Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 11:00 PM EDT

PDB ID	:	7N02
Title	:	X-ray crystallographic structure model of Lactococcus lactis prolidase mutant D36S
Authors	:	Xu, S.; Grochulski, P.; Tanaka, T.
Deposited on	:	2021-05-24
Resolution	:	2.35 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 ⓘ 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 ⓘ) were used in the production of this report:

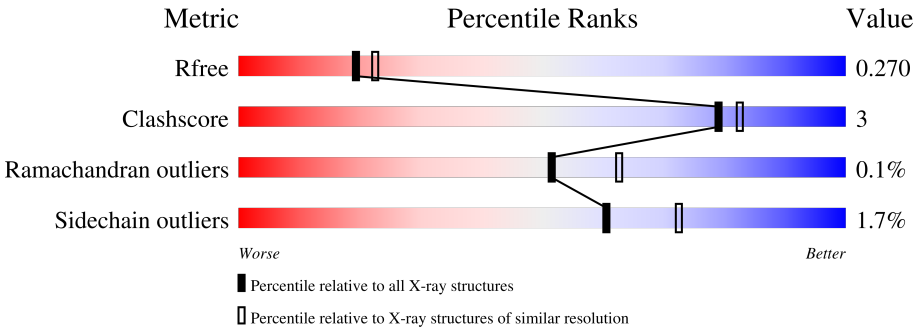
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)

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%

Mol	Chain	Length	Quality of chain
1	A	362	94% 6%
1	B	362	92% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5790 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 Aminopeptidase P family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2804	1776	466	545	17			
1	B	362	Total	C	N	O	S	0	0	0
			2804	1776	466	545	17			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	36	SER	ASP	engineered mutation	UNP A8WBX8
B	36	SER	ASP	engineered mutation	UNP A8WBX8

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mn	0	0
			2	2		
2	B	2	Total	Mn	0	0
			2	2		

- Molecule 3 is water.

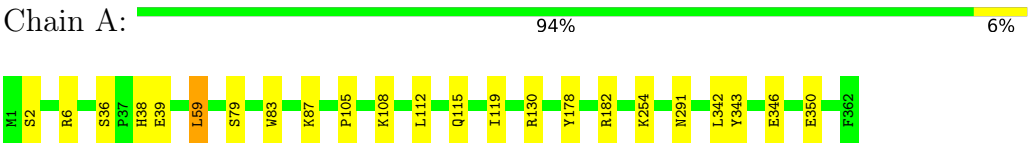
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	90	Total	O	0	0
			90	90		
3	B	88	Total	O	0	0
			88	88		

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.

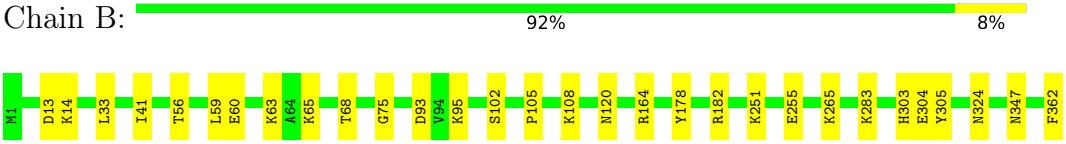
• Molecule 1: Aminopeptidase P family protein

Chain A:



• Molecule 1: Aminopeptidase P family protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	83.12Å 87.27Å 119.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.63 – 2.35 43.64 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.3 (43.63-2.35) 63.5 (43.64-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.07 (at 1.89Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.212 , 0.270 0.212 , 0.270	Depositor DCC
R_{free} test set	3232 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	30.3	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 43.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5790	wwPDB-VP
Average B, all atoms (Å^2)	57.0	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 27.52 % 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 2.1679e-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: MN

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	0/2859	0.61	2/3867 (0.1%)
1	B	0.46	0/2859	0.61	1/3867 (0.0%)
All	All	0.46	0/5718	0.61	3/7734 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	130	ARG	NE-CZ-NH2	-9.85	115.38	120.30
1	B	59	LEU	CA-CB-CG	8.28	134.33	115.30
1	A	130	ARG	NE-CZ-NH1	5.53	123.07	120.30

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	2804	0	2757	12	0
1	B	2804	0	2757	19	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	90	0	0	1	0
3	B	88	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5790	0	5514	31	1

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 3.

All (31) 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:93:ASP:HB3	1:B:95:LYS:NZ	2.06	0.70
1:B:251:LYS:O	1:B:255:GLU:HG3	1.99	0.62
1:B:120:ASN:OD1	3:B:501:HOH:O	2.16	0.61
1:B:105:PRO:HD2	1:B:108:LYS:HD2	1.85	0.57
1:A:36:SER:HB2	1:A:38:HIS:CD2	2.43	0.54
1:A:59:LEU:H	1:A:59:LEU:HD23	1.73	0.54
1:B:93:ASP:HB3	1:B:95:LYS:HZ2	1.73	0.54
1:B:93:ASP:HB3	1:B:95:LYS:HZ1	1.75	0.52
1:B:41:ILE:HD12	1:B:108:LYS:HE2	1.93	0.50
1:B:63:LYS:H	1:B:63:LYS:HD2	1.78	0.49
1:B:93:ASP:OD2	1:B:95:LYS:CE	2.63	0.47
1:B:65:LYS:NZ	1:B:68:THR:O	2.28	0.46
1:B:178:TYR:CZ	1:B:182:ARG:HD2	2.51	0.46
1:A:254:LYS:NZ	3:A:503:HOH:O	2.37	0.46
1:B:95:LYS:N	1:B:95:LYS:HD2	2.31	0.46
1:A:119:ILE:HD13	1:A:119:ILE:HA	1.82	0.45
1:A:87:LYS:HG3	1:A:115:GLN:HB3	2.00	0.43
1:B:56:THR:O	1:B:75:GLY:HA2	2.19	0.43
1:B:13:ASP:OD2	1:B:14:LYS:HE2	2.18	0.43
1:B:283:LYS:HB3	1:B:283:LYS:HE3	1.81	0.43
1:A:36:SER:HB2	1:A:38:HIS:NE2	2.34	0.43
1:A:87:LYS:O	1:A:87:LYS:HD3	2.19	0.43
1:B:265:LYS:HB3	1:B:265:LYS:HE2	1.90	0.43
1:A:343:TYR:CZ	1:A:350:GLU:HB2	2.55	0.42
1:B:60:GLU:HA	1:B:63:LYS:HD3	2.01	0.42
1:A:105:PRO:HD2	1:A:108:LYS:HD2	2.02	0.41
1:A:83:TRP:CZ3	1:A:112:LEU:HD13	2.55	0.41
1:A:178:TYR:CZ	1:A:182:ARG:HD2	2.54	0.41
1:A:291:ASN:OD1	1:A:291:ASN:N	2.53	0.41
1:B:304:GLU:HB2	1:B:305:TYR:H	1.74	0.41
1:B:164:ARG:NH1	1:B:362:PHE:O	2.51	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:95:LYS:NZ	1:B:347:ASN:ND2[4_455]	2.00	0.20

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	360/362 (99%)	347 (96%)	12 (3%)	1 (0%)	41 47
1	B	360/362 (99%)	346 (96%)	14 (4%)	0	100 100
All	All	720/724 (99%)	693 (96%)	26 (4%)	1 (0%)	51 63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	SER

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	303/303 (100%)	297 (98%)	6 (2%)	55 66
1	B	303/303 (100%)	299 (99%)	4 (1%)	69 80
All	All	606/606 (100%)	596 (98%)	10 (2%)	60 72

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	ARG
1	A	39	GLU
1	A	59	LEU
1	A	79	SER
1	A	342	LEU
1	A	346	GLU
1	B	33	LEU
1	B	102	SER
1	B	303	HIS
1	B	324	ASN

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

Mol	Chain	Res	Type
1	A	67	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

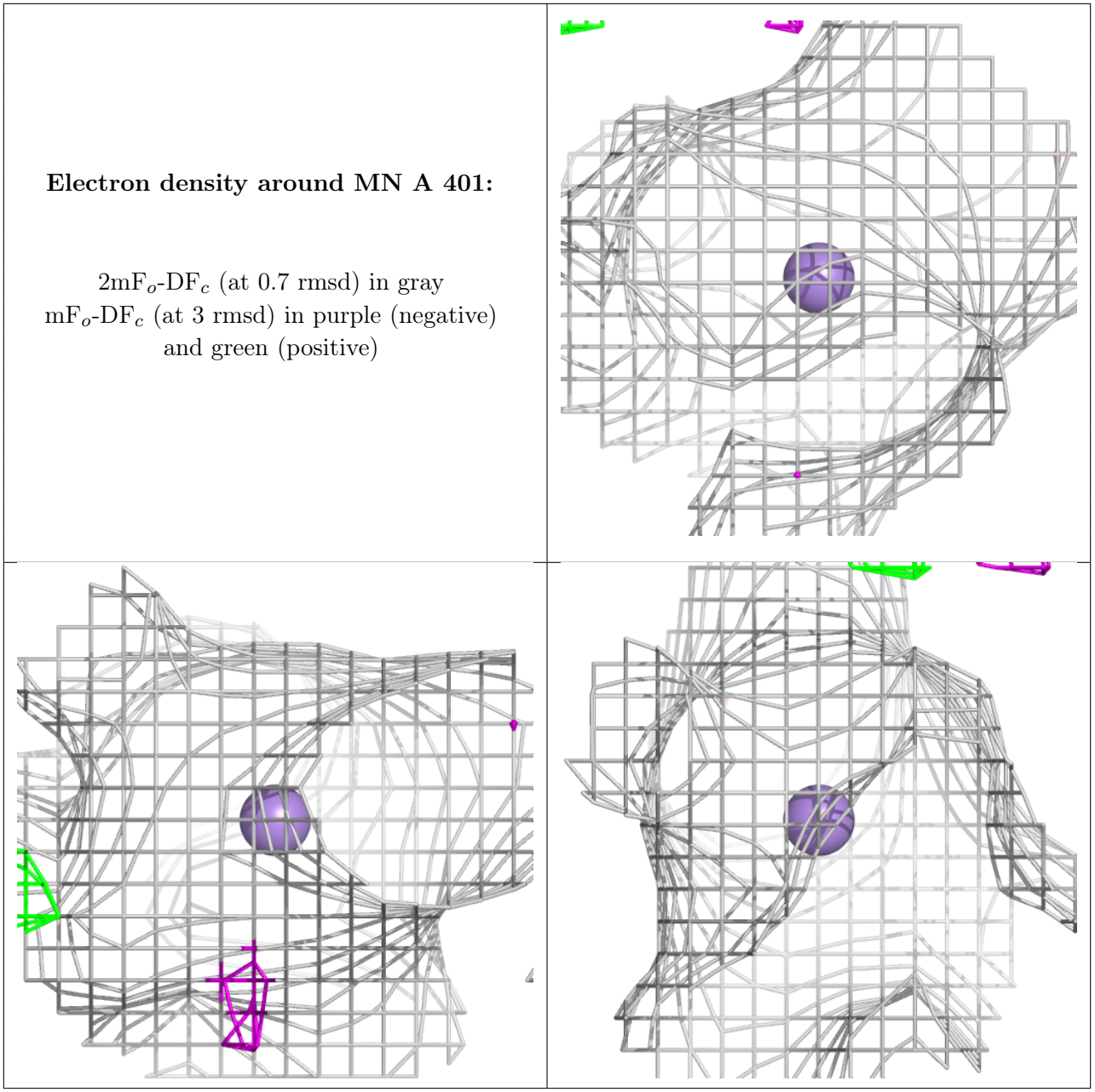
6.4 Ligands

Unable to reproduce the depositors R factor - 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.

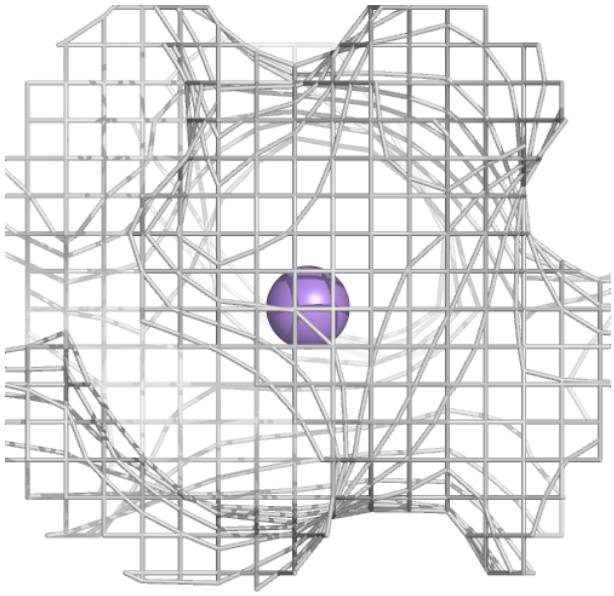
Electron density around MN A 401:

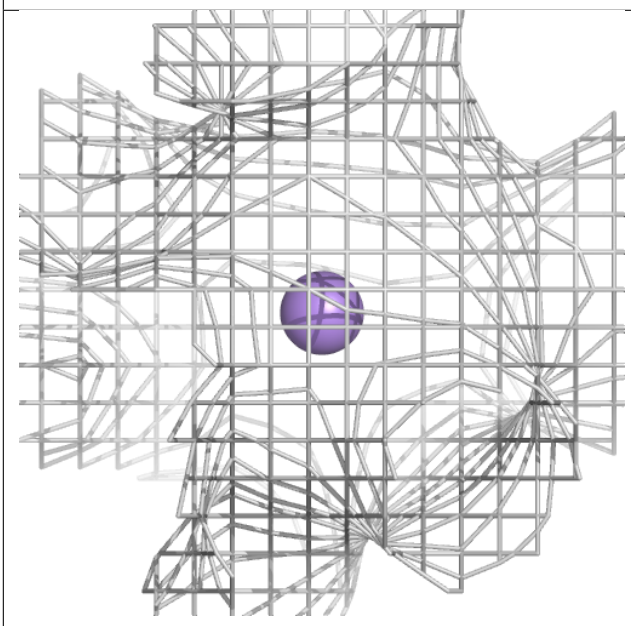
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

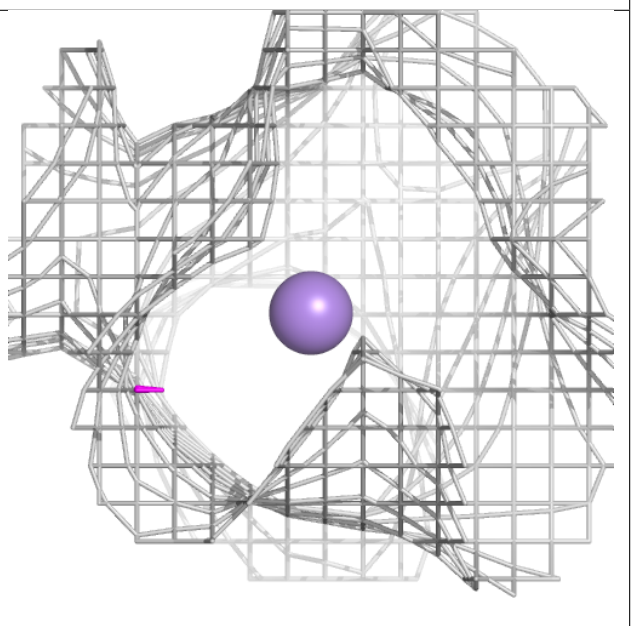


Electron density around MN A 402:

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

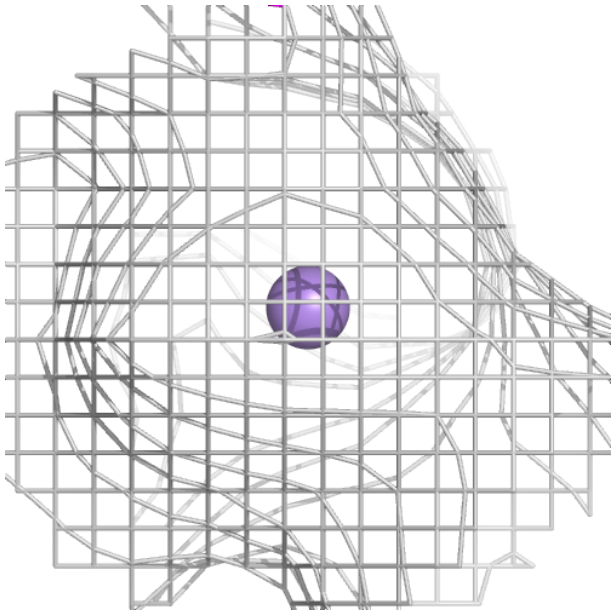


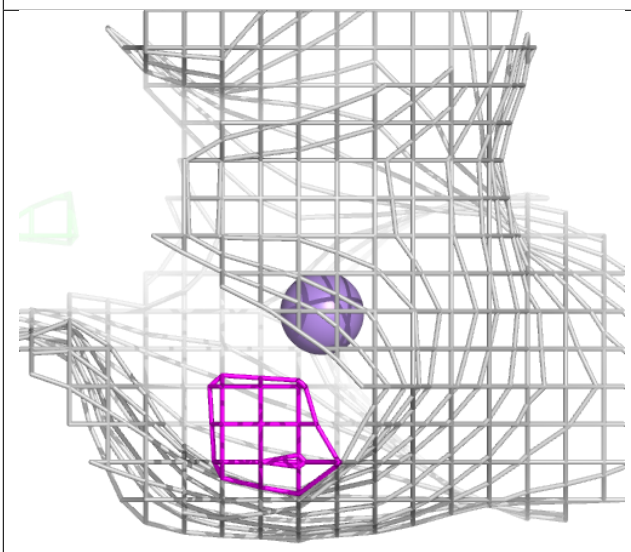


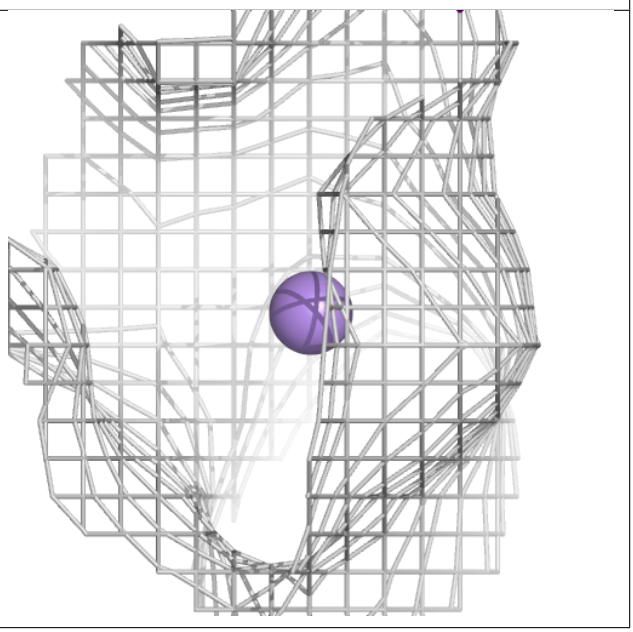


Electron density around MN B 401:

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

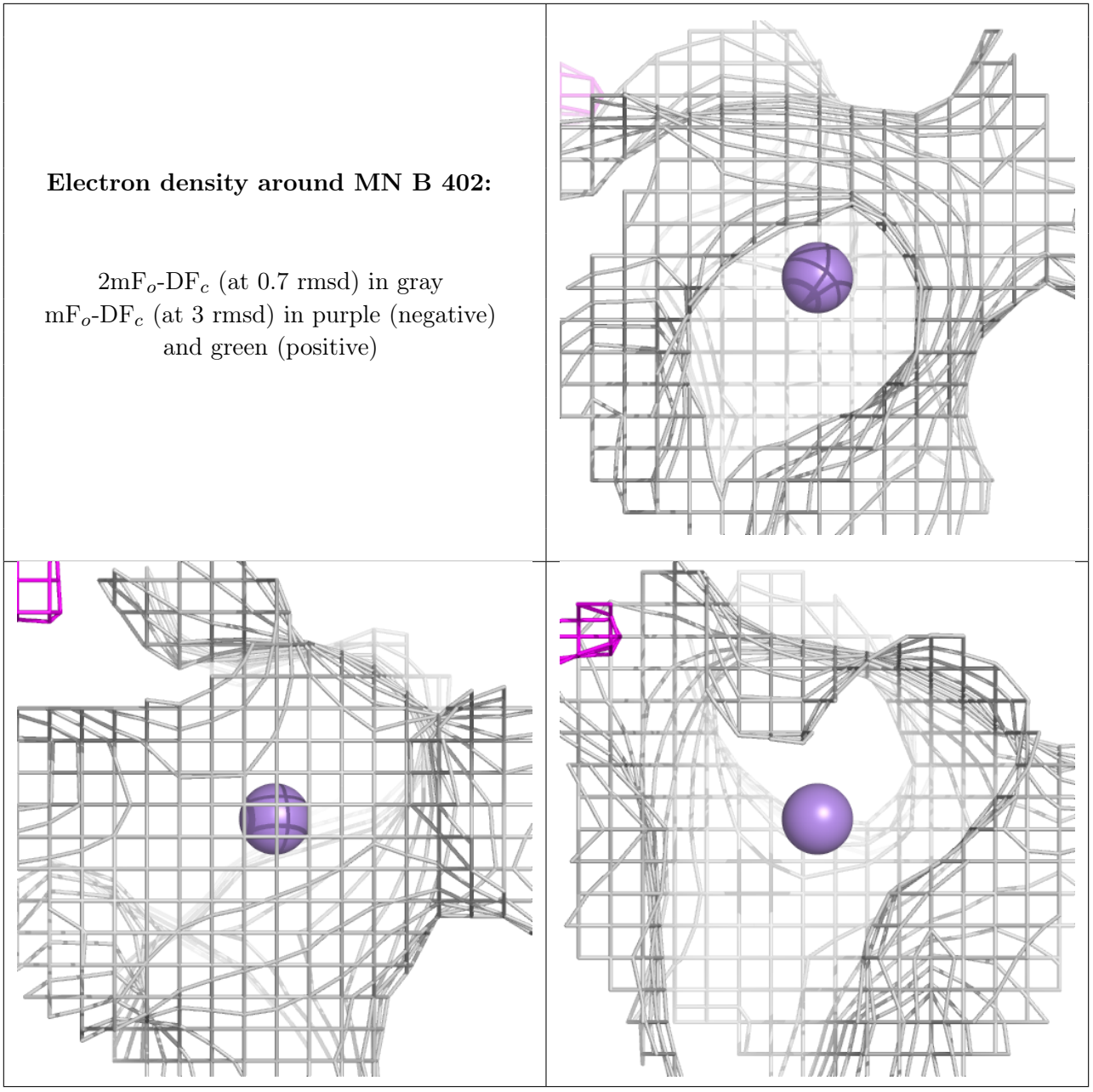






Electron density around MN B 402:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.