



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:28 AM UTC

PDB ID : 6VDE / pdb_00006vde
Title : Full-length M. smegmatis Pol1
Authors : Shuman, S.; Goldgur, Y.; Ghosh, S.
Deposited on : 2019-12-24
Resolution : 2.71 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 ⓘ 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 : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

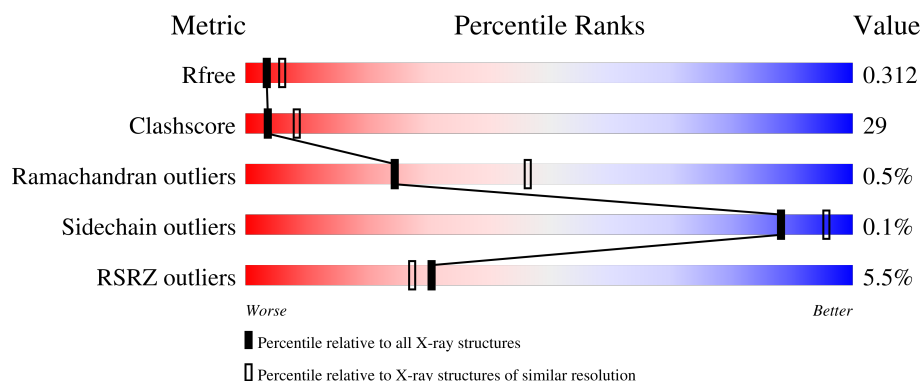
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.71 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}	180053	4348 (2.74-2.70)
Clashscore	190562	4665 (2.74-2.70)
Ramachandran outliers	187476	4584 (2.74-2.70)
Sidechain outliers	187428	4585 (2.74-2.70)
RSRZ outliers	180081	4348 (2.74-2.70)

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 $\leq 5\%$. 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	908	5% 55% 38% •• 5%
1	B	908	6% 50% 42% • 5%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13366 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 DNA polymerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	860	Total	C	N	O	S	0	0	0
			6672	4181	1171	1304	16			
1	B	860	Total	C	N	O	S	0	0	0
			6672	4181	1171	1304	16			

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

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

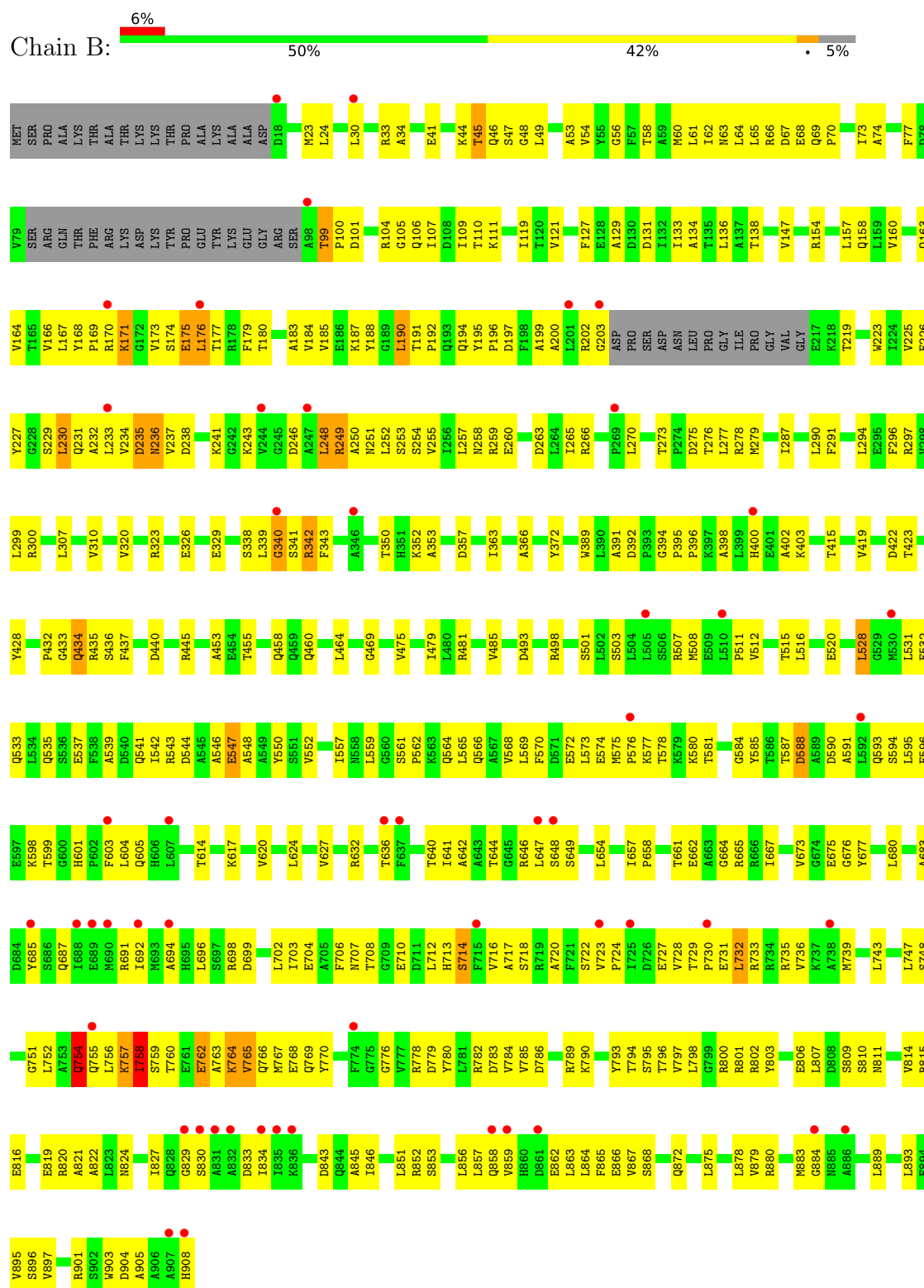
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	6	Total	O	0	0
			6	6		
3	B	10	Total	O	0	0
			10	10		

- Molecule 1: DNA polymerase I



• Molecule 1: DNA polymerase I



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	70.00Å 80.94Å 117.91Å 98.68° 105.13° 109.36°	Depositor
Resolution (Å)	50.37 – 2.71 50.37 – 2.71	Depositor EDS
% Data completeness (in resolution range)	97.4 (50.37-2.71) 90.5 (50.37-2.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 2.73Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.234 , 0.304 0.241 , 0.312	Depositor DCC
R_{free} test set	2000 reflections (3.33%)	wwPDB-VP
Wilson B-factor (Å ²)	80.0	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 63.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	13366	wwPDB-VP
Average B, all atoms (Å ²)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.52	1/6783 (0.0%)	0.96	25/9208 (0.3%)
1	B	0.50	0/6783	0.95	27/9208 (0.3%)
All	All	0.51	1/13566 (0.0%)	0.96	52/18416 (0.3%)

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	13
1	B	0	13
All	All	0	26

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	598	LYS	CD-CE	-5.32	1.36	1.52

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	434	GLN	CA-CB-CG	8.47	131.05	114.10
1	B	806	GLU	CA-CB-CG	7.97	130.04	114.10
1	A	300	ARG	CG-CD-NE	7.50	128.50	112.00
1	A	598	LYS	CD-CE-NZ	-7.48	87.96	111.90
1	A	98	ALA	CA-C-N	7.47	132.89	122.06

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	233	LEU	Peptide
1	A	235	ASP	Peptide
1	A	248	LEU	Peptide
1	A	258	ASN	Peptide
1	A	45	THR	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	6672	0	6604	391	3
1	B	6672	0	6608	388	2
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	6	0	0	0	0
3	B	10	0	0	8	0
All	All	13366	0	13212	773	3

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 29.

The worst 5 of 773 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:A:731:GLU:OE2	1:A:735:ARG:NH2	1.59	1.33
1:A:288:HIS:CG	1:A:300:ARG:HH21	1.48	1.31
1:B:129:ALA:O	1:B:133:ILE:HD12	1.29	1.29
1:A:191:THR:HG23	1:A:194:GLN:OE1	1.18	1.26
1:A:288:HIS:HA	1:A:300:ARG:NH2	1.47	1.25

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:762:GLU:CD	1:B:598:LYS:NZ[1_454]	1.66	0.54
1:A:762:GLU:OE1	1:B:598:LYS:NZ[1_454]	1.67	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:292:ASP:OD2	1:A:908:HIS:NE2[1_545]	2.16	0.04

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	854/908 (94%)	764 (90%)	85 (10%)	5 (1%)	21	42
1	B	854/908 (94%)	763 (89%)	87 (10%)	4 (0%)	24	46
All	All	1708/1816 (94%)	1527 (89%)	172 (10%)	9 (0%)	24	46

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	ASN
1	A	249	ARG
1	A	342	ARG
1	B	236	ASN
1	B	249	ARG

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	703/742 (95%)	702 (100%)	1 (0%)	88	95
1	B	703/742 (95%)	702 (100%)	1 (0%)	88	95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1406/1484 (95%)	1404 (100%)	2 (0%)	88 95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	748	SER
1	B	714	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	236	ASN
1	B	476	GLN
1	B	882	HIS
1	B	458	GLN
1	A	593	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	860/908 (94%)	0.37	42 (4%) 35 32	62, 109, 171, 213	0
1	B	860/908 (94%)	0.41	52 (6%) 27 24	59, 116, 184, 215	0
All	All	1720/1816 (94%)	0.39	94 (5%) 30 27	59, 112, 180, 215	0

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	176	LEU	7.5
1	A	199	ALA	5.9
1	B	647	LEU	5.0
1	A	224	ILE	4.9
1	A	201	LEU	4.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

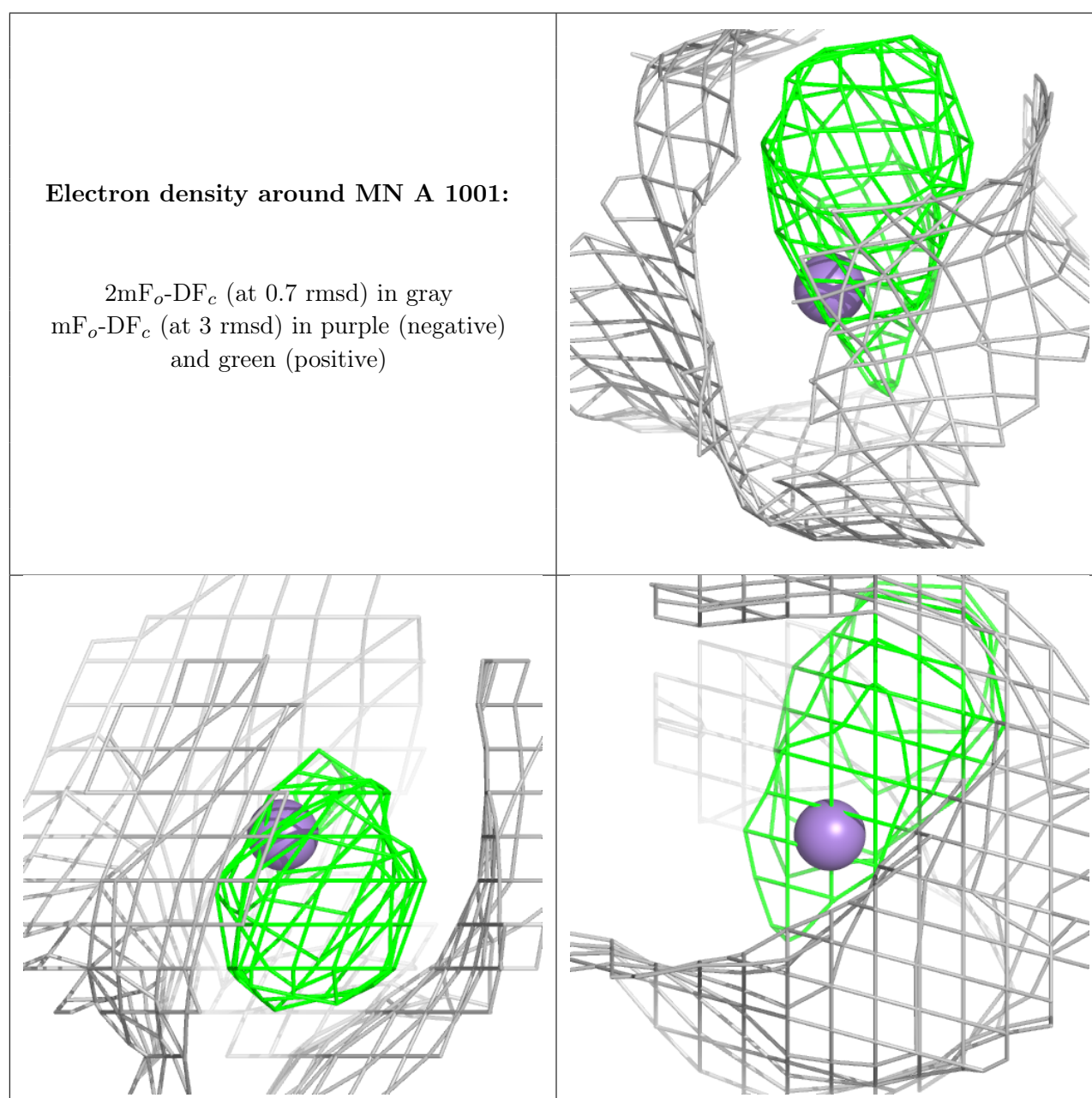
There are no oligosaccharides in this entry.

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, 95th 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.

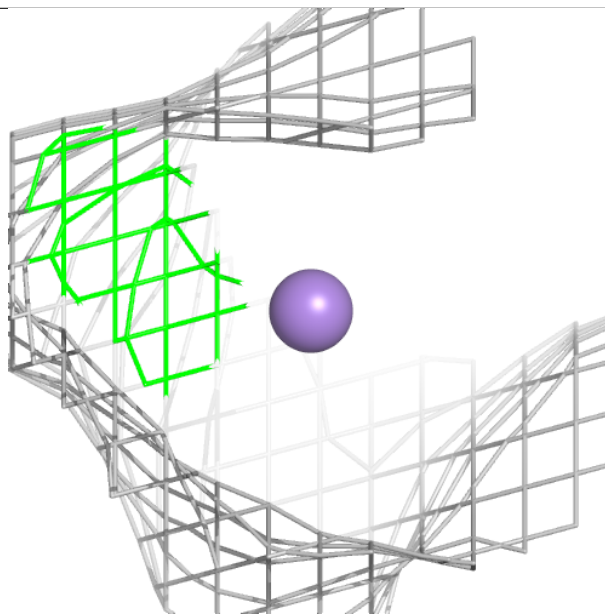
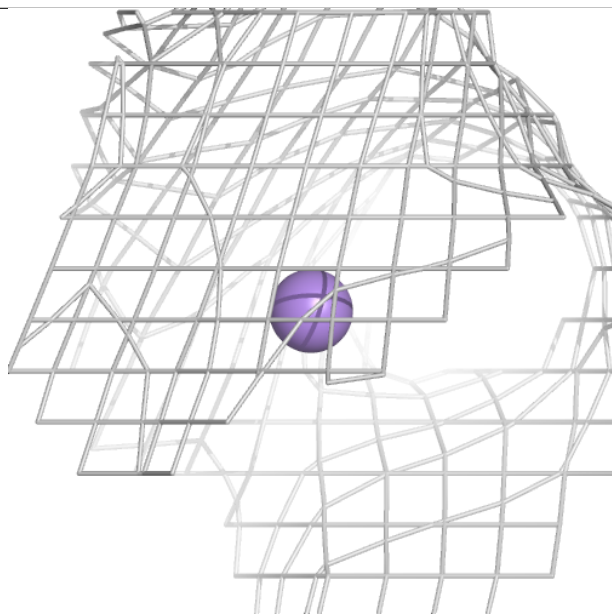
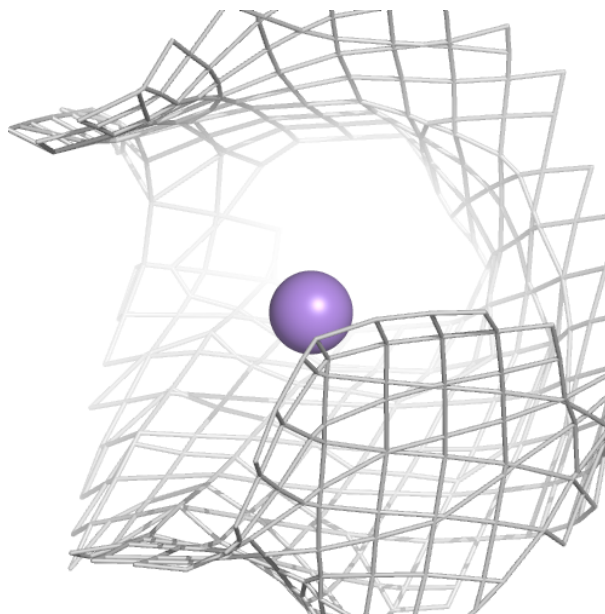
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	MN	A	1001	1/1	0.70	0.13	314,314,314,314	0
2	MN	A	1002	1/1	0.77	0.12	164,164,164,164	1
2	MN	B	1001	1/1	0.79	0.14	266,266,266,266	1
2	MN	B	1002	1/1	0.91	0.12	197,197,197,197	0
2	MN	B	1003	1/1	0.94	0.13	95,95,95,95	0
2	MN	A	1003	1/1	0.97	0.18	150,150,150,150	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.



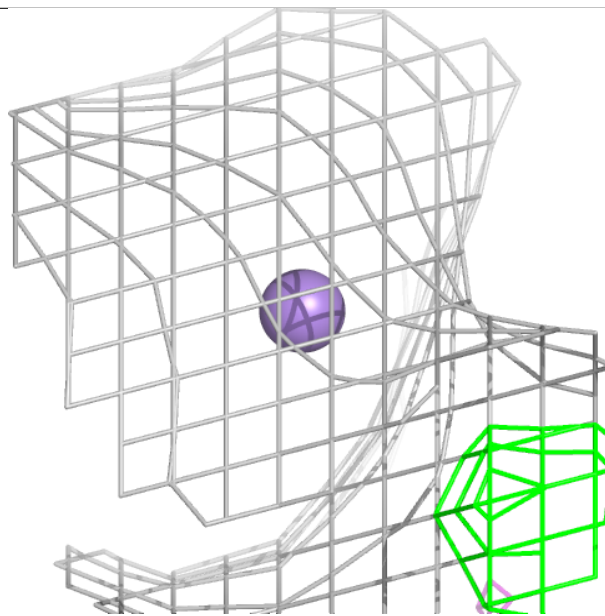
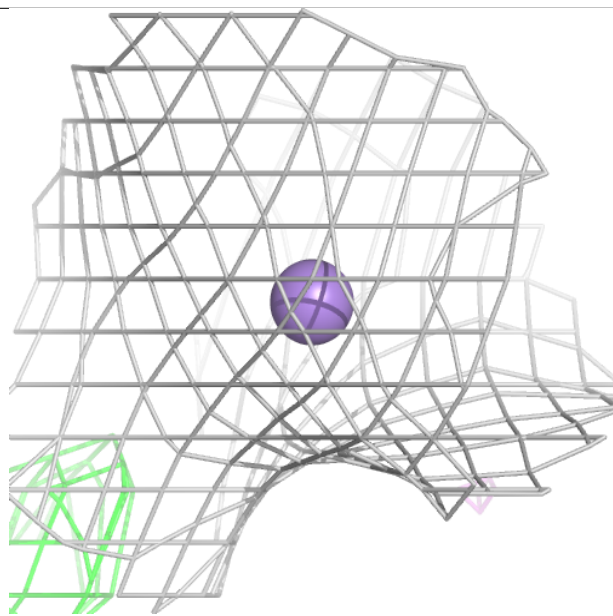
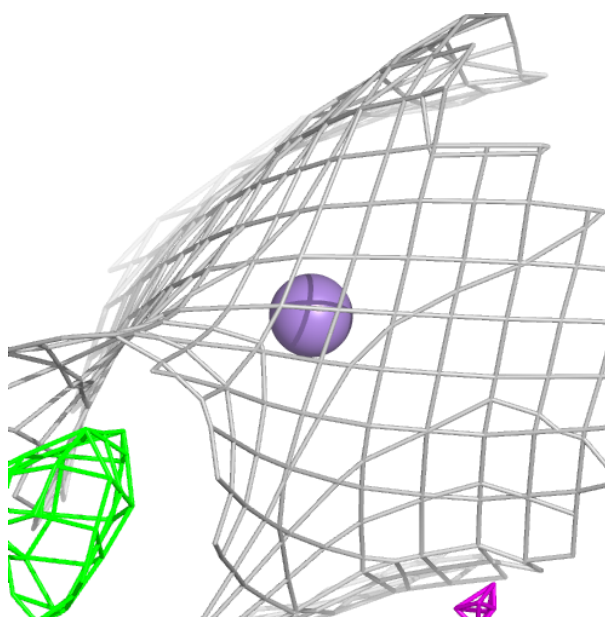
Electron density around MN A 1002:

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



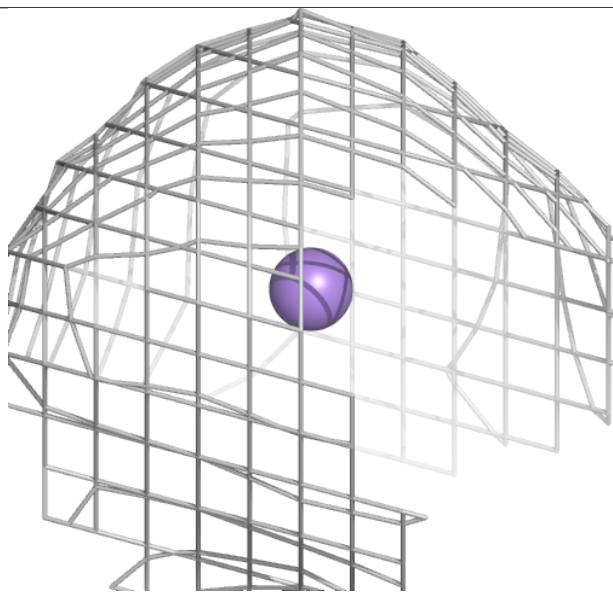
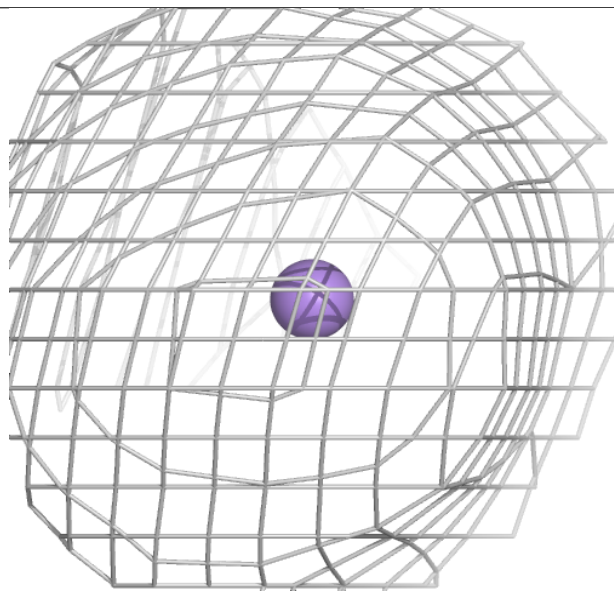
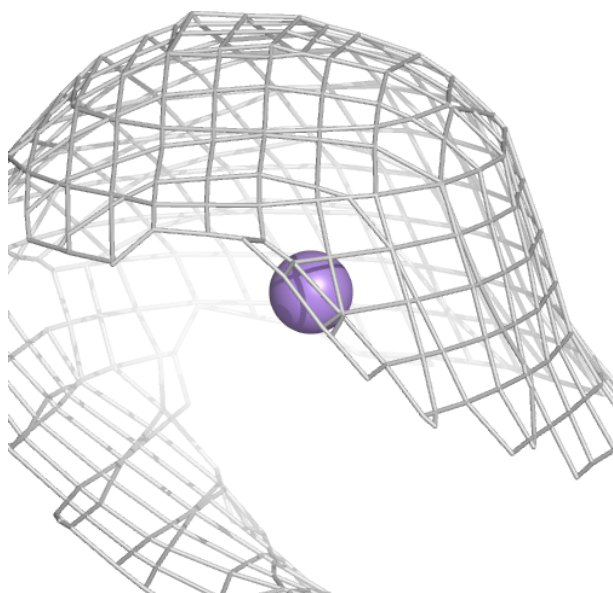
Electron density around MN B 1001:

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



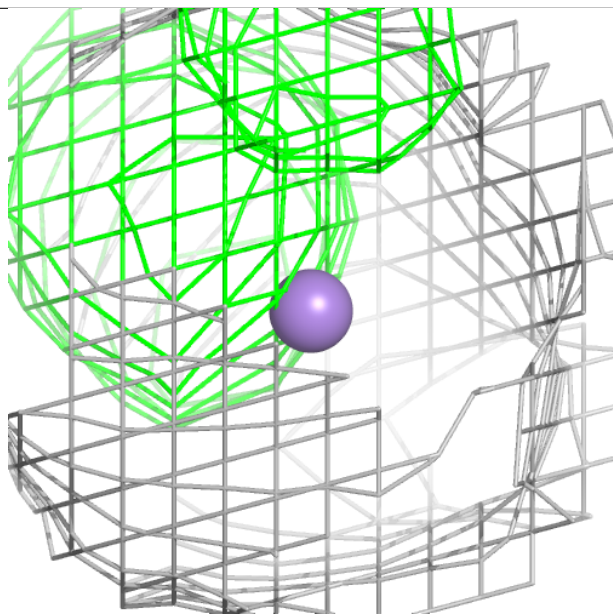
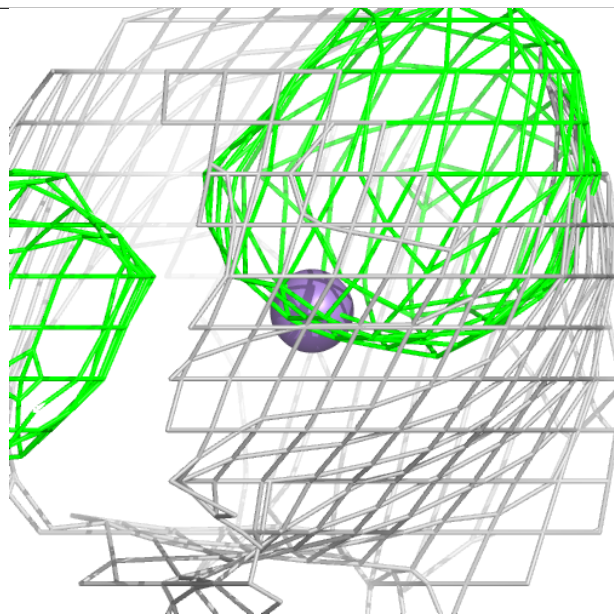
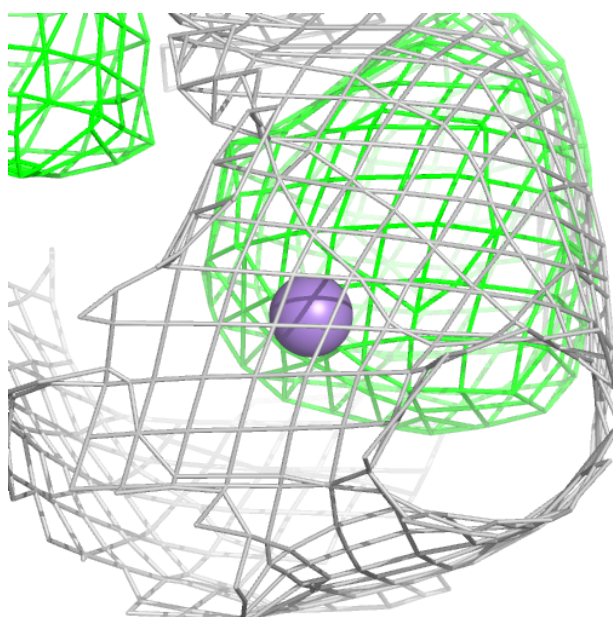
Electron density around MN B 1002:

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



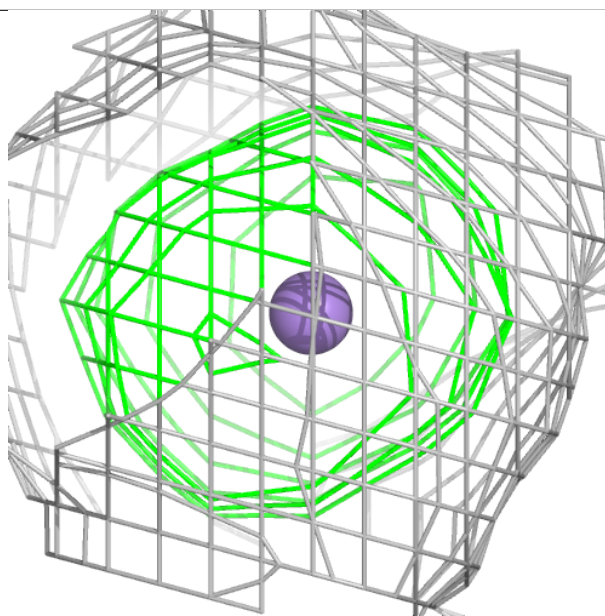
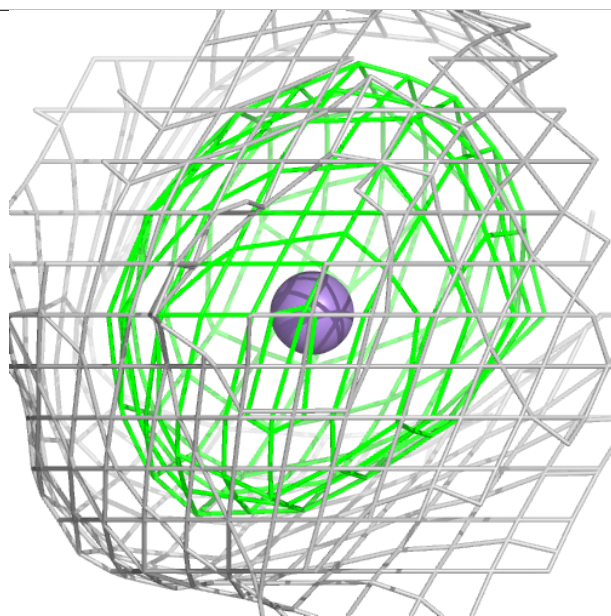
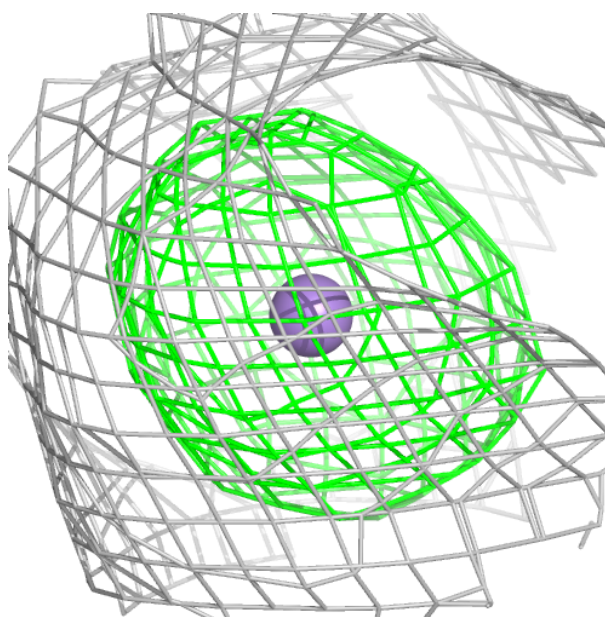
Electron density around MN B 1003:

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



Electron density around MN A 1003:

$2mF_o-DF_c$ (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.