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 ⓘ

Nov 21, 2023 – 03:11 AM JST

PDB ID	:	7EK7
Title	:	prawn ferritin to coordinate with heavy metal ions
Authors	:	Wang, Y.; Zang, J.
Deposited on	:	2021-04-04
Resolution	:	2.70 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

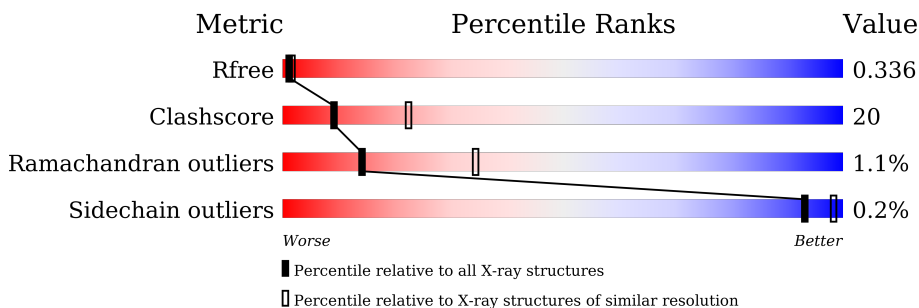|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.70 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		0.336
Clashscore		20
Ramachandran outliers		1.1%
Sidechain outliers		0.2%

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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	169	50% 48% •
1	B	169	58% 41% •
1	C	169	63% 37% •
1	D	169	59% 40% •
1	E	169	57% 43% •
1	F	169	61% 37% •

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8354 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 Ferritin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	169	Total	C	N	O	S	0	0	0
			1361	853	230	271	7			
1	F	169	Total	C	N	O	S	3	2	0
			1370	860	231	272	7			
1	A	169	Total	C	N	O	S	0	1	0
			1368	858	232	271	7			
1	B	169	Total	C	N	O	S	0	1	0
			1367	858	231	271	7			
1	D	169	Total	C	N	O	S	0	6	0
			1391	875	233	276	7			
1	E	169	Total	C	N	O	S	0	1	0
			1367	857	230	273	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	89	ARG	GLN	engineered mutation	UNP T2B7E1
F	89	ARG	GLN	engineered mutation	UNP T2B7E1
A	89	ARG	GLN	engineered mutation	UNP T2B7E1
B	89	ARG	GLN	engineered mutation	UNP T2B7E1
D	89	ARG	GLN	engineered mutation	UNP T2B7E1
E	89	ARG	GLN	engineered mutation	UNP T2B7E1

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	Hg	0	0
			1	1		
2	A	1	Total	Hg	0	0
			1	1		
2	D	1	Total	Hg	0	0
			1	1		

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	1	Total 1	Hg 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	23	Total 23	O 23	0	0
3	F	21	Total 21	O 21	0	0
3	A	24	Total 24	O 24	0	0
3	B	21	Total 21	O 21	0	0
3	D	17	Total 17	O 17	0	0
3	E	20	Total 20	O 20	0	0

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. 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.
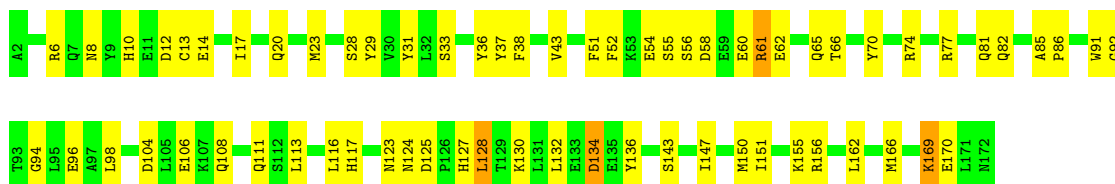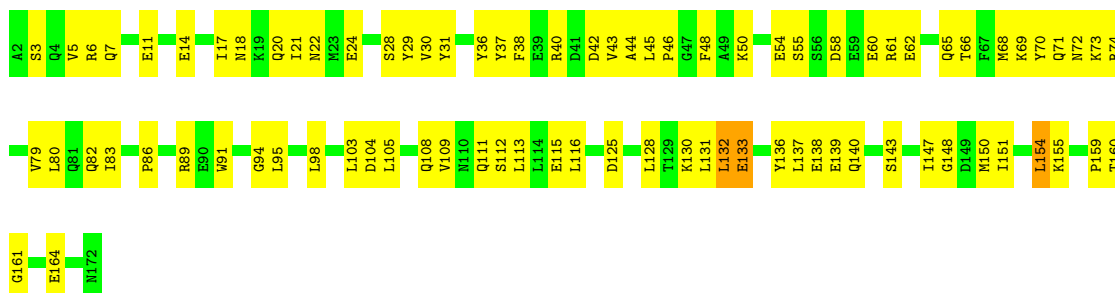
- Molecule 1: Ferritin

Chain C:
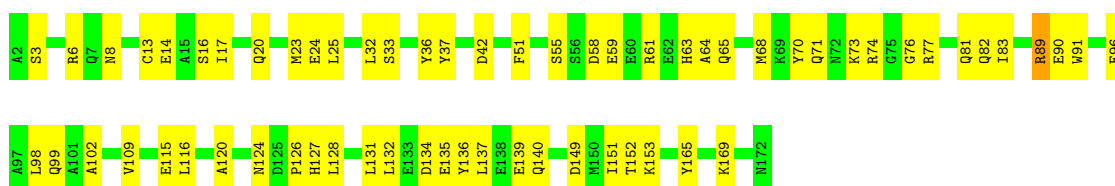
- Molecule 1: Ferritin

Chain F:
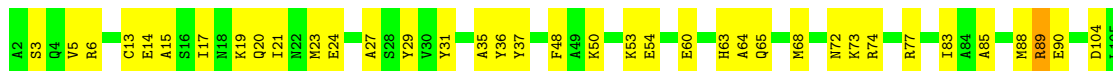
- Molecule 1: Ferritin

Chain A:

- Molecule 1: Ferritin

Chain B:

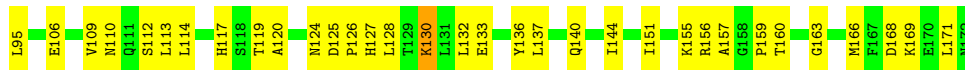Q108 V109 N110 Q111 S112 L113 L114 E115 L116 H117 S118 T119 N123 N124 D125 L128 T129 K130 L131 L132 E135 Y136 L137 E138 E139 Q140 V141 D142 K145 K146 D149 M150 K153 R156 A157 G158 G161 L162 G163 N172

• Molecule 1: Ferritin

Chain D:

59%   40%   ·

A2 S3 Q4 V5 R6 C13 E14 A15 S16 I17 N18 K19 Q20 I21 N22 M23 E24 A27 S28 Y29 V30 Y31 A35 Y36 Y37 F48 A49 K50 K53 E54 E60 H63 A64 Q65 M68 N72 K73 R74 R77 I83 A84 A85 M88 R89 E90 D104 L105

E106 K107 Q108 V109 N110 Q111 S112 L113 L114 E115 L116 A120 H127 L128 T129 K130 L131 L132 E133 D134 E135 Y136 L137 Q140 V141 D142 S143 I144 K145 K146 M150 L154 G158 G161 L162 G163 K169 E170 L171 N172

• Molecule 1: Ferritin

Chain E:

57%   43%   ·

A2 S3 Q4 Q7 N8 Y9 H10 E11 D12 C13 N18 K19 Q20 I21 N22 Y29 V30 Y31 M34 A35 Y36 Y37 F38 E39 R40 D41 D42 V43 A44 L45 F48 F52 S56 Q65 T66 Y70 Q71 N72 K73 R74 G75 G76 R77 I78 Q81 Q82 R89 E90

L95 E106 V109 N110 Q111 S112 L113 L114 H117 S118 T119 A120 N124 D125 P126 H127 L128 T129 K130 L131 L132 E133 Y136 L137 Q140 I144 I151 K155 R156 A157 G158 P159 T160 G163 M166 F167 D168 K169 E170 L171 N172

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	126.28Å 126.28Å 177.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.02 – 2.70 30.02 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.02-2.70) 99.9 (30.02-2.70)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 2.72Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.254 , 0.336 0.254 , 0.336	Depositor DCC
R_{free} test set	2056 reflections (5.42%)	wwPDB-VP
Wilson B-factor (Å^2)	50.2	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 11.3	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.39, < L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.116 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k 0.115 for -1/2*h+1/2*k+1/2*l,1/2*h-1/2*k+1/2*l,h+k 0.097 for -1/2*h-1/2*k+1/2*l,-1/2*h-1/2*k-1/2*l,h-k 0.096 for -1/2*h-1/2*k-1/2*l,-1/2*h-1/2*k+1/2*l,-h+k 0.447 for -h,k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8354	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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: HG

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	0.51	0/1396	0.76	3/1876 (0.2%)
1	B	0.52	0/1394	0.68	1/1872 (0.1%)
1	C	0.49	0/1385	0.63	0/1861
1	D	0.58	3/1433 (0.2%)	0.77	3/1923 (0.2%)
1	E	0.52	0/1394	0.70	2/1873 (0.1%)
1	F	0.59	2/1400 (0.1%)	0.71	3/1880 (0.2%)
All	All	0.54	5/8402 (0.1%)	0.71	12/11285 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	13	CYS	CB-SG	10.60	2.00	1.82
1	D	145	LYS	CD-CE	-6.96	1.33	1.51
1	D	145	LYS	CB-CG	-5.67	1.37	1.52
1	D	145	LYS	CG-CD	-5.32	1.34	1.52
1	F	169	LYS	CE-NZ	5.02	1.61	1.49

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	145	LYS	CD-CE-NZ	-14.61	78.09	111.70
1	D	145	LYS	CB-CG-CD	-8.36	89.88	111.60
1	D	107	LYS	CD-CE-NZ	-7.03	95.53	111.70
1	B	130	LYS	CD-CE-NZ	6.97	127.74	111.70
1	E	130	LYS	CD-CE-NZ	-6.48	96.81	111.70

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	1368	0	1322	75	3
1	B	1367	0	1329	64	0
1	C	1361	0	1315	48	1
1	D	1391	0	1365	61	3
1	E	1367	0	1320	57	1
1	F	1370	0	1334	60	1
2	A	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
3	A	24	0	0	6	0
3	B	21	0	0	4	0
3	C	23	0	0	5	0
3	D	17	0	0	5	0
3	E	20	0	0	5	0
3	F	21	0	0	8	0
All	All	8354	0	7985	323	5

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 20.

The worst 5 of 323 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:13:CYS:SG	1:B:123:ASN:ND2	2.27	1.05
1:D:53:LYS:NZ	3:D:301:HOH:O	1.94	1.00
1:A:138:GLU:OE2	3:A:301:HOH:O	1.81	0.97
1:A:154:LEU:HB2	3:A:302:HOH:O	1.75	0.85
1:F:123:ASN:ND2	3:F:201:HOH:O	2.06	0.85

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:7:GLN:OE1	1:D:107:LYS:NZ[6_444]	1.85	0.35
1:A:72:ASN:O	1:D:145:LYS:NZ[6_444]	1.99	0.21
1:A:130:LYS:NZ	1:D:133:GLU:O[6_444]	2.02	0.18
1:E:43:VAL:O	1:E:156:ARG:NH2[3_455]	2.02	0.18
1:C:134:ASP:OD1	1:F:130:LYS:NZ[3_555]	2.10	0.10

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	
1	A	168/169 (99%)	153 (91%)	13 (8%)	2 (1%)	13	32
1	B	168/169 (99%)	155 (92%)	10 (6%)	3 (2%)	8	21
1	C	167/169 (99%)	156 (93%)	9 (5%)	2 (1%)	13	32
1	D	173/169 (102%)	163 (94%)	9 (5%)	1 (1%)	25	50
1	E	168/169 (99%)	154 (92%)	12 (7%)	2 (1%)	13	32
1	F	169/169 (100%)	160 (95%)	8 (5%)	1 (1%)	25	50
All	All	1013/1014 (100%)	941 (93%)	61 (6%)	11 (1%)	14	34

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	171	LEU
1	C	115	GLU
1	F	134	ASP
1	A	132	LEU
1	B	138	GLU

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	
1	A	146/145 (101%)	146 (100%)	0	100	100
1	B	146/145 (101%)	146 (100%)	0	100	100
1	C	145/145 (100%)	144 (99%)	1 (1%)	84	94
1	D	151/145 (104%)	150 (99%)	1 (1%)	84	94
1	E	146/145 (101%)	146 (100%)	0	100	100
1	F	147/145 (101%)	147 (100%)	0	100	100
All	All	881/870 (101%)	879 (100%)	2 (0%)	93	98

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

Mol	Chain	Res	Type
1	C	89	ARG
1	D	89	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	110	ASN
1	E	4	GLN
1	E	140	GLN
1	A	82	GLN
1	B	71	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

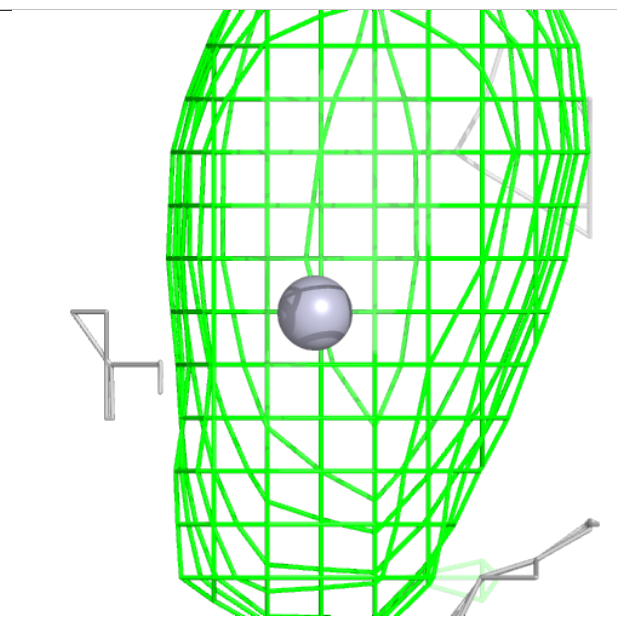# 6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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.
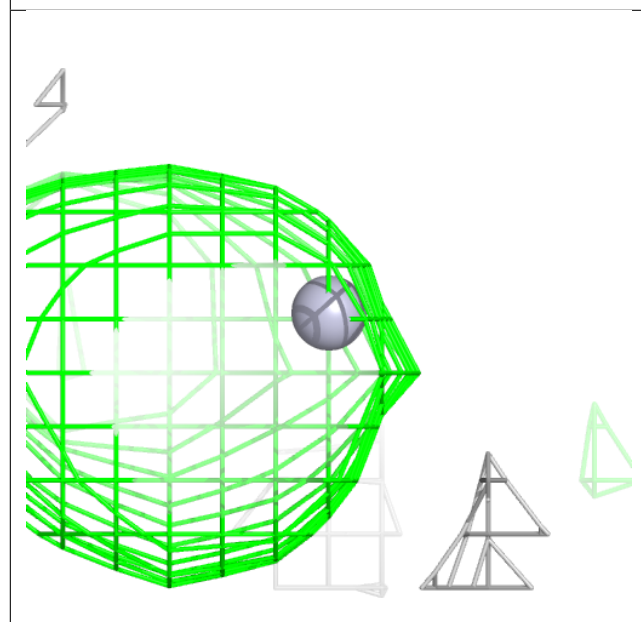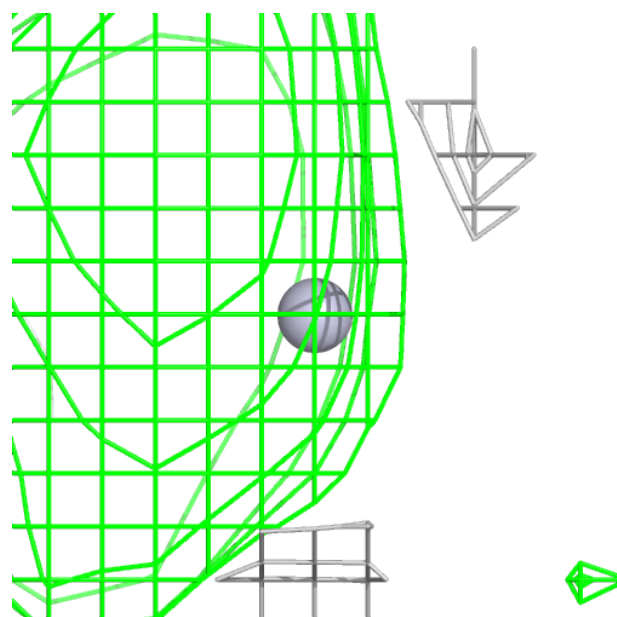
6.4 Ligands ⓘ

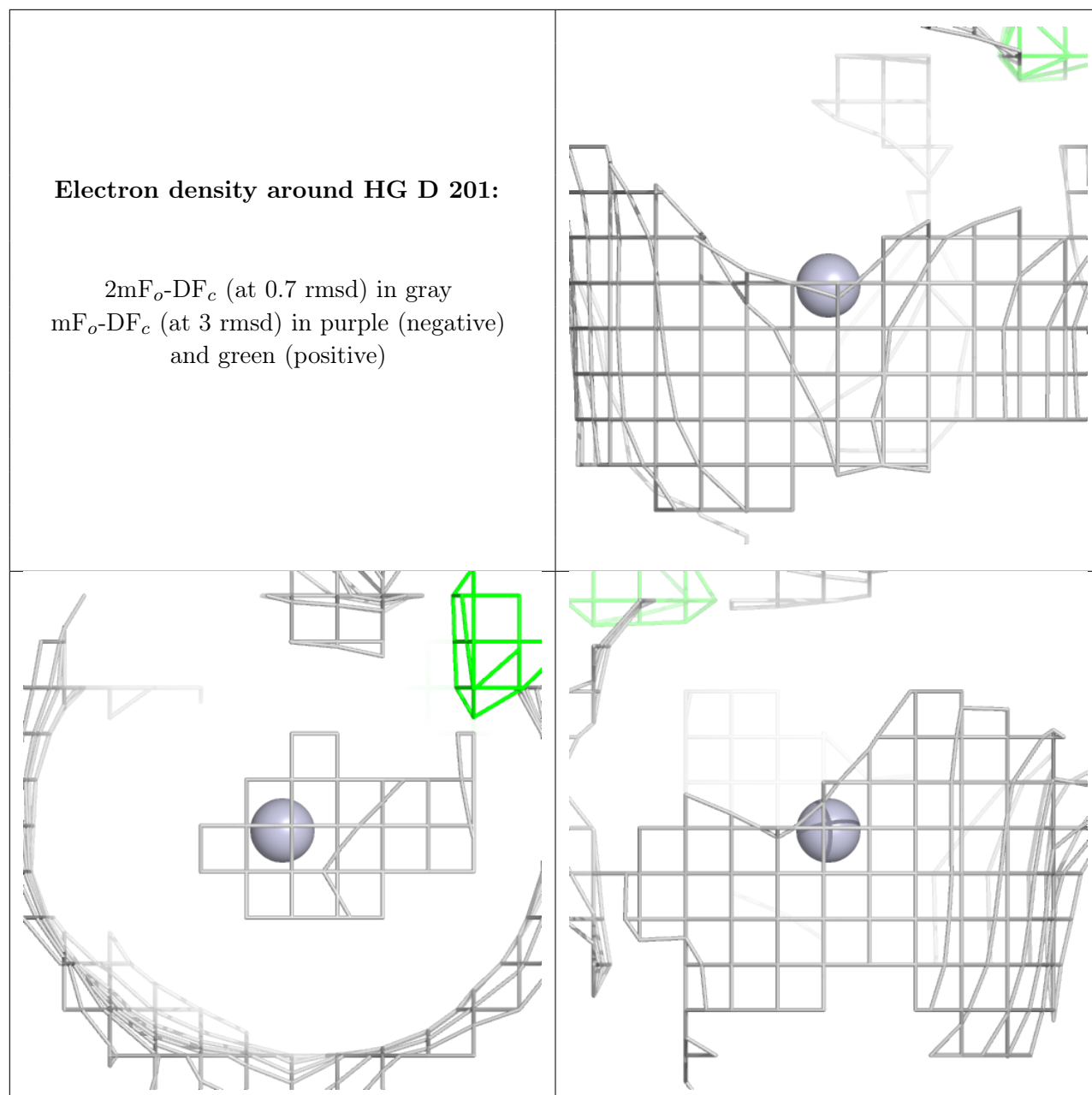Unable to reproduce the depositors R factor - this section is therefore empty.

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.
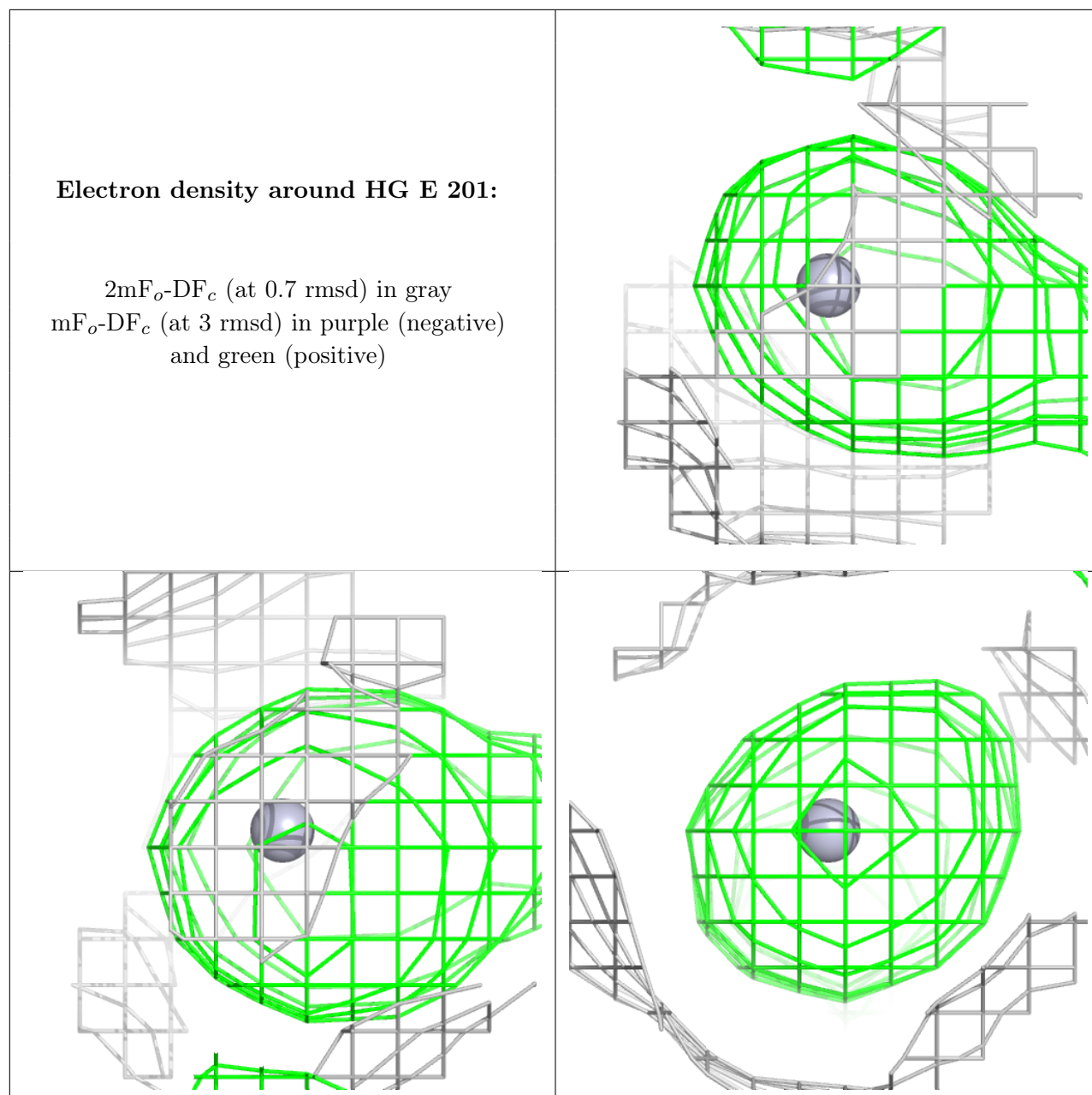
Electron density around HG C 201:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
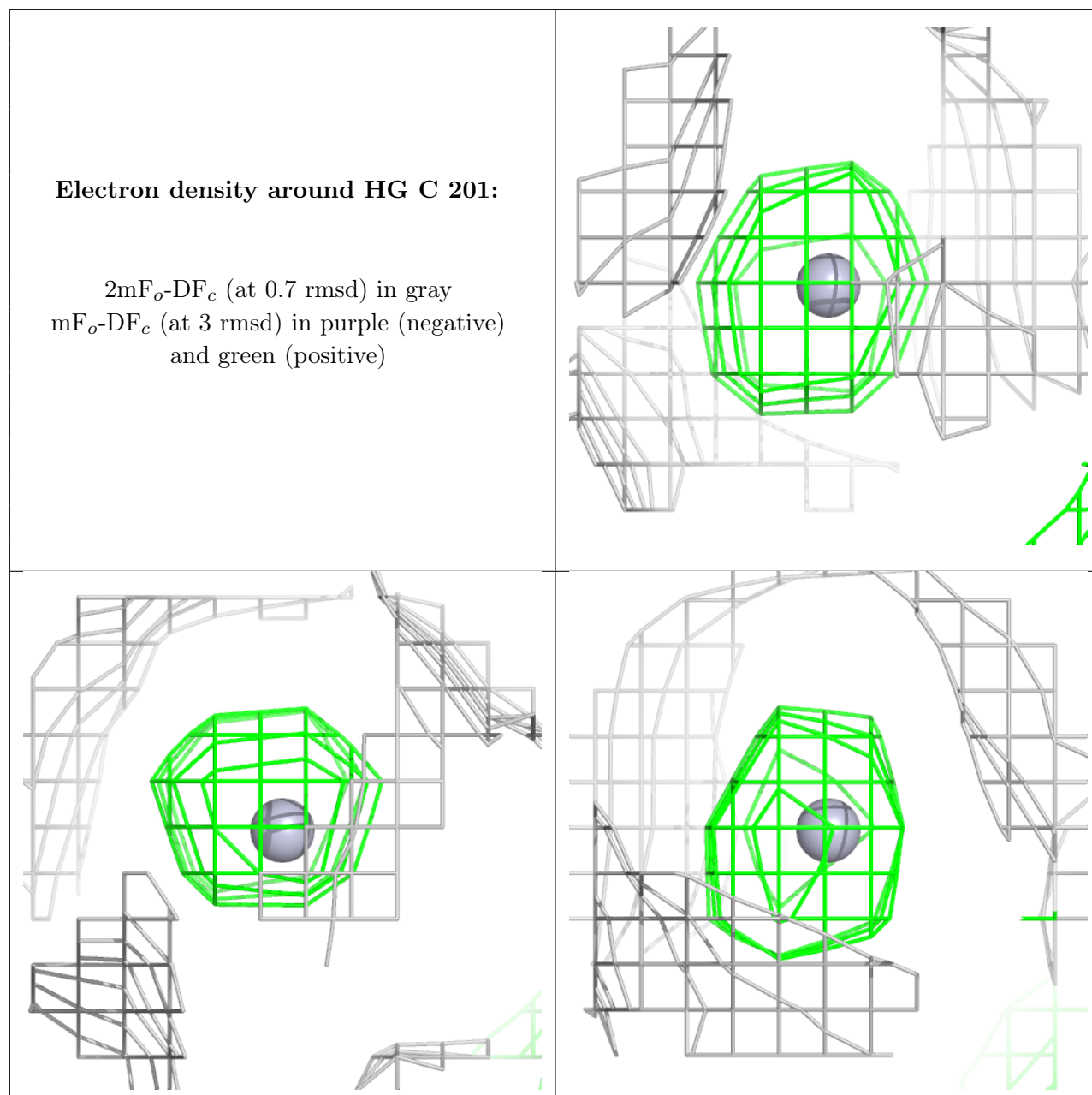
Electron density around HG A 201:

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

Electron density around HG D 201:

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 HG E 201:

$2mF_o$-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 ⓘ

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