W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Jul 3, 2024 – 10:46 AM EDT

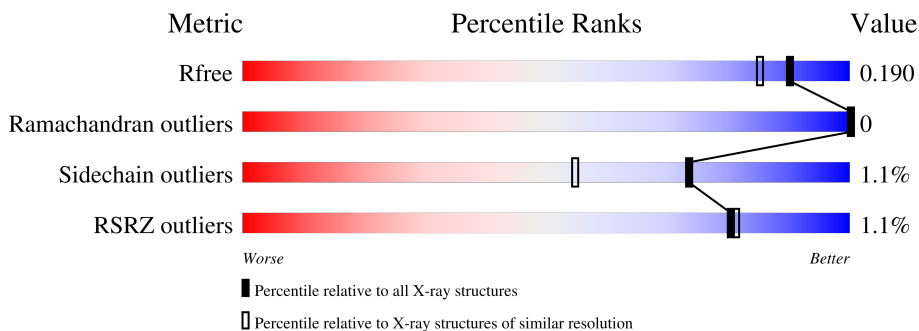|		
PDB ID	:	8VNL
Title	:	Homing endonuclease I-PpoI-DNA complex:reaction at pH6.0 (K+ MES) with 500 uM Mn2+ for 240s
Authors	:	Chang, C.; Gao, Y.
Deposited on	:	2024-01-13
Resolution	:	1.64 Å(reported)

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 1.64 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	Percentile Ranks	Value
Rfree		0.190
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		1.1%

Worse — Better

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3122 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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	C	13	92% 8%
1	D	13	100%
2	c	8	100%
2	d	8	100%
3	A	162	.% 100%
3	B	162	2% 98% .

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 3876 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 DNA chain called DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*TP*CP*TP* TP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	13	Total	C	N	O	P	0	1	0
			280	137	46	84	13			
1	D	13	Total	C	N	O	P	0	1	0
			280	137	46	84	13			

- Molecule 2 is a DNA chain called DNA (5'-D(P*GP*AP*GP*AP*GP*TP*CP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	c	8	Total	C	N	O	P	0	1	0
			190	89	40	52	9			
2	d	8	Total	C	N	O	P	0	1	0
			190	89	40	52	9			

- Molecule 3 is a protein called Intron-encoded endonuclease I-PpoI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	162	Total	C	N	O	S	0	1	0
			1249	789	232	220	8			
3	B	162	Total	C	N	O	S	0	1	0
			1249	789	232	220	8			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

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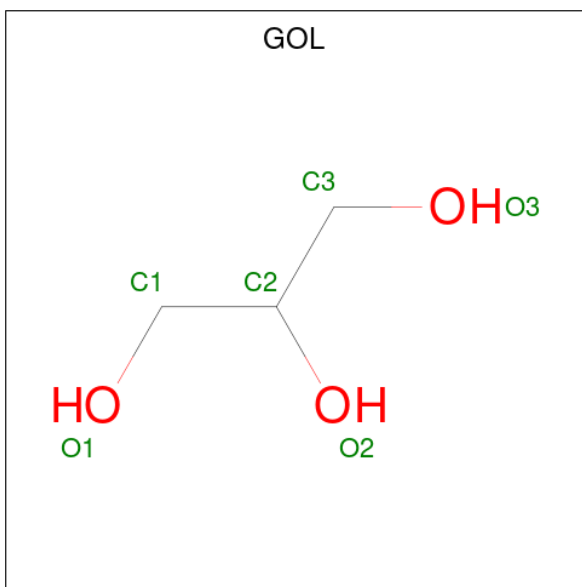
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total 6	C 3	O 3	0	0
4	D	1	Total 6	C 3	O 3	0	0
4	A	1	Total 6	C 3	O 3	0	0
4	A	1	Total 6	C 3	O 3	0	0
4	B	1	Total 6	C 3	O 3	0	0

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	c	1	Total 1	Mn 1	0	0
5	D	1	Total 1	Mn 1	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	c	1	Total 1	Na 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total 1	Na 1	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	30	Total 30	O 30	0	0
8	c	15	Total 15	O 15	0	0
8	D	36	Total 36	O 36	0	0
8	d	17	Total 17	O 17	0	0
8	A	151	Total 151	O 151	0	1
8	B	151	Total 151	O 151	0	1

3   Residue-property plots ⓘ

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: DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*TP*CP*TP*TP*AP*A)-3')

Chain C:

92%   8%

T401 T411 A412 A413

• Molecule 1: DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*TP*CP*TP*TP*AP*A)-3')

Chain D:

100%

There are no outlier residues recorded for this chain.

• Molecule 2: DNA (5'-D(P*GP*AP*GP*AP*GP*TP*CP*A)-3')

Chain c:

100%

There are no outlier residues recorded for this chain.

• Molecule 2: DNA (5'-D(P*GP*AP*GP*AP*GP*TP*CP*A)-3')

Chain d:

100%

There are no outlier residues recorded for this chain.
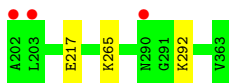
• Molecule 3: Intron-encoded endonuclease I-PpoI

Chain A:

·%

100%

A2 V163

• Molecule 3: Intron-encoded endonuclease I-PpoI

Chain B:

2%

98%   ·

A202 L203 E217 K265 N290 G291 K292 V363

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	114.06Å 114.06Å 88.02Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.33 – 1.64 37.33 – 1.64	Depositor EDS
% Data completeness (in resolution range)	99.3 (37.33-1.64) 99.3 (37.33-1.64)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 1.63Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.179 , 0.188 0.180 , 0.190	Depositor DCC
R_{free} test set	3930 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	17.3	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 39.7	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3876	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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, NA, GOL, ZN

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	C	0.88	0/311	1.09	1/475 (0.2%)
1	D	0.84	0/311	1.09	0/475
2	c	0.82	0/214	0.92	0/329
2	d	0.87	0/214	0.94	0/329
3	A	0.36	0/1293	0.58	0/1770
3	B	0.40	0/1293	0.61	0/1770
All	All	0.56	0/3636	0.76	1/5148 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	411	DT	O4'-C1'-N1	-5.29	104.29	108.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	
3	A	161/162 (99%)	161 (100%)	0	0	100	100
3	B	161/162 (99%)	161 (100%)	0	0	100	100
All	All	322/324 (99%)	322 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	
3	A	134/133 (101%)	134 (100%)	0	100	100
3	B	134/133 (101%)	131 (98%)	3 (2%)	52	25
All	All	268/266 (101%)	265 (99%)	3 (1%)	73	55

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	217	GLU
3	B	265	LYS
3	B	292	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 8 are monoatomic - leaving 5 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	#$|Z| > 2$
4	GOL	D	603	-	5,5,5	0.09	0	5,5,5	0.32	0
4	GOL	A	603	-	5,5,5	0.09	0	5,5,5	0.44	0
4	GOL	C	501	-	5,5,5	0.08	0	5,5,5	0.31	0
4	GOL	B	403	-	5,5,5	0.07	0	5,5,5	0.29	0
4	GOL	A	604	-	5,5,5	0.10	0	5,5,5	0.41	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
4	GOL	D	603	-	-	1/4/4/4	-
4	GOL	A	603	-	-	3/4/4/4	-
4	GOL	C	501	-	-	0/4/4/4	-
4	GOL	B	403	-	-	0/4/4/4	-
4	GOL	A	604	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	603	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	A	604	GOL	C1-C2-C3-O3
4	A	603	GOL	O1-C1-C2-O2
4	A	604	GOL	O2-C2-C3-O3
4	A	603	GOL	O2-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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(Å²)	Q<0.9
1	C	13/13 (100%)	-0.60	0	100	100	16, 22, 27, 28	0
1	D	13/13 (100%)	-0.43	0	100	100	17, 24, 30, 33	0
2	c	8/8 (100%)	-0.40	0	100	100	17, 28, 32, 32	0
2	d	8/8 (100%)	-0.49	0	100	100	16, 23, 27, 29	0
3	A	162/162 (100%)	-0.35	1 (0%)	89	89	13, 18, 28, 38	0
3	B	162/162 (100%)	-0.39	3 (1%)	66	67	14, 18, 36, 48	0
All	All	366/366 (100%)	-0.38	4 (1%)	80	81	13, 19, 32, 48	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	203	LEU	3.3
3	B	290	ASN	2.7
3	B	202	ALA	2.3
3	A	2	ALA	2.2

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 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(Å^2)	Q<0.9
4	GOL	A	604	6/6	0.78	0.28	37,40,41,43	0
4	GOL	C	501	6/6	0.81	0.15	23,23,31,41	0
4	GOL	D	603	6/6	0.88	0.15	22,26,32,38	0
4	GOL	A	603	6/6	0.89	0.11	18,27,37,38	0
4	GOL	B	403	6/6	0.96	0.06	17,18,19,25	0
5	MN	D	601	1/1	0.99	0.09	14,14,14,14	1
6	NA	D	602	1/1	0.99	0.09	14,14,14,14	1
7	ZN	A	602	1/1	0.99	0.07	15,15,15,15	0
5	MN	c	501	1/1	1.00	0.06	15,15,15,15	1
7	ZN	A	601	1/1	1.00	0.04	16,16,16,16	0
6	NA	c	502	1/1	1.00	0.06	15,15,15,15	1
7	ZN	B	401	1/1	1.00	0.05	16,16,16,16	0
7	ZN	B	402	1/1	1.00	0.06	14,14,14,14	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.
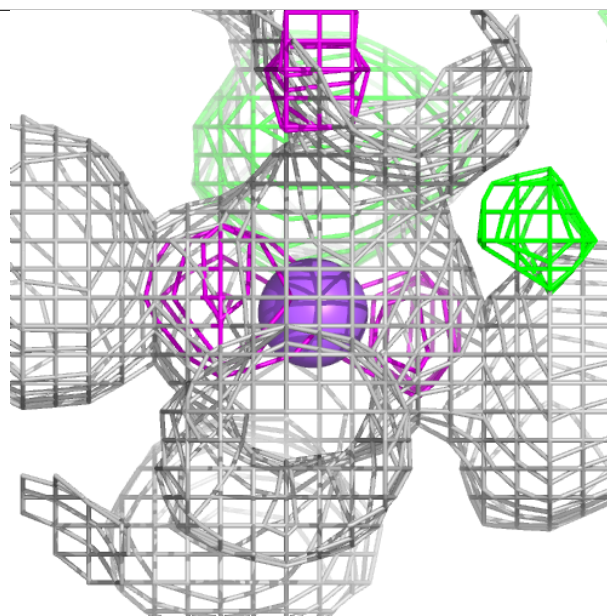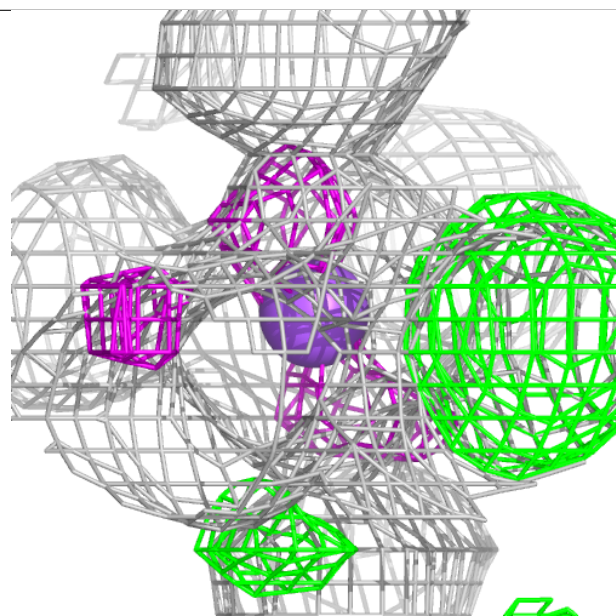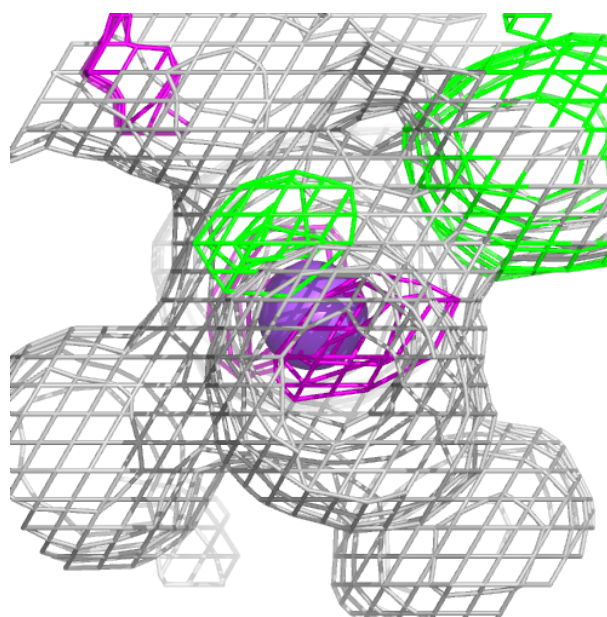
Electron density around MN D 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
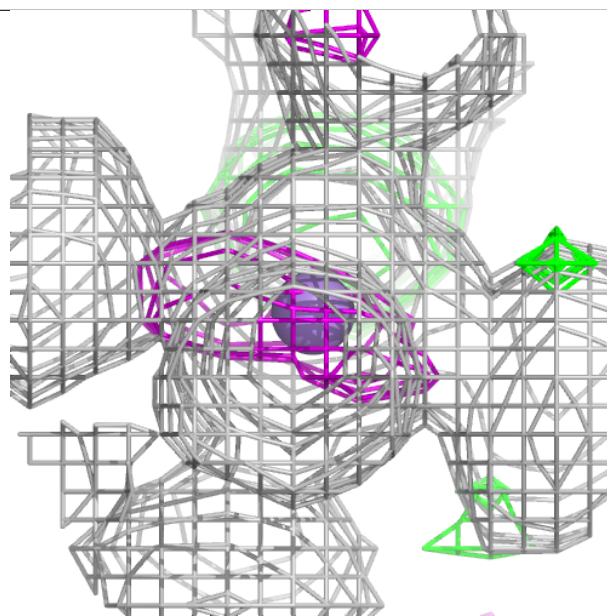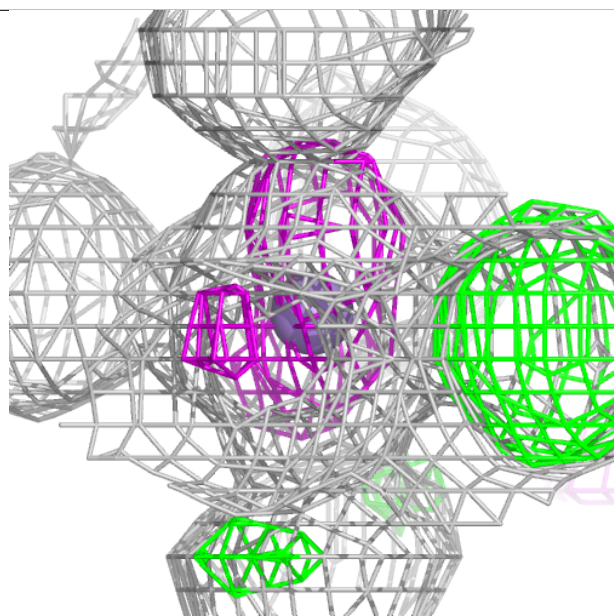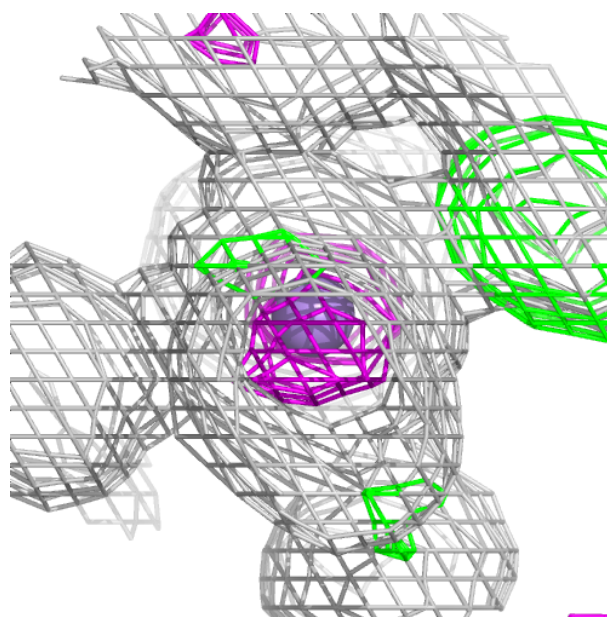mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
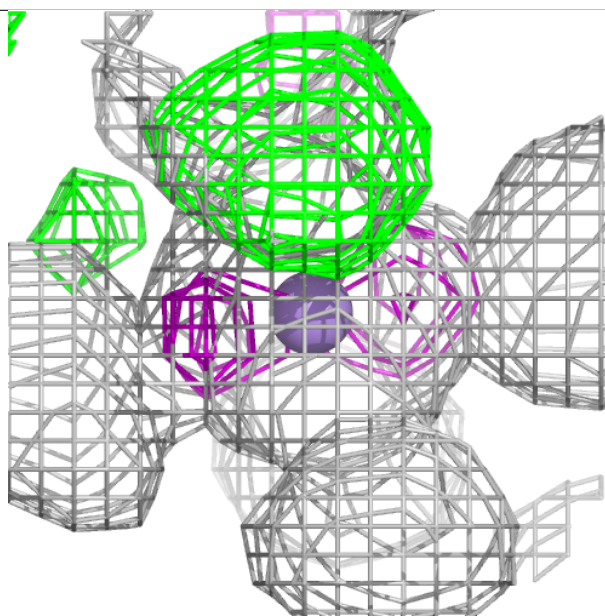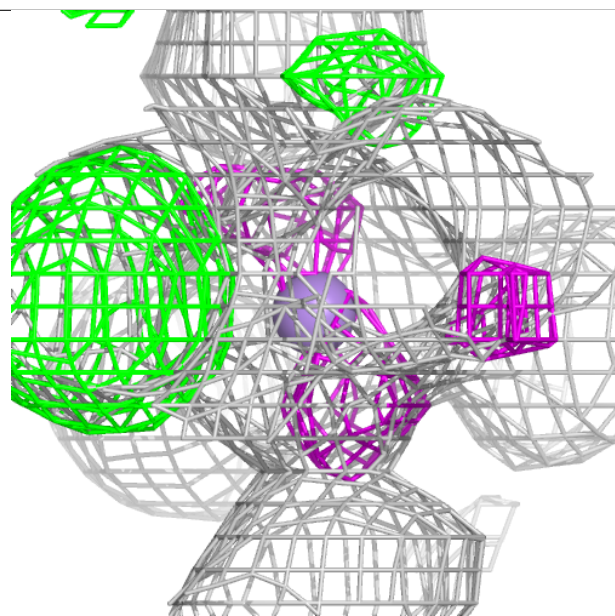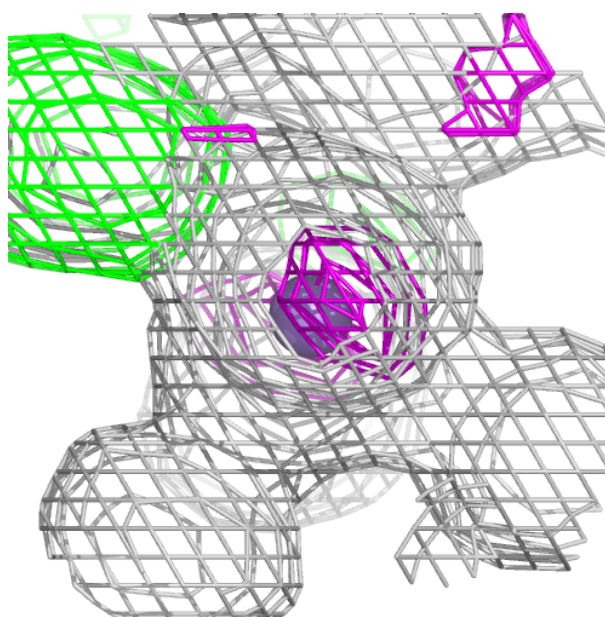
Electron density around NA D 602:

$2mF_o$-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 c 501:

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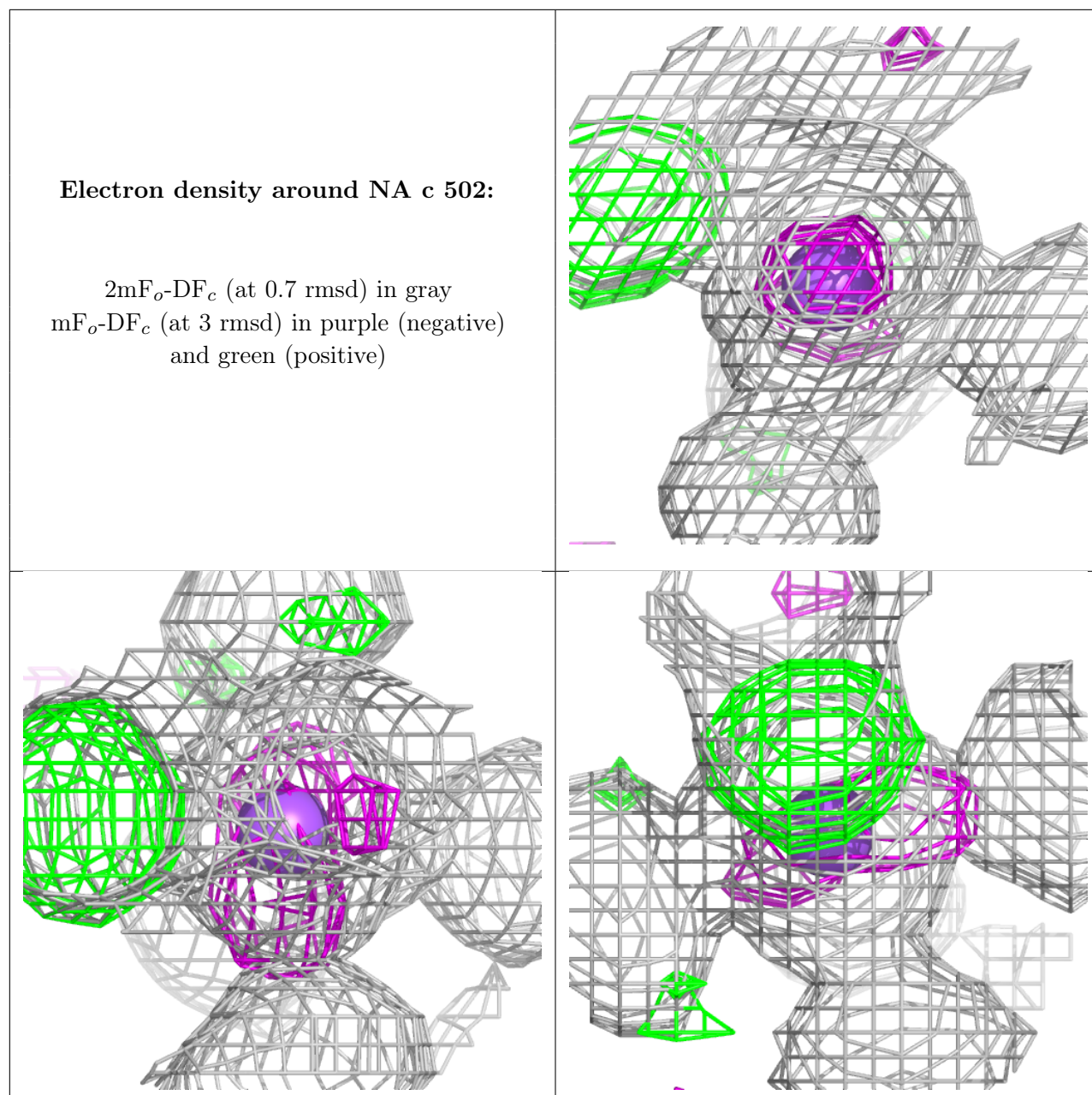and green (positive)

Electron density around NA c 502:

$2mF_o$-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.