Full wwPDB X-ray Structure Validation Report ⓘ

Jul 24, 2024 – 10:10 pm BST

PDB ID	:	9FXN
Title	:	Human mitochondrial nuclease EXOG (hEXOG)
Authors	:	Karasinska, J.M.; Szymanski, M.R.
Deposited on	:	2024-07-02
Resolution	:	1.62 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:

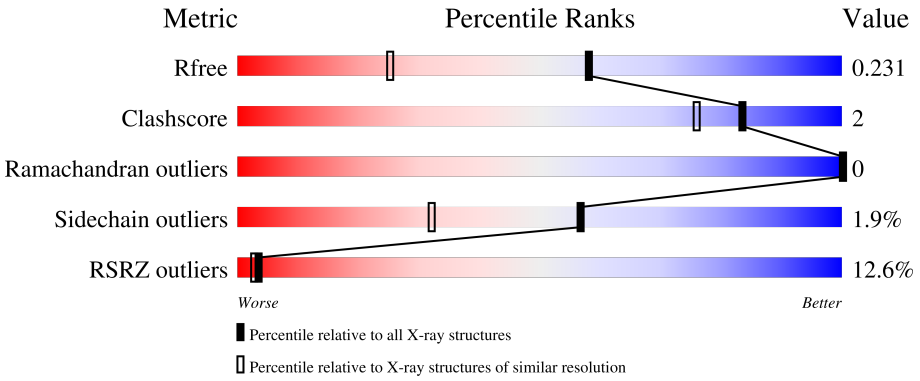
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.62 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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	317	14% 90% • 5%
1	B	317	10% 85% 9% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9840 atoms, of which 4751 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 Nuclease EXOG, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	300	Total	C	H	N	O	S	200	0	0
			4782	1535	2369	412	455	11			
1	B	301	Total	C	H	N	O	S	170	0	0
			4804	1541	2382	414	456	11			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	58	MET	-	initiating methionine	UNP Q9Y2C4
A	369	HIS	-	expression tag	UNP Q9Y2C4
A	370	HIS	-	expression tag	UNP Q9Y2C4
A	371	HIS	-	expression tag	UNP Q9Y2C4
A	372	HIS	-	expression tag	UNP Q9Y2C4
A	373	HIS	-	expression tag	UNP Q9Y2C4
A	374	HIS	-	expression tag	UNP Q9Y2C4
B	58	MET	-	initiating methionine	UNP Q9Y2C4
B	369	HIS	-	expression tag	UNP Q9Y2C4
B	370	HIS	-	expression tag	UNP Q9Y2C4
B	371	HIS	-	expression tag	UNP Q9Y2C4
B	372	HIS	-	expression tag	UNP Q9Y2C4
B	373	HIS	-	expression tag	UNP Q9Y2C4
B	374	HIS	-	expression tag	UNP Q9Y2C4

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		
2	B	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

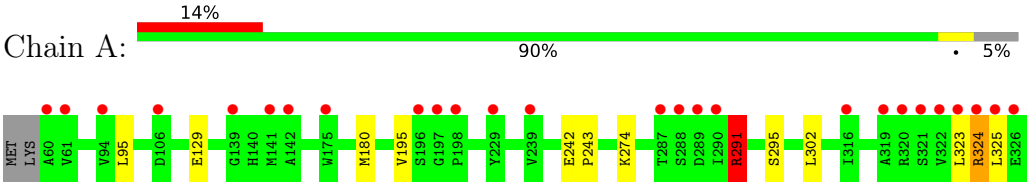
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	120	Total O 120 120	0	0
3	B	132	Total O 132 132	0	0

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.

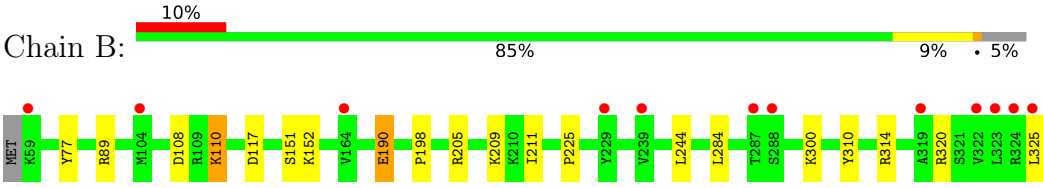
• Molecule 1: Nuclease EXOG, mitochondrial

Chain A:



• Molecule 1: Nuclease EXOG, mitochondrial

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	73.55Å 83.56Å 74.39Å 90.00° 113.51° 90.00°	Depositor
Resolution (Å)	40.55 – 1.62 40.55 – 1.62	Depositor EDS
% Data completeness (in resolution range)	93.1 (40.55-1.62) 93.1 (40.55-1.62)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.01 (at 1.62Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.157 , 0.228 0.159 , 0.231	Depositor DCC
R_{free} test set	2100 reflections (2.15%)	wwPDB-VP
Wilson B-factor (Å^2)	34.1	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 55.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.005 for l,-k,h	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	9840	wwPDB-VP
Average B, all atoms (Å^2)	54.0	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 25.16 % 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 3.3423e-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: 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	0.31	1/2470 (0.0%)	0.53	1/3341 (0.0%)
1	B	0.34	1/2479 (0.0%)	0.63	7/3352 (0.2%)
All	All	0.33	2/4949 (0.0%)	0.59	8/6693 (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	A	0	2
1	B	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	190	GLU	CB-CG	-6.24	1.40	1.52
1	A	291	ARG	CG-CD	5.26	1.65	1.51

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	190	GLU	OE1-CD-OE2	-14.73	105.62	123.30
1	B	333	LYS	CD-CE-NZ	7.77	129.57	111.70
1	A	324	ARG	CA-CB-CG	-7.07	97.84	113.40
1	B	333	LYS	CG-CD-CE	6.47	131.32	111.90
1	B	333	LYS	CB-CG-CD	-6.37	95.05	111.60
1	B	190	GLU	CG-CD-OE2	-5.46	107.38	118.30
1	B	190	GLU	CG-CD-OE1	5.17	128.64	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	333	LYS	CA-CB-CG	5.08	124.59	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	291	ARG	Sidechain
1	A	324	ARG	Sidechain
1	B	190	GLU	Sidechain

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	2413	2369	2369	10	0
1	B	2422	2382	2382	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	120	0	0	0	0
3	B	132	0	0	0	0
All	All	5089	4751	4751	23	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.

All (23) 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:243:PRO:HG2	1:A:291:ARG:NE	2.17	0.59
1:A:95:LEU:HD13	1:A:195:VAL:HG22	1.88	0.56
1:A:180:MET:HE1	1:A:302:LEU:HD12	1.90	0.53
1:B:77:TYR:OH	1:B:117:ASP:OD2	2.16	0.51
1:B:310:TYR:CZ	1:B:314:ARG:HD2	2.46	0.50
1:B:351:LEU:O	1:B:355:LYS:HG2	2.11	0.50
1:B:205:ARG:NH2	1:B:209:LYS:HD2	2.30	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:108:ASP:OD1	1:B:110:LYS:HG2	2.16	0.46
1:B:205:ARG:HH21	1:B:209:LYS:HD2	1.81	0.45
1:B:332:LEU:O	1:B:335:ALA:HB3	2.16	0.45
1:A:274:LYS:HG3	1:B:89:ARG:HH22	1.81	0.45
1:B:244:LEU:HD13	1:B:284:LEU:HD11	1.97	0.45
1:A:274:LYS:HG3	1:B:89:ARG:NH2	2.31	0.45
1:B:205:ARG:HD3	1:B:211:ILE:HG13	1.98	0.44
1:A:243:PRO:HG2	1:A:291:ARG:HE	1.81	0.44
1:A:243:PRO:HB2	1:A:291:ARG:HD3	1.99	0.43
1:A:242:GLU:HG3	1:A:243:PRO:HD2	1.99	0.43
1:B:339:PRO:HB3	1:B:343:PHE:CD1	2.54	0.42
1:A:325:LEU:HG	1:A:347:TYR:CE2	2.55	0.42
1:A:346:ARG:HA	1:A:346:ARG:HD2	1.86	0.42
1:B:198:PRO:HB2	1:B:225:PRO:HB3	2.03	0.41
1:B:334:ASN:C	1:B:336:GLU:H	2.23	0.41
1:B:325:LEU:HG	1:B:347:TYR:CE2	2.57	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	298/317 (94%)	296 (99%)	2 (1%)	0	100	100
1	B	299/317 (94%)	296 (99%)	3 (1%)	0	100	100
All	All	597/634 (94%)	592 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	266/282 (94%)	261 (98%)	5 (2%)	57 32
1	B	267/282 (95%)	262 (98%)	5 (2%)	57 32
All	All	533/564 (94%)	523 (98%)	10 (2%)	57 32

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

Mol	Chain	Res	Type
1	A	129	GLU
1	A	291	ARG
1	A	295	SER
1	A	323	LEU
1	A	349	LYS
1	B	110	LYS
1	B	151	SER
1	B	152	LYS
1	B	300	LYS
1	B	320	ARG

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

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

Of 2 ligands modelled in this entry, 2 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

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	300/317 (94%)	0.65	44 (14%)	2	1	30, 44, 81, 96	300 (100%)
1	B	301/317 (94%)	0.48	32 (10%)	6	5	31, 45, 71, 96	300 (99%)
All	All	601/634 (94%)	0.56	76 (12%)	3	3	30, 45, 75, 96	600 (99%)

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	60	ALA	8.7
1	A	323	LEU	7.3
1	A	334	ASN	6.5
1	A	335	ALA	6.2
1	A	328	ILE	6.1
1	B	323	LEU	5.8
1	A	329	MET	5.7
1	A	325	LEU	5.7
1	A	290	ILE	5.6
1	A	324	ARG	5.4
1	A	337	ILE	5.1
1	B	333	LYS	5.1
1	A	332	LEU	5.0
1	A	333	LYS	4.8
1	A	326	GLU	4.7
1	A	322	VAL	4.2
1	B	322	VAL	4.2
1	A	330	GLU	4.2
1	A	327	LYS	4.1
1	A	289	ASP	4.0
1	A	331	ASN	3.7
1	B	330	GLU	3.7
1	A	287	THR	3.6
1	B	359	GLN	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	347	TYR	3.5
1	B	287	THR	3.4
1	B	324	ARG	3.4
1	B	326	GLU	3.4
1	A	336	GLU	3.3
1	B	347	TYR	3.2
1	B	328	ILE	3.2
1	B	337	ILE	3.2
1	B	327	LYS	3.2
1	B	343	PHE	3.2
1	A	106	ASP	3.0
1	B	345	SER	3.0
1	B	325	LEU	3.0
1	B	338	GLU	2.9
1	A	338	GLU	2.8
1	A	343	PHE	2.8
1	A	288	SER	2.8
1	A	319	ALA	2.7
1	B	104	MET	2.7
1	A	351	LEU	2.7
1	B	351	LEU	2.7
1	B	341	ASP	2.7
1	B	329	MET	2.6
1	B	348	GLU	2.6
1	A	139	GLY	2.6
1	A	229	TYR	2.5
1	B	331	ASN	2.5
1	B	344	MET	2.5
1	A	198	PRO	2.5
1	B	288	SER	2.5
1	A	197	GLY	2.4
1	B	164	VAL	2.4
1	B	346	ARG	2.4
1	A	239	VAL	2.3
1	A	316	ILE	2.3
1	B	319	ALA	2.3
1	A	142	ALA	2.2
1	A	141	MET	2.2
1	A	94	VAL	2.2
1	B	332	LEU	2.2
1	A	359	GLN	2.2
1	A	196	SER	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	321	SER	2.2
1	A	348	GLU	2.2
1	B	229	TYR	2.1
1	B	334	ASN	2.1
1	A	320	ARG	2.1
1	A	175	TRP	2.1
1	A	349	LYS	2.1
1	B	59	LYS	2.1
1	A	61	VAL	2.1
1	B	239	VAL	2.1

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 monosaccharides in this entry.

6.4 Ligands (i)

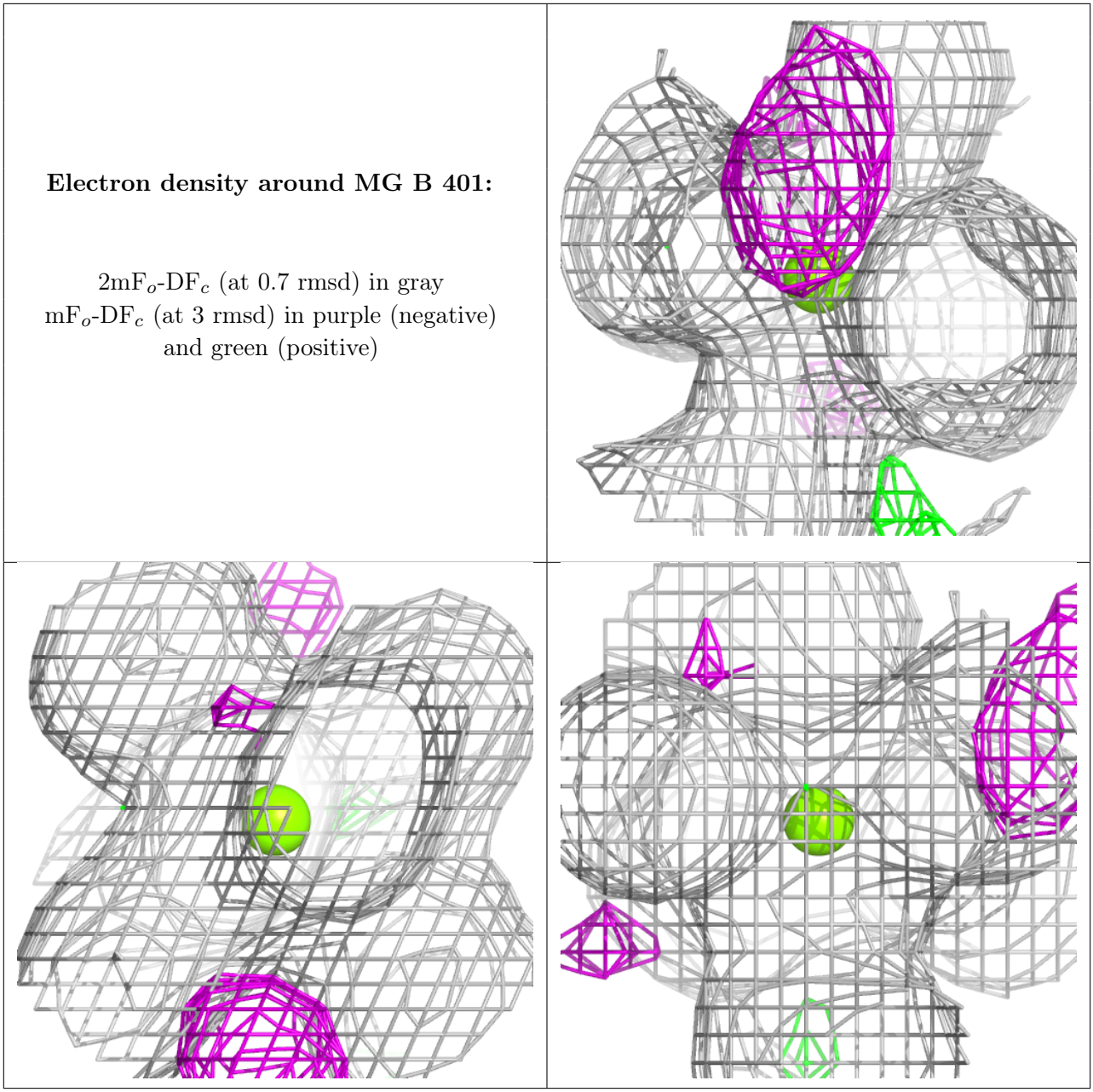
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	MG	B	401	1/1	0.99	0.09	27,27,27,27	1
2	MG	A	401	1/1	1.00	0.07	33,33,33,33	1

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.

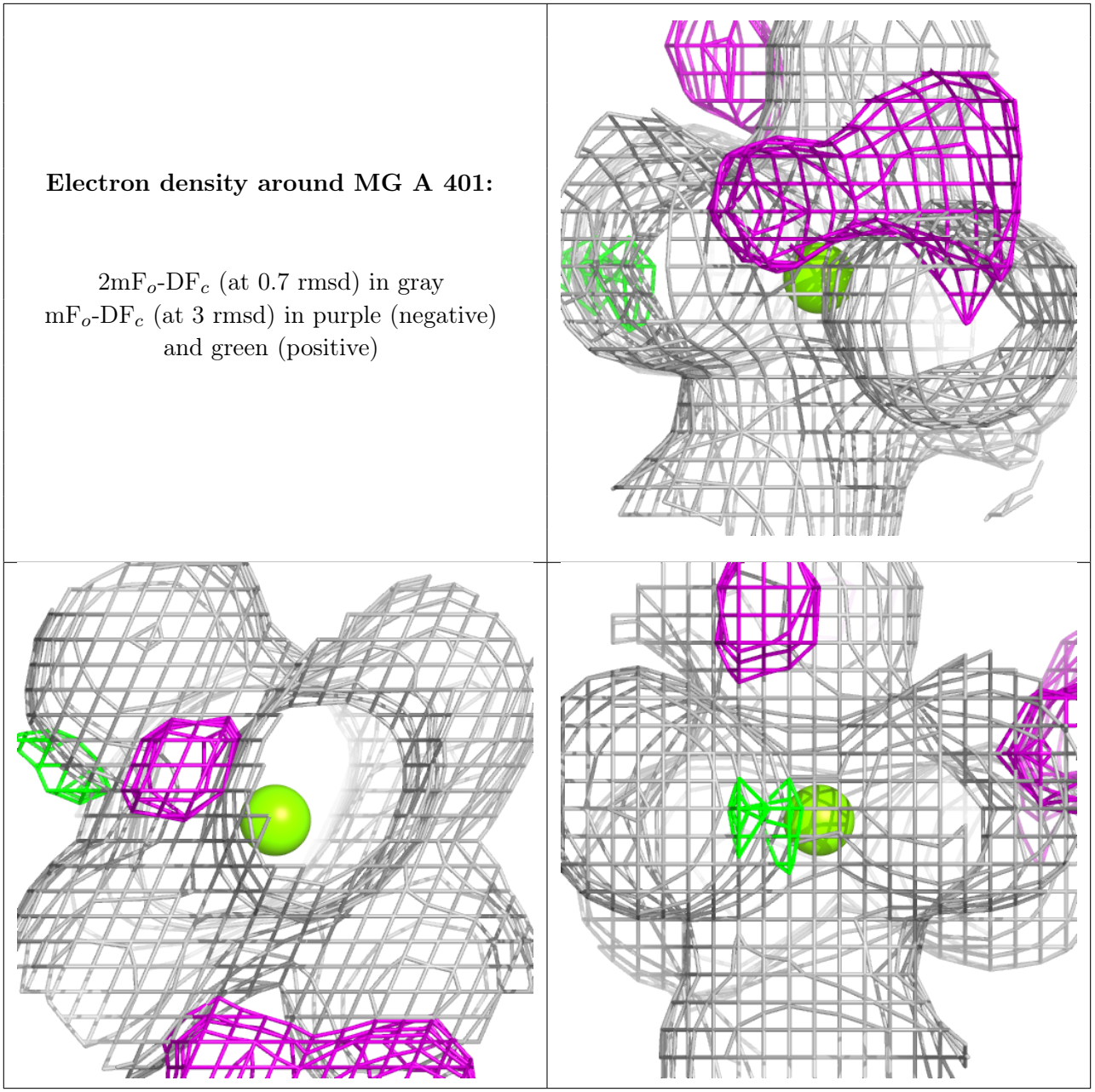
Electron density around MG B 401:

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



Electron density around MG 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)



6.5 Other polymers

There are no such residues in this entry.