wwPDB X-ray Structure Validation Summary Report (i)

Sep 2, 2024 – 07:17 pm BST

PDB ID	:	8OSI
Title	:	Genetically encoded green ratiometric calcium indicator FNCaMP in calcium-bound state
Authors	:	Varfolomeeva, L.A.; Boyko, K.M.; Nikolaeva, A.Y.; Subach, O.M.; Subach, F.V.
Deposited on	:	2023-04-19
Resolution	:	2.42 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 (i) 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 (i)) were used in the production of this report:

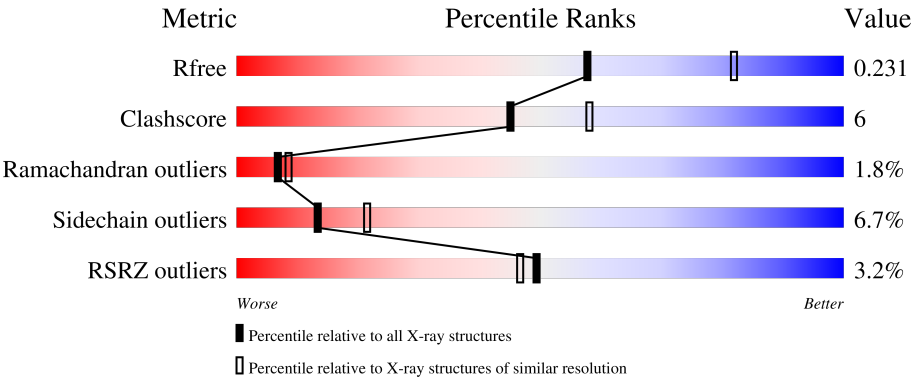
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

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 2.42 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}	164625	5670 (2.44-2.40)
Clashscore	180529	6299 (2.44-2.40)
Ramachandran outliers	177936	6232 (2.44-2.40)
Sidechain outliers	177891	6233 (2.44-2.40)
RSRZ outliers	164620	5670 (2.44-2.40)

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	461	3% 70% 12% 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3251 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 mNeonGreen,Calmodulin,Contig An16c0100, genomic contig.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	401	Total	C	N	O	S	0	0	0
			3198	2003	532	639	24			

There are 79 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-41	MET	-	initiating methionine	UNP A0A1S4NYF2
A	-40	GLY	-	expression tag	UNP A0A1S4NYF2
A	-39	GLY	-	expression tag	UNP A0A1S4NYF2
A	-38	SER	-	expression tag	UNP A0A1S4NYF2
A	-37	HIS	-	expression tag	UNP A0A1S4NYF2
A	-36	HIS	-	expression tag	UNP A0A1S4NYF2
A	-35	HIS	-	expression tag	UNP A0A1S4NYF2
A	-34	HIS	-	expression tag	UNP A0A1S4NYF2
A	-33	HIS	-	expression tag	UNP A0A1S4NYF2
A	-32	HIS	-	expression tag	UNP A0A1S4NYF2
A	-31	GLY	-	expression tag	UNP A0A1S4NYF2
A	-10	GLU	-	linker	UNP A0A1S4NYF2
A	-9	ASN	-	linker	UNP A0A1S4NYF2
A	-8	LEU	-	linker	UNP A0A1S4NYF2
A	-7	TYR	-	linker	UNP A0A1S4NYF2
A	-6	PHE	-	linker	UNP A0A1S4NYF2
A	-5	GLN	-	linker	UNP A0A1S4NYF2
A	-4	GLY	-	linker	UNP A0A1S4NYF2
A	-3	HIS	-	linker	UNP A0A1S4NYF2
A	-2	MET	-	linker	UNP A0A1S4NYF2
A	-1	ARG	-	linker	UNP A0A1S4NYF2
A	0	SER	-	linker	UNP A0A1S4NYF2
A	8	GLU	ASP	conflict	UNP A0A1S4NYF2
A	28	ILE	VAL	conflict	UNP A0A1S4NYF2
A	53	MET	LYS	conflict	UNP A0A1S4NYF2
A	68	CR2	GLY	chromophore	UNP A0A1S4NYF2
A	68	CR2	TYR	chromophore	UNP A0A1S4NYF2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	68	CR2	GLY	chromophore	UNP A0A1S4NYF2
A	107	THR	SER	conflict	UNP A0A1S4NYF2
A	128	GLU	LYS	conflict	UNP A0A1S4NYF2
A	146	GLU	-	linker	UNP A0A1S4NYF2
A	147	ALA	-	linker	UNP A0A1S4NYF2
A	148	GLN	-	linker	UNP A0A1S4NYF2
A	164	LEU	ALA	conflict	UNP A2QJG6
A	170	LEU	LYS	conflict	UNP A2QJG6
A	189	ASP	GLY	conflict	UNP A2QJG6
A	190	LEU	GLN	conflict	UNP A2QJG6
A	207	GLY	ASP	conflict	UNP A2QJG6
A	208	GLY	ASN	conflict	UNP A2QJG6
A	209	ASP	ASN	conflict	UNP A2QJG6
A	222	THR	ALA	conflict	UNP A2QJG6
A	224	GLU	LYS	conflict	UNP A2QJG6
A	227	TYR	ASP	conflict	UNP A2QJG6
A	228	ARG	THR	conflict	UNP A2QJG6
A	230	THR	SER	conflict	UNP A2QJG6
A	237	LEU	ALA	conflict	UNP A2QJG6
A	238	CYS	PHE	conflict	UNP A2QJG6
A	246	ASP	ASN	conflict	UNP A2QJG6
A	250	VAL	SER	conflict	UNP A2QJG6
A	257	ALA	VAL	conflict	UNP A2QJG6
A	264	GLU	LYS	conflict	UNP A2QJG6
A	294	LYS	MET	conflict	UNP A2QJG6
A	298	GLY	-	linker	UNP A2QJG6
A	299	GLY	-	linker	UNP A2QJG6
A	300	GLY	-	linker	UNP A2QJG6
A	301	GLY	-	linker	UNP A2QJG6
A	302	SER	-	linker	UNP A2QJG6
A	303	GLY	-	linker	UNP A2QJG6
A	304	GLY	-	linker	UNP A2QJG6
A	305	SER	-	linker	UNP A2QJG6
A	306	GLY	-	linker	UNP A2QJG6
A	307	MET	ALA	conflict	UNP A2R7C5
A	310	LEU	THR	conflict	UNP A2R7C5
A	312	LYS	HIS	conflict	UNP A2R7C5
A	323	ARG	LYS	conflict	UNP A2R7C5
A	328	MET	-	linker	UNP A2R7C5
A	329	TYR	-	linker	UNP A2R7C5
A	330	PHE	-	linker	UNP A2R7C5
A	335	VAL	ARG	conflict	UNP A0A1S4NYF2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	340	CYS	TYR	conflict	UNP A0A1S4NYF2
A	342	ASP	ASN	conflict	UNP A0A1S4NYF2
A	347	VAL	ILE	conflict	UNP A0A1S4NYF2
A	350	LEU	PHE	conflict	UNP A0A1S4NYF2
A	353	ALA	SER	conflict	UNP A0A1S4NYF2
A	354	PHE	TYR	conflict	UNP A0A1S4NYF2
A	355	ILE	THR	conflict	UNP A0A1S4NYF2
A	357	ASP	GLY	conflict	UNP A0A1S4NYF2
A	395	ILE	LYS	conflict	UNP A0A1S4NYF2
A	413	GLY	VAL	conflict	UNP A0A1S4NYF2

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total Ca 4 4	0	0

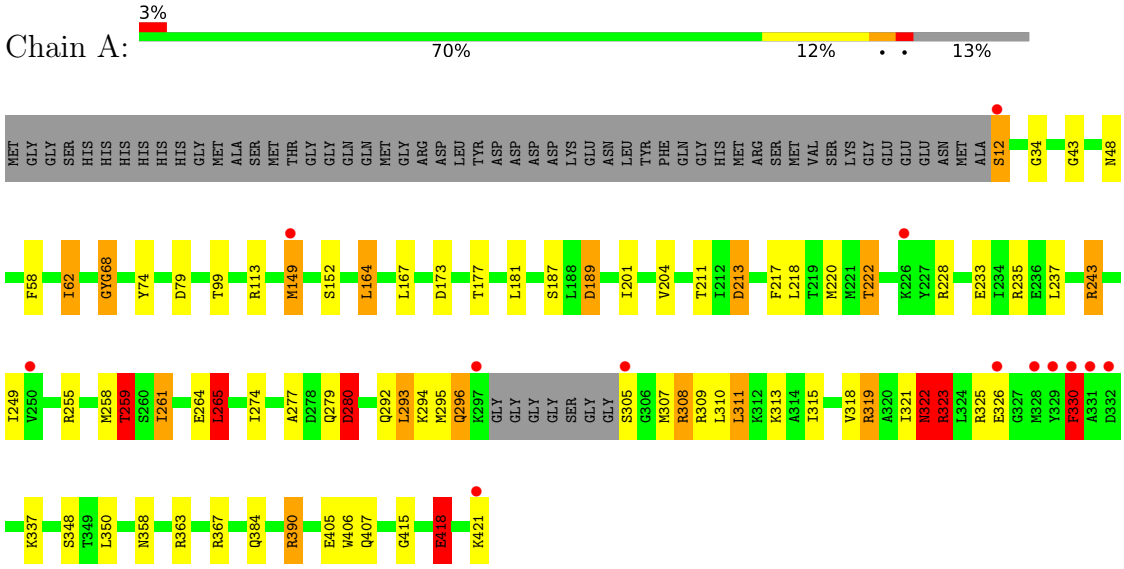
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	49	Total O 49 49	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: mNeonGreen,Calmodulin,Contig An16c0100, genomic contig



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	151.40Å 151.40Å 61.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.95 – 2.42 47.95 – 2.42	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.95-2.42) 99.4 (47.95-2.42)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.196 , 0.231 0.196 , 0.231	Depositor DCC
R_{free} test set	1539 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	63.6	Xtriage
Anisotropy	0.376	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3251	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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: CA, CR2

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	0.72	1/3241 (0.0%)	1.36	32/4371 (0.7%)

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	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	213	ASP	CG-OD1	5.75	1.38	1.25

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	322	ASN	N-CA-CB	8.79	126.43	110.60
1	A	384	GLN	CB-CA-C	-8.49	93.42	110.40
1	A	177	THR	N-CA-CB	7.59	124.73	110.30
1	A	319	ARG	NE-CZ-NH2	-7.45	116.57	120.30
1	A	177	THR	CB-CA-C	-7.27	91.96	111.60

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	243	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	305	SER	Peptide
1	A	308	ARG	Sidechain
1	A	309	ARG	Sidechain
1	A	319	ARG	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	3198	0	3040	36	0
2	A	4	0	0	0	0
3	A	49	0	0	0	0
All	All	3251	0	3040	36	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 6.

The worst 5 of 36 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:321:ILE:O	1:A:322:ASN:HB2	1.77	0.84
1:A:415:GLY:HA2	1:A:418:GLU:HG3	1.61	0.82
1:A:296:GLN:HG3	1:A:308:ARG:HH21	1.47	0.79
1:A:321:ILE:O	1:A:322:ASN:CB	2.36	0.70
1:A:58:PHE:CD1	1:A:62:ILE:HD11	2.26	0.70

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	394/461 (86%)	373 (95%)	14 (4%)	7 (2%)	7 8

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	280	ASP
1	A	296	GLN
1	A	322	ASN
1	A	330	PHE
1	A	149	MET

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	344/390 (88%)	321 (93%)	23 (7%)	13 21

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	261	ILE
1	A	293	LEU
1	A	274	ILE
1	A	295	MET
1	A	187	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	35	GLN
1	A	175	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	$\#\|Z\| > 2$
1	CR2	A	68	1	20,20,21	0.99	1 (5%)	25,27,29	2.56	10 (40%)

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
1	CR2	A	68	1	-	0/6/25/26	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	68	CR2	C1-N3	2.35	1.40	1.37

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	68	CR2	O2-C2-CA2	-5.58	127.83	130.96
1	A	68	CR2	C1-CA1-N1	-5.15	101.45	112.85
1	A	68	CR2	CA2-C2-N3	4.62	105.56	103.37
1	A	68	CR2	CA1-C1-N2	-4.04	118.86	124.28
1	A	68	CR2	CG2-CB2-CA2	3.48	134.20	129.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	68	CR2	3	0

5.5 Carbohydrates (i)

There are no oligosaccharides 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

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	400/461 (86%)	-0.11	13 (3%) 49 47	41, 63, 110, 138	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	329	TYR	3.9
1	A	331	ALA	3.8
1	A	330	PHE	3.8
1	A	250	VAL	2.9
1	A	226	LYS	2.6

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

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
1	CR2	A	68	19/20	0.98	0.09	37,52,62,65	0

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,

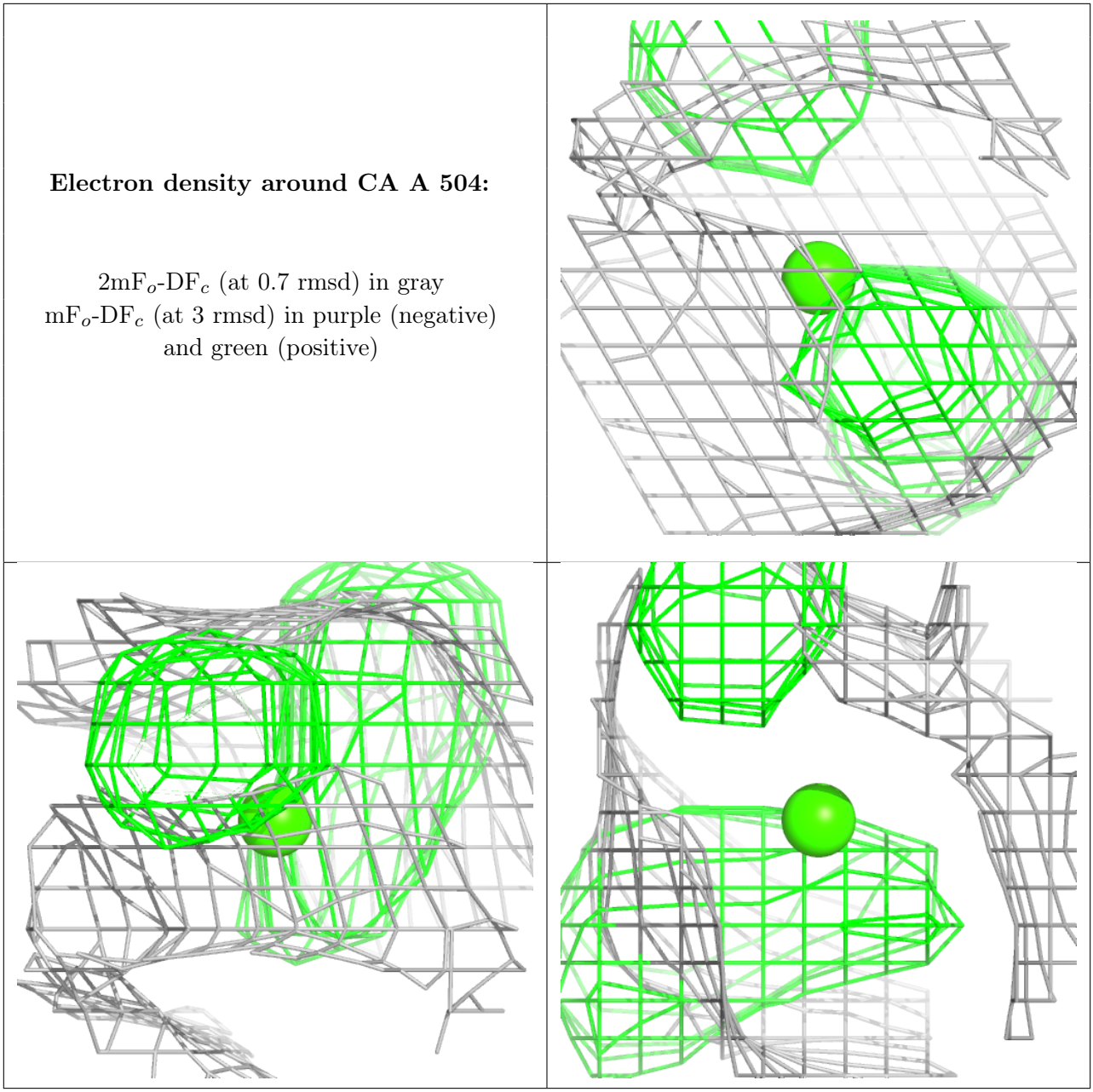
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CA	A	504	1/1	0.96	0.09	74,74,74,74	0
2	CA	A	502	1/1	0.99	0.04	98,98,98,98	0
2	CA	A	503	1/1	0.99	0.04	70,70,70,70	0
2	CA	A	501	1/1	0.99	0.04	69,69,69,69	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.

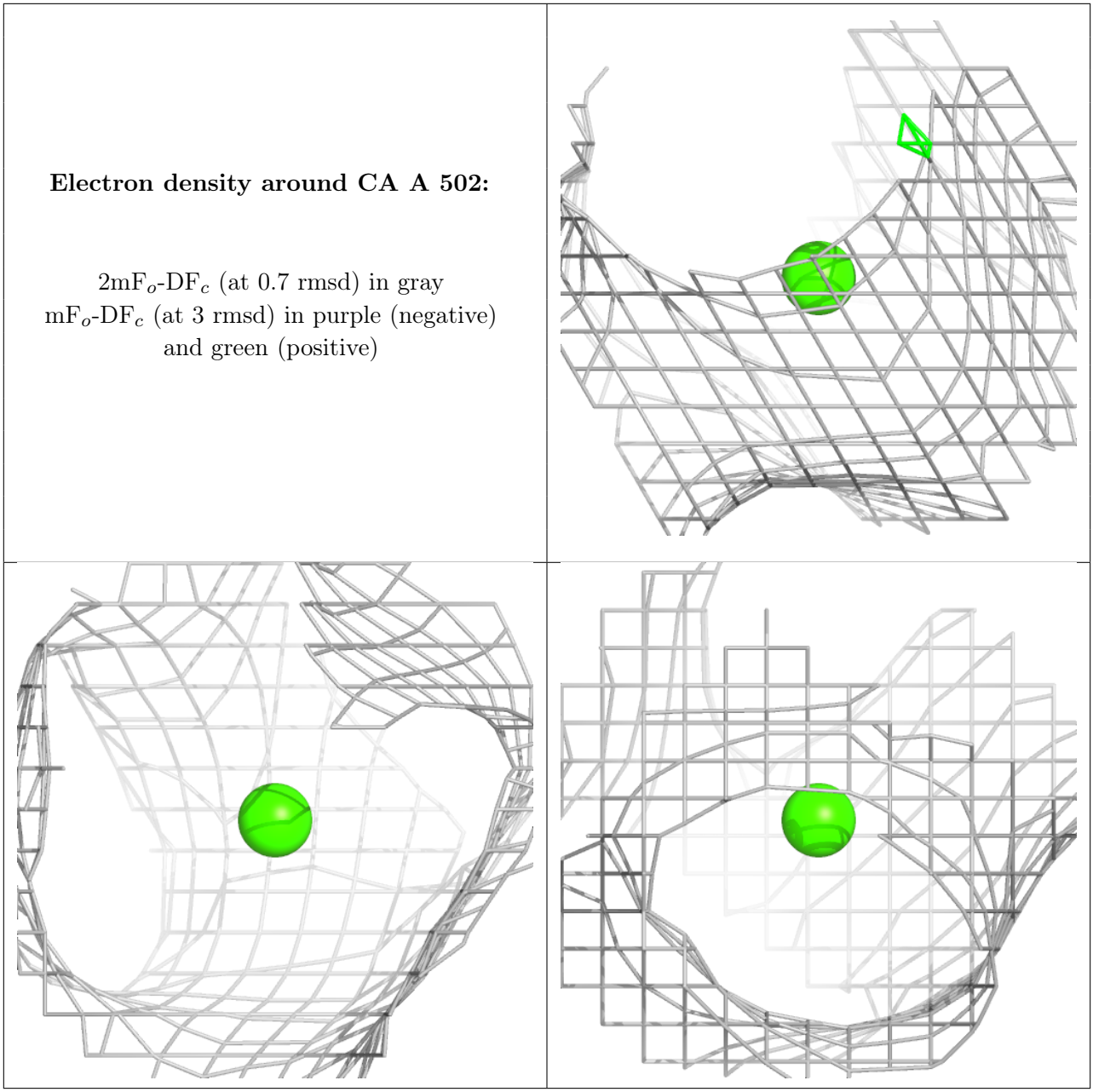
Electron density around CA A 504:

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



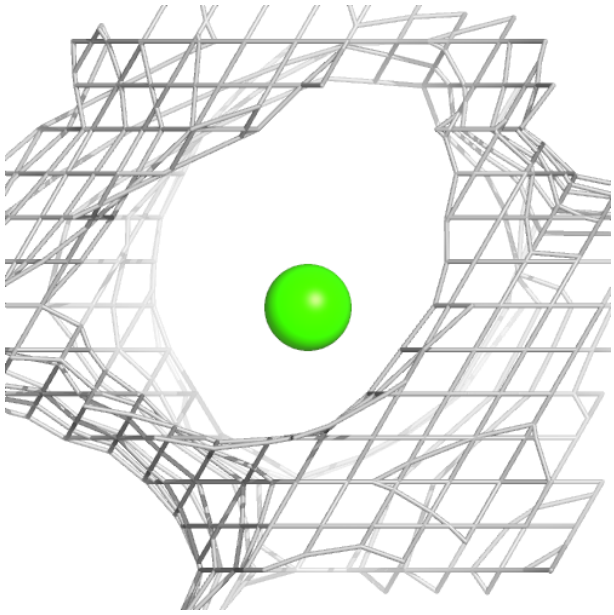
Electron density around CA A 502:

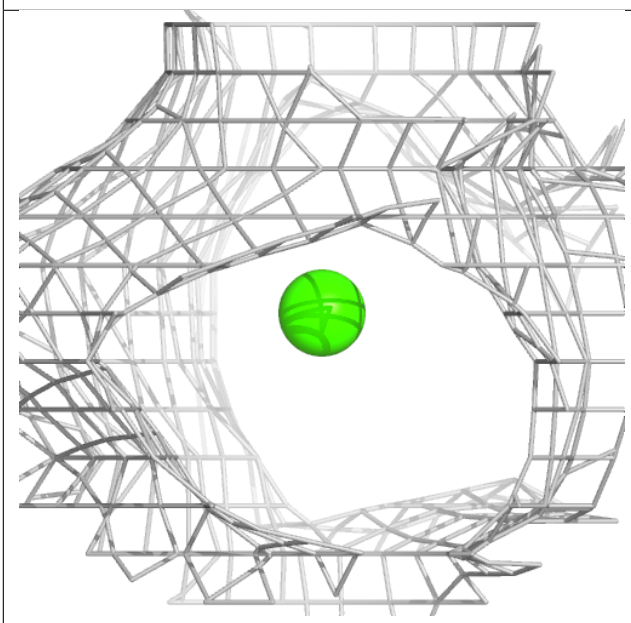
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

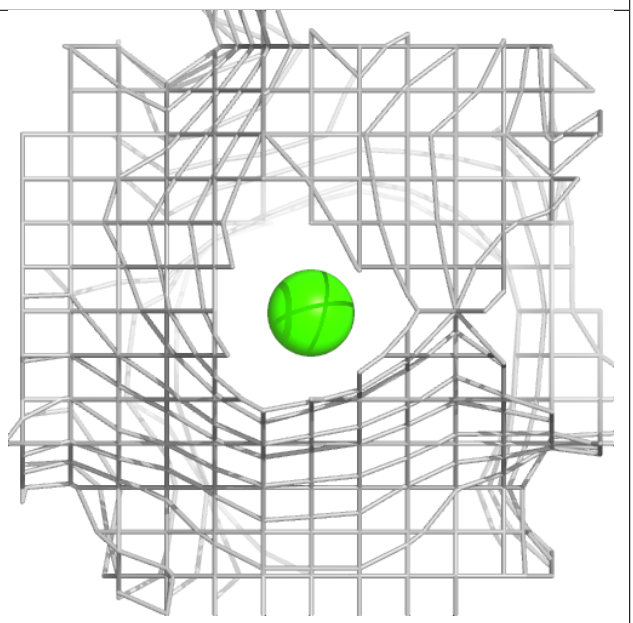


Electron density around CA A 503:

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

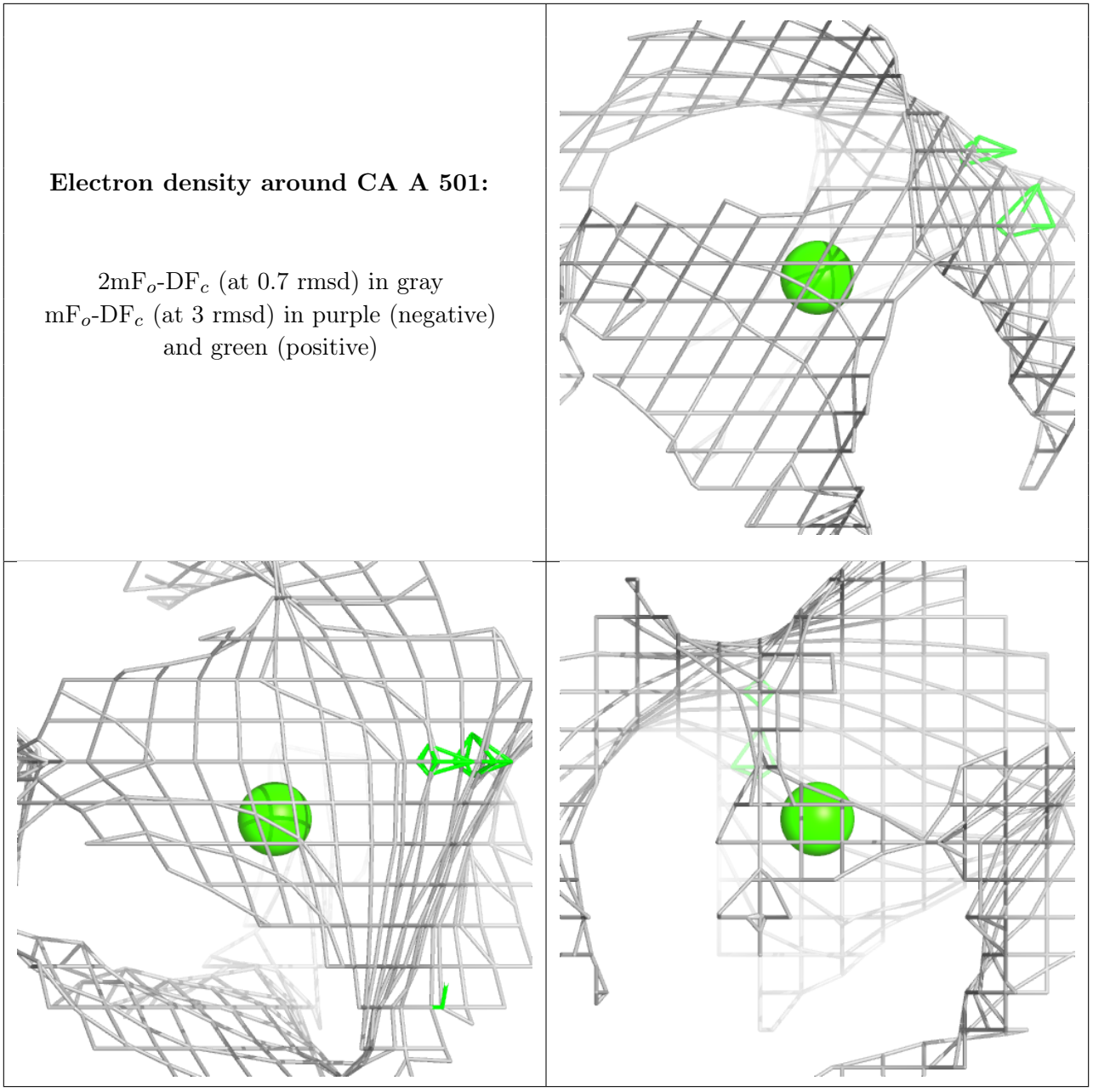






Electron density around CA A 501:

$2mF_o$-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.