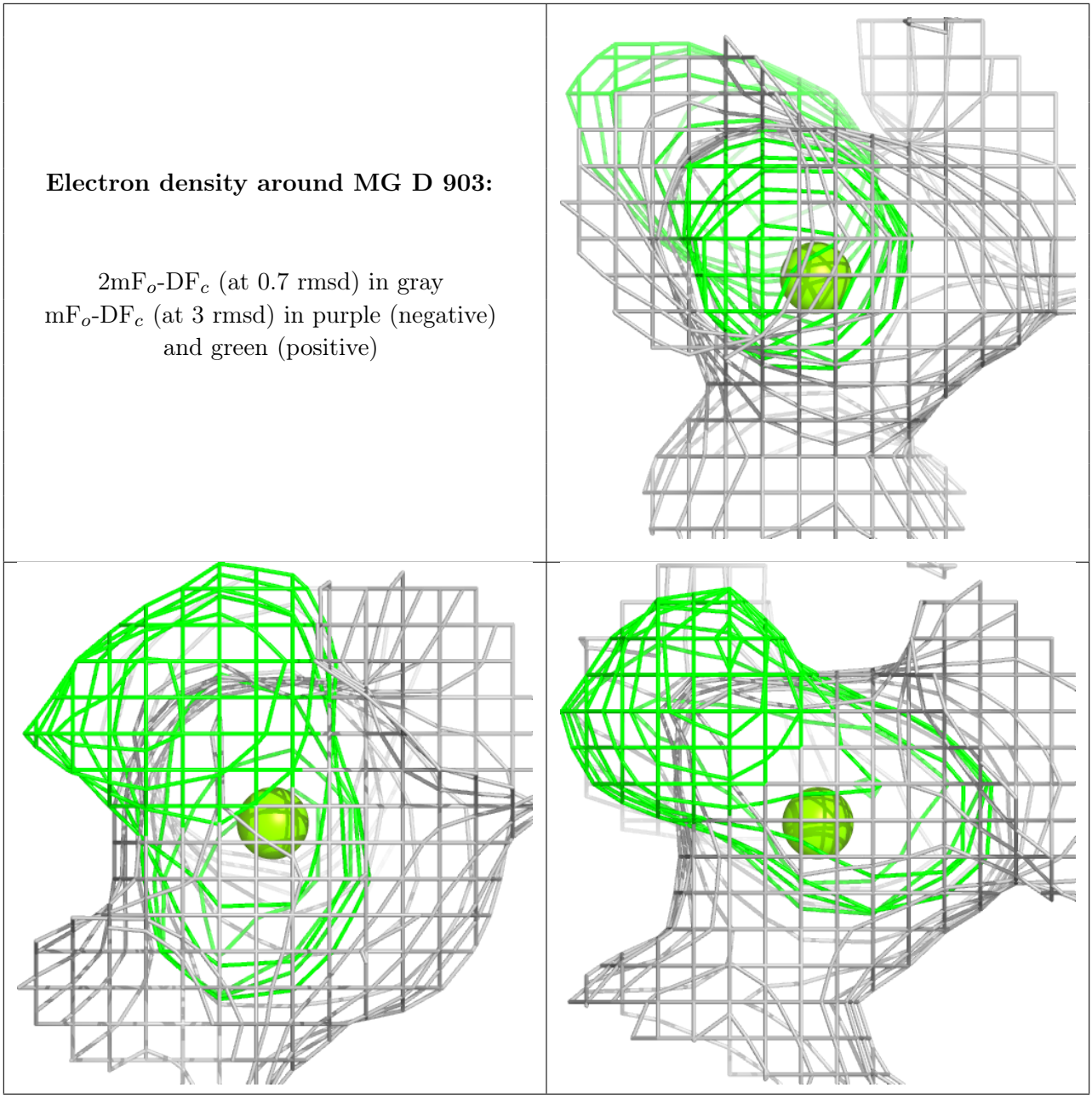
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





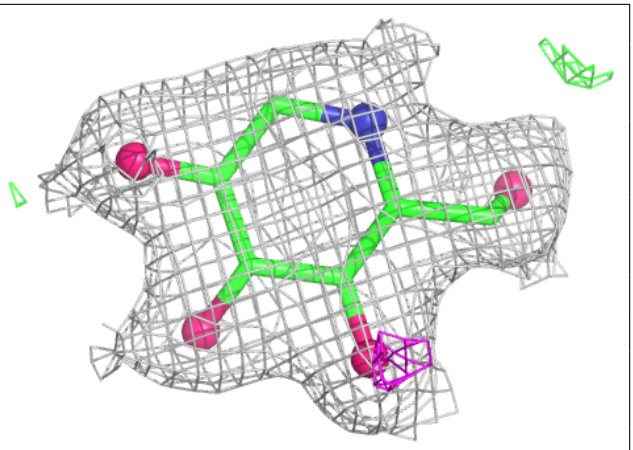
Electron density around MG D 903:

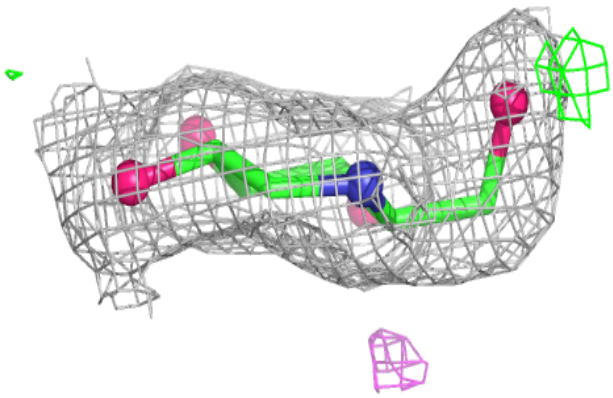
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

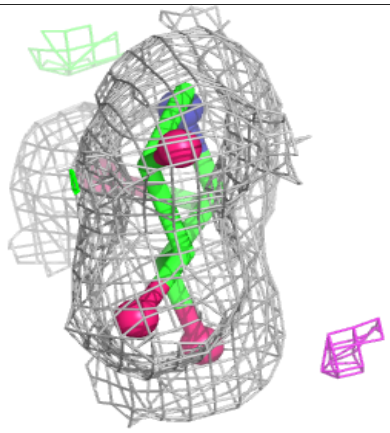


Electron density around NOJ A 903:

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

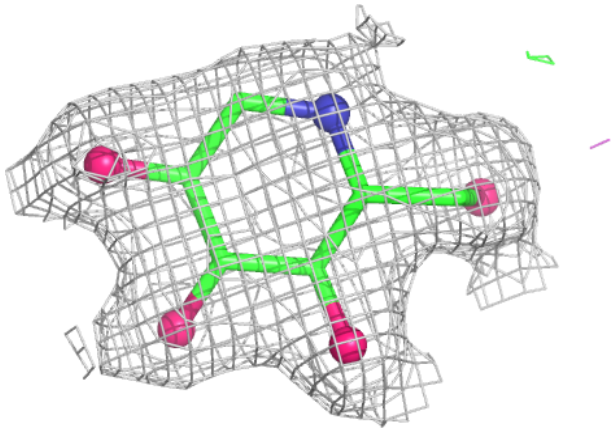


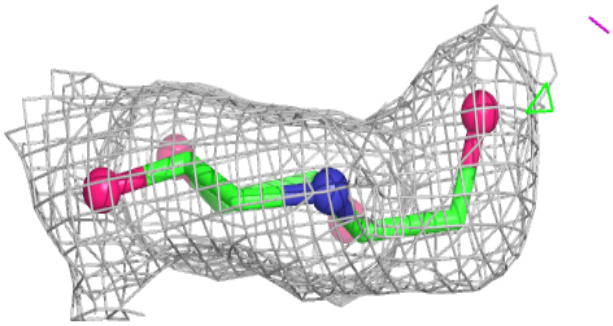


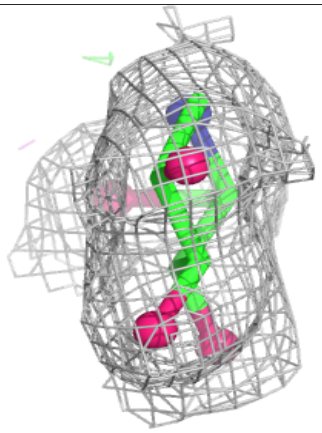


Electron density around NOJ B 904:

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

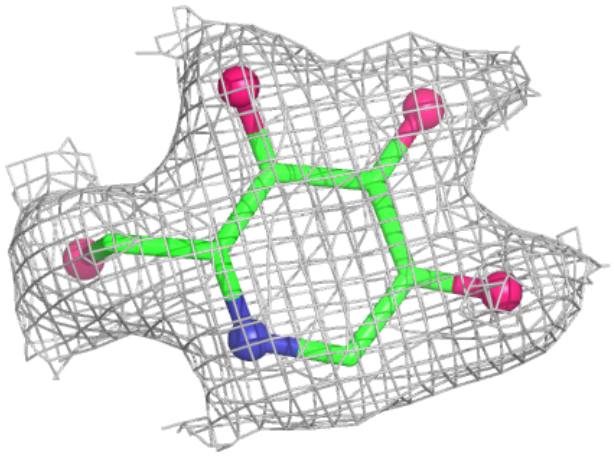


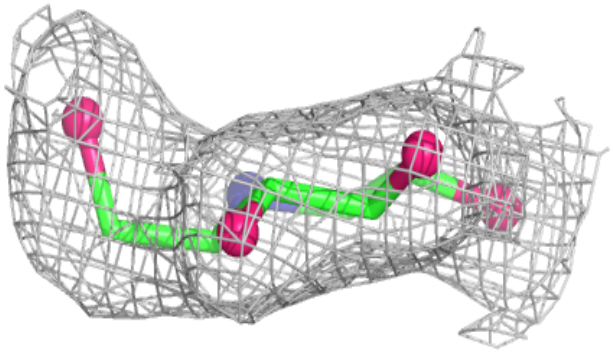


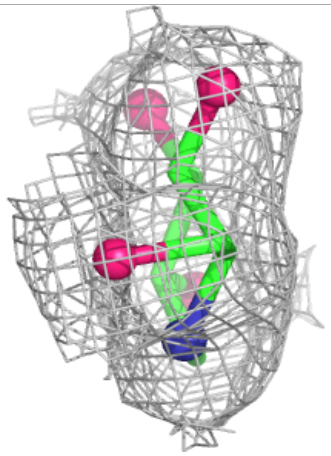


Electron density around NOJ C 903:

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

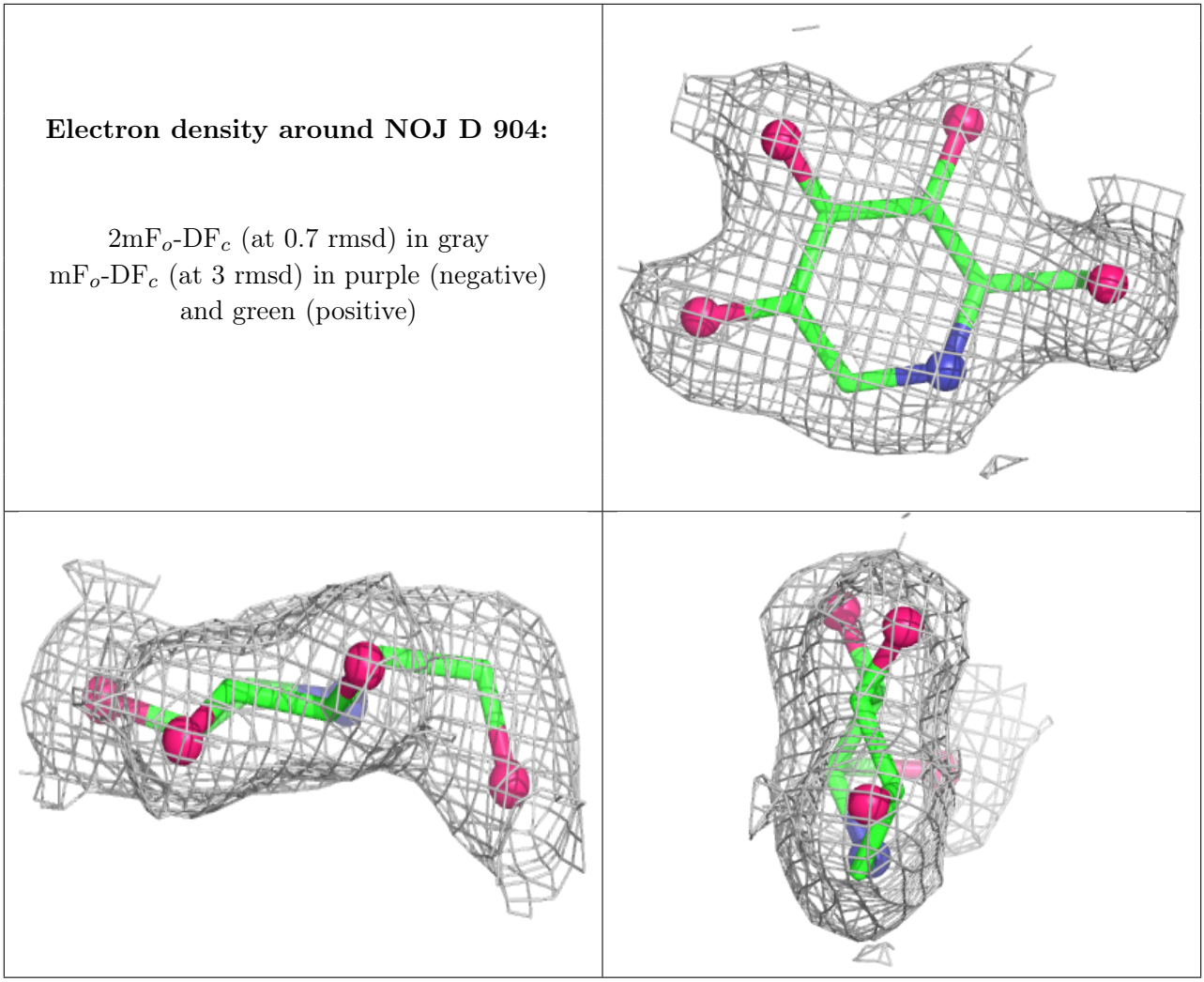






Electron density around NOJ D 904:

$2mF_o$-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

EDS was not executed - this section is therefore empty.